Temperature and vernalisation regulation of seed properties in *Brassica napus*

Rebecca Ellen Doherty

A thesis presented for the degree of Doctor of Philosophy

University of East Anglia, UK

John Innes Centre, UK

December 2022

Table of Contents

Abstract

As the global climate changes and temperatures increase in coming decades, crop yields are likely to be threatened. This includes in the UK, where warmer winters and summer heatwaves are expected to become more common. To ensure food production remains stable, it will be necessary to breed crops which can withstand these changing temperatures.

In *Brassica napus*, seed properties including seed size, the number of seeds produced, and percentage of seed oil are all factors contributing to the yield. It is known that yield stability in *Brassica napus* is impacted by temperature during seed development and vernalisation. The effects of temperature on seed properties have been seen in *Arabidopsis thaliana*. Studies in *Brassica napus* have examined the genetic basis of yield separately to investigating how yield changes under different temperatures. However, the genetic basis of differences in how yield is affected by temperature has not yet been explored in-depth in *Brassica napus*.

This thesis uses phenotypic approaches, transcriptome profiling and associative transcriptomics to describe the effect of vernalisation temperature, as well as temperature during seed set and seed maturation, on seed traits and gene expression in *Brassica napus*. This work reports a decrease in seed weight and seed number per pod associated with increasing seed maturation temperature from 18°C to 24°C, as well as with increasing vernalisation temperature from 5°C to 15°C. These results indicate that warmer temperatures lead to reductions in yield-related traits, particularly in Winter OSR. The decrease in mature seed size of Winter OSR resulting from an increased vernalisation temperature of 15°C were found to arise during a developmental window between 14 and 21 days after pollination. Associated changes in gene expression were described, including reduced expression of *FLC* genes in seeds produced by plants vernalised at 15°C compared to 5°C.

List of Figures

List of Tables

Acknowledgements

I would first like to thank my supervisory team, who have been an endless well of knowledge and learning over the last four years. None of this thesis would exist without Steve Penfield, my supervisor, who has taught me an incredible amount over the course of this PhD. I am incredibly grateful for the opportunity to have studied in his lab at the John Innes Centre and will never forget his dedication to science and the example he has set as a group leader. I would also like to express my deepest appreciation to Rachel Wells; without her incredible expertise in organising the BRAVO phenotyping experiment I can hardly imagine what my thesis would have looked like. I would also of course like to thank my secondary supervisor, Lars Østergaard, whose regular feedback and optimism about my experiments was a welcome part of many meetings.

I would further like to thank the many members of the Penfield group, past and present, whom I met during my PhD. I have enjoyed my time working with every person who passed through the office and consider myself fortunate to have been able to learn from them. I would like to first thank Carmel O'Neill for her constant enthusiasm for both science and catching up over a coffee. I would also like to thank Lucy Thursfield, Laura Wilkinson and Sam Warner, for both their help in the lab and friendship outside of it. I am also extremely grateful to have gotten help and advice on many experiments from Xiaochao Chen, Thiago Barros-Galvão and Francesca Stefanato.

Many thanks also go to the entire BRAVO team, who I was incredibly lucky to collaborate with during my PhD. Particular thanks go to Catherine Chinoy, Emilie Knight and Mikhaela Neequaye, for their contributions in harvesting and phenotyping thousands of plants. Special thanks also go to Hugh Woolfenden, for his invaluable help in preparing transcriptome data for analysis as well as contributing R scripts which made my associative transcriptomics a much smoother experience.

A number of other collaborators and teams also contributed to the data I produced over the last four years. I would like to thank Peter Eastmond, Smita Kurup and Laura Siles-Suarez at Rothamsted Research for their help in analysing seed oil content, as well as Fiona Corke and John Doonan at IBERS for carrying out the seed germination

experiment. The work done by Horticultural Services, Genotyping Services and BRACT at JIC was also key to so much of the data I produced during my time here.

Last, but certainly not least, a massive thank you to my family and friends for being hugely supportive during my PhD. The friends I've made from the 2018 PhD student cohort have been an incredible source of support and advice, who have filled the last four years with laughter and kindness (and karaoke). Beyond academia, I cannot thank Sam enough for constantly making sure I remembered to be a person and not just a PhD candidate – and for reminding me every week that I could do this. Finally, my eternal thanks to Amy. She has believed in me enough for both of us and more through this whole process, and I cannot thank her enough for her endless support and patience over the last four years.

Chapter 1 - Introduction

1.1 General Introduction

1.1.1 Effects of climate change on crop yields

Increasing global temperatures, predicted over the next 50 years, create a high risk for decreases in food production (IPCC, 2014). The average global temperature today has already risen by at least 1°C compared to 1850-1900, and all areas of the UK are expected to be warmer by the end of the 21st century (Morice et al, 2020). In a high emissions scenario (represented by the UN IPCC RCP8.5 scenario), winter temperatures in central England could be up to 4.2°C warmer by the 2070s (UKCP18 Land Report).

In areas important for crop growth, predicted temperature changes will have an impact on production of crops. In the UK, changes in temperature are expected to disrupt patterns and timing of growth in wheat and negatively impact yield due to the risk of increased temperatures during flowering (Semenov, 2008). Climate estimates during the growing seasons of maize, wheat, rice, and soybeans in major regions for crop growth indicate an increase in temperature since 1980, having particularly negative impacts on the yield of maize and wheat (Lobell et al, 2011).

Changes in temperature have already been shown to have an impact on how and where farmers grow crops. In China, multiple cropping areas for maize, wheat and rice are moving northwards (Yang et al, 2015). In Europe, farmers are attempting to mitigate the effects of the changing climate on crop yield by changing the timing and variety of crops they grow (Olesen et al, 2011).

Without such adaptations, it has been suggested that production of wheat, rice and maize in temperate and tropical regions will be negatively affected by increases of

temperature of 2°C (Challinor et al, 2014). Climate change is also expected to reduce the range of areas suitable for cultivating winter annual crops, such as oilseed rape in Western Europe (Jaime et al, 2018).

Further warming is indicated to have a global negative impact on the yields of major crops (Rosenzweig et al, 2014; Zhao et al, 2017).

1.1.2 Abiotic stress and crop yield

Plants must develop physiological and genetic strategies for dealing with abiotic stresses such as high temperatures and drought. Understanding how plants respond to stresses and withstand them is a useful source of information for crop breeding, as it can enable plant breeders to create more resilient crops. It is vital to create crops which are able to withstand high temperatures in light of expected temperature increases due to climate change.

Crop yields are negatively impacted by abiotic stresses, including exposure to high temperatures. Multiple stresses can occur at the same time, such as heat and drought, leading to further negative yield impacts (Mittler, 2006). High temperatures are known to affect yields across a range of crops (Moriondo et al, 2011; Deryng et al, 2014). Specifically in oilseed rape high temperature stress is known to cause reductions in yield (Morrison and Stewart, 2002; Rashid et al, 2018; Huang et al, 2019). There is a need to adapt growth strategies for *Brassica napus* in order for yield to withstand the effects of climate change (reviewed in Ahmad et al, 2021).

Drought stress is another abiotic stress which can reduce crop yield. Drought stress is known to affect overall yield in OSR, with water shortages leading to a reduction in crop growth and yield (Wan et al, 2009; Khan et al, 2019). Establishment rates of oilseed rape in the field are important for subsequent plant growth and final yield (Finch-Savage and Bassel, 2016). In oilseed rape, high levels of drought are associated with reduced seedling establishment and survival as well as impacting rate of germination

(Yang et al, 2007). Reduced germination and vigour can also be seen when the seed matures under high temperatures (Chen et al, 2021).

Temperature stress can affect yield, as extreme heat scenarios decrease yield; these scenarios may become more common in crop growth regions due to the effects of climate change. However, temperature increases below the level of stressing the plant can still have an impact on yield. Reduced yields in *Brassica napus* can be associated with changes in temperature during winter (Brown et al, 2019; He et al., 2017). Small changes in overnight temperature can cause yield losses in rice and winter wheat (Peng et al, 2004; Shibo et al, 2010). Even when plants are grown below the threshold for extreme temperature stress caused by events such as heatwaves, changes in temperature within normal growth ranges can be associated with yield loss.

1.1.3 Brassica napus as a crop species

Brassica napus is an allotetraploid crop, arising from a hybridization of *Brassica oleracea* and *Brassica rapa* (Østergaard and King, 2008; Nagaharu, 1935). It has been cultivated and bred to grow with different life histories and under different climates. The *Brassica napus* crop has been cultivated into several crop types: winter oilseed rape (OSR) has an obligate vernalisation requirement and is sown in autumn prior to growth over winter before flowering in spring; spring OSR varieties have functionally no vernalisation requirement and can be sown and harvested in a single season; semi-winter OSR has a low vernalisation requirement and can be grown in climates with mild winters (Akter et al, 2021; Calderwood et al, 2021; Qiu et al, 2006; Ferreira et al, 1995).

Different varieties of *Brassica napus* are grown for different purposes, such as growing kale varieties for leaves or swede varieties for the tuberous vegetable. Alongside food production, the crop can also be grown as animal fodder. As most *Brassica napus* is grown as oilseed rape, there is a focus on seeds as a determinant of yield in order to produce rapeseed oil (Allender and King, 2010). The oil produced from rapeseed is also an important source of biodiesel for transportation fuel (Bušić et al, 2018).

Although it is an important crop in the UK, there are a number of issues with growing *Brassica napus*. There have been recent decreases in how much oilseed rape is grown along with a decreasing contribution to the UK economy because of these issues (DEFRA, 2021). The issues affecting oilseed rape growth include pests such as cabbage stem flea beetle, cabbage stem weevils, aphids, and pollen beetles (Dewar et al, 2016). Banning of neonicotinoid pesticides has led to increased pest burden on the crop (Nicholls, 2016). The inability to use neonicotinoids to treat OSR has also led to increased incidence of yellow turnip virus, which negatively affects OSR yield (Stevens et al, 2008; Hackenburg et al, 2020).

Alongside these biotic stresses, oilseed rape already has issues with yield reliability. Yield reliability is primarily caused through temperature. It has been found that cold temperatures during a specific window in winter (November 27 – Dec 21) are highly correlated with increased yield, such that an increase of 1°C can lead to a 113 kg/hectare yield loss (Brown et al, 2019).

1.2 Seed development

1.2.1 The stages of seed development in angiosperms

Brassica species are angiosperms. Angiosperms undergo double fertilization which leads to the development of a diploid zygote and a triploid endosperm (Chaudhury et al, 2001). The diploid zygote is formed when a sperm cell fuses with the egg cell. The resulting zygote goes on to establish the basic pattern of the embryo, which consists of a shoot meristem, cotyledons, the hypocotyl, roots, and root meristem (Li and Li, 2015). The embryo develops through several stages, from the early globular stage through heart, torpedo, bent cotyledon and finally the mature green embryo stage, before the seed matures and the embryo enters dormancy until germination.

A second sperm cell forms the triploid endosperm by fertilizing the diploid central cell (Berger, 1999). The endosperm starts off in a syncytial stage, undergoing growth by

many nuclear divisions before cellularization occurs (Olsen, 2004; Sundaresan et al, 2005). This is eventually followed by differentiation and death (Berger, 1999; Li and Li, 2015). In dicots, such as *Arabidopsis thaliana* and *Brassica napus,* endosperm growth is initially rapid before being consumed to allow growth of the embryo which then makes up the majority of the mature seed mass (Chaudhury et al, 2001; Li and Li, 2015). Growth of the endosperm syncytium has a significant impact on the final size of the seed (Boisnard-Lorig et al, 2001; Ohto et al, 2009) and so the syncytial growth phase is important in seed development and determining yield. Paternal dosage effects have been shown to impact seed size based on the timing of endosperm cellularization. Excess maternal contribution leads to earlier cellularization and smaller seeds (Scott et al, 1998).

The seed coat is formed from the maternal integuments, which surround the endosperm and embryo during seed development (Chaudhury et al, 2001; Sundaresan et al, 2005). The integuments initially consist of five cell layers, which follow specific fates including tannin synthesis by the endothelial layer and formation of the columella (Haughn and Chaudhury, 2005). Tannins in the seed coat have also been shown to be related to seed dormancy: more tannin accumulation is associated with increased seed dormancy (MacGregor et al, 2015). As the maternal integuments surround the rest of the seed and eventually form the seed coat, they provide a cavity for growth of the embryo and endosperm (Li and Li, 2016).

1.2.2 Genetic control of seed development

Genes controlling seed development are key to seed size, and a number of seed size pathways which act to control the timing of seed development as well as cell expansion and proliferation are known. These varied modes of seed size control include ubiquitin signalling, responses to phytohormones, parental effects, G-protein signalling, and the roles of transcription factors (reviewed in Li et al, 2015; Le et al, 2010).

DA1 is a gene which acts as part of the ubiquitin pathway to limit seed size. This gene acts maternally alongside *ARF2* to control seed size by restricting cell proliferation in the maternal integuments (Fang et al, 2012). *DAR1* (*DA1-RELATED PROTEIN 1)* also acts redundantly with *DA1* to restrict seed and organ growth (Li and Li, 2015). *DA1* is induced in response to ABA, and *da1-1* mutants have been shown to be resistant to developmental inhibition by ABA (Li et al, 2008). As ABA is involved in a wide variety of responses to abiotic stress in plants, this suggests that *DA1* and related genes could regulate cell proliferation to impact seed size in response to environmental perturbation through the ABA response (Finkelstein et al, 2002; Li et al, 2008).

ARF2/MNT is a transcription factor which responds to auxin and so regulates gene expression of some genes containing an auxin response element in the promoter in an auxin-responsive manner (Li and Li, 2015). This transcription factor regulates seed size by restricting cell proliferation in the maternal integuments (Schruff et al, 2006). It has also been linked to regulation of *ANT* expression (Li and Li, 2015). *ANT* encodes an AP2-like transcription factor which promotes cell proliferation and regulates integument growth (Elliott et al, 1996; Klucher et al, 1996).

APETALA2 (AP2) is a transcription factor which acts to control seed size (Jofuku et al, 1994). In Arabidopsis, *ap2* mutants have increased seed mass due to enhanced endosperm proliferation and delayed endosperm cellularization, highlighting the role of *AP2* in negatively regulating cell expansion in maternal tissues (Jofuku et al, 2005; Ohto et al, 2005; Ingram, 2010). RNA -induced silencing of a *Brassica napus* homologue of AP2 results in similar increases in seed size as well as alterations to seed and floral development (Yan et al, 2012). Analysis of seed development in *ap2* mutants (Ohto et al, 2009) reveals a number of phenotypes including larger seeds, later cellularization of the embryo, elongated integument cells and a greater size and number of cells in the mature embryo. This suggests that *AP2* restricts seed size by inhibiting growth of the seed by these means.

AP2 is also involved in floral organ development and the floral transition (Yant et al, 2010) and has been shown to prevent flowering before vernalisation in the *Arabidopsis*

perennial relative *Arabis alpina* (Bergonzi et al, 2013).This involvement in vernalisation and flowering pathways may suggest that control of seed properties by *AP2* and other AP2/ERF transcription factors is sensitive to vernalisation and environmental perturbation.

Another transcription factor which works to control seed size in Arabidopsis is *TRANSPARENT TESTA GLABRA 2 (TTG2)*. The *TTG2* gene encodes a WRKY family transcription factor (Johnson et al, 2002) which upregulates cell expansion in maternal tissues, contributing to increased seed size. In Arabidopsis, the *ttg2* mutation impacts the seed coat and restricts the growth of the endosperm, and hence limits seed growth by preventing cell elongation in the integuments (Garcia et al, 2005). It is also possible that reduced growth in the integuments acts an indirect physical restrictor of endosperm growth (Li and Li, 2015). The action of *TTG2* also integrates with the HAIKU pathway as the *iku2 ttg2* double mutant shows further reduced seed size compared to single mutants. The seed phenotype resulting from the two mutations can be attributed to the addition of the maternal effect of *ttg2* to the zygotic effect of *iku2* (Garcia et al, 2005).

The HAIKU pathway has three main components which are expressed in the endosperm and contribute to its development (Garcia et al, 2003; Luo et al, 2005; Wang et al, 2010). *IKU1* is a VQ motif protein expressed in the endosperm during early seed development, where it regulates endosperm growth. Seeds of homozygous *iku1* mutants are small and show reduced endosperm growth (Wang et al, 2010). *IKU2* is a leucine-rich repeat kinase which alongside *IKU1* has been identified as a gene contributing to seed size, with *iku2* mutants producing small seeds in Arabidopsis (Garcia et al, 2003). The endosperm of *iku1* and *iku2* mutants has been found to cellularize early and show reduced cell proliferation, as well as showing reduced cell proliferation in the embryo (Garcia et al, 2003; Luo et al, 2005).

MINI3, also known as *WRKY10*, is a WRKY class transcription factor which is also expressed in the endosperm at an early stage of seed development. Its expression forms a gradient where it is highly expressed at the micropylar endosperm and expression

reduced towards the chalazal endosperm. *MINI3* expression is affected by the inverse gradient of *MPK10* across the endosperm. *MINI3* and *MPK10* work to control seed size, with MPK10 acting negatively and MINI3 acting positively to control seed size (Xi et al, 2021). Homologues of *TTG2* and *MINI3* have been associated with seed weight QTLs in *Brassica napus* (Fan et al, 2010).

IKU2 and *MINI3* can be activated by *SHORT HYPOCOTYL UNDER BLUE 1 (SHB1)*. Overexpression of *AtSHB1* in *Brassica napus* leads to increased expression of the *Brassica napus* ortholog of *IKU2*, and production of larger seeds. It was additionally shown that expression of this *IKU2* ortholog from *Brassica napus* can rescue the small seed phenotype of *iku2* mutants in *Arabidopsis thaliana*, suggesting that comparisons can be drawn between their function (Xiao et al, 2016).

Changes in the growth of the integument may act as a physical method of constraining seed growth. A gene which negatively regulates plant growth and seed size, *EPA1*, increases tensile stress in the integuments (Creff et al, 2015; Zhang et al, 2011). In an opposing mechanism, endosperm turgor related to movement of sugars into the endosperm is a key feature of seed growth and turgor pressure from the endosperm has been shown to drive seed expansion (Beauzamy et al, 2016). Additionally, this mode of seed growth is impacted in Arabidopsis *prc2* mutants, suggesting a role of genetic imprinting as PRC2 is a regulatory complex in plants which contributes to regulation of genomic imprinting (Raissig et al, 2013). The effect of endosperm turgor is also impacted in *iku2* mutants, contributing to their reduced seed size. This is linked to a constant turgor pressure in *iku2* mutants compared to wild-type Arabidopsis seeds, where the pressure decreases after the early rapid growth of the embryo (Creff et al, 2021).

The *MINI3* cascade also relates to the function of *AN3* (*ANGUSTIFOLIA3*). It has been shown that in Arabidopsis, *AN3* represses *MINI3* to control seed size through regulation of seed expansion (Meng et al, 2016). *AN3* also functions in a cascade with *YDA*, influencing seed size through sugar accumulation. When *AN3* function is lost, control of seed expansion through *MINI3* is impacted and seeds are able to grow to a larger size.

Arabidopsis mutants in *AN3* with increased seed size were also found to have a reduced seed number per silique (Meng et al., 2016). There is a known inverse correlation between seed size and number (Alonso-Blanco et al, 1999). However, this study suggested that the majority of the seed size increase in *AN3* mutants is due to the function of *AN3* rather than solely due to compensating for the reduced fertility of the mutants.

AN3, alternatively known as *GIF1*, is part of the GIF family of transcriptional co-activators (Kim and Kende, 2004). This family of genes is required for normal development of reproductive organs in *Arabidopsis thaliana*, including the carpel margin meristem. As a result, mutants in *AN3* have abnormal development of the ovule and gynoecium (Lee et al, 2014). A lower number of fertile ovules, as well as reduced numbers of ovules per ovary, have been linked to decreased numbers of seed per pod, a measure of seed yield, in *Brassica napus* (Bouttier & Morgan, 1992).

1.2.3 Genetic background of seed maturation processes

Seed maturation occurs once the seed has developed and accumulated storage reserves, and the embryo has reached the mature green stage. This involves a number of processes including the development of desiccation tolerance and seed dormancy (Angelovici et al,2010).

LEC1, *LEC2*, *FUS3* and *ABI3* are key transcriptional regulators of seed maturation. *LEC1* encodes a HAP3 subunit of the CCAAT-binding transcription factor (CBF/NF-Y). The activity of *LEC1* during embryogenesis is key for activating the transcription of developmental genes (Lotan et al, 1998; Holdsworth et al, 2008). *LEC2* is a master regulator of genes involved in seed maturation, with targets including *DOG1* and *FUS3* (Kroj et al, 2003; Holdsworth et al, 2008). *DOG1* is a gene required for the establishment of seed dormancy (Bentsink et al, 2006).

These genes are involved in initiating and maintaining the seed maturation stage of seed development. This has been shown through plants mutated in these genes, which do not undergo normal maturation processes (Santos-Mendoza et al, 2008). All four *abi3*, *lec1*, *lec2* and *fus3* mutants are severely affected in seed maturation and share some common phenotypes related to seed quality, including decreased dormancy at maturation (Raz et al, 2001). Additionally, mutants *ec2, fus3,* and *abi3* display similar alterations of the maturation process, alongside additive double mutant phenotypes. This suggests these genes function in several overlapping regulatory pathways (Santos-Mendoza et al, 2008). As *LEC1* is also required for *FLC* resetting, there is evidence that seed maturation and vernalisation requirement are controlled by a common process.

1.3 Temperature and plant development

1.3.1 Vernalisation

Vernalisation is the process by which overwintering plants use an epigenetic memory of the temperatures they have experienced in order to correctly time flowering in spring.

The appropriate timing of developmental transitions requires accurate sensing of seasonal cues (Auge et al, 2017). A requirement for vernalisation confers the winter annual habit by delaying flowering until after prolonged cold during winter. Vernalisation is distinct from the mechanism by which the plant responds to cold through stress pathways such as freezing tolerance and cold acclimation (Thomashow, 1999; Gendall et al, 2001).

In the *Brassicaceae* family, *FLC* is the major determinant of flowering time in response to vernalisation (Sheldon et al, 2000). *FLC* encodes a MADS-box transcription factor which acts to repress flowering (Michaels and Amasino, 1999). In response to the extended period of cold experienced during vernalisation, expression levels of *FLC* drop. This allows plants to gain the competency to flower in response to environmental signals and upregulation of floral activators.

FLC is silenced by methylation using H3K27me3 histone marks. These methylation marks are then reset during seed development, allowing the next generation of plants to experience vernalisation. Prolonged periods of cold promote the expression of *COOLAIR*, the antisense *FLC* transcript. Prolonged cold also leads to *FLC* repression via the PRC2 complex and H3K27me3 marks (Swiezewski et al, 2009; Csorba et al, 2014).

Importantly, it has been shown that vernalisation of the maternal rosette before reproductive development of the seed occurs can influence the germination response of the seed, showing that vernalisation has a trans-generational impact on development (Auge et al, 2017). The vernalisation state is maintained through mitotic divisions but is reset in the next generation. As vernalisation epigenetically silences *FLC*, *FLC* expression must be reactivated during seed development. This ensures the next generation of plants also has a vernalisation requirement (Finnegan et al, 2000; Sheldon et al., 2008; Crevillén et al, 2014). The regulation of *FLC* and its ability to control flowering time is the result of a number of pathways (Figure 1.1, from Whittaker and Dean, 2017). The activity of the autonomous pathway occurs independently of temperature (Koorneef et al, 1998).

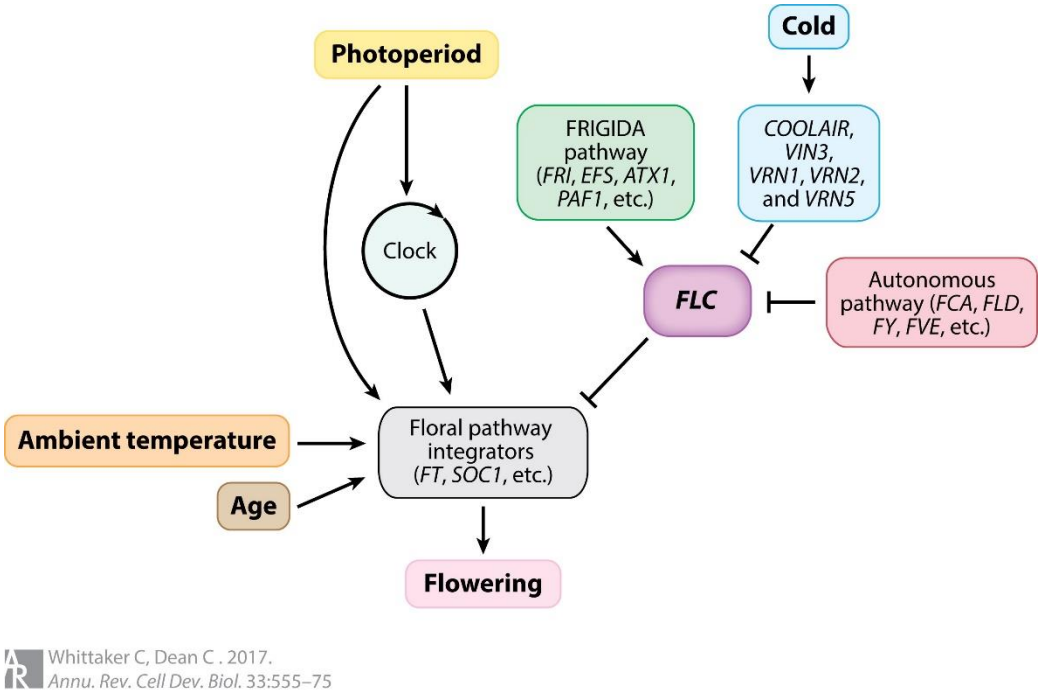

Figure 1.1: Role of FLC and other floral pathway integrators in promotion of flowering. From Whittaker and Dean, 2017.

FLC is also regulated by a number of cold-responsive genes. *VRN1* and *VRN2* have roles in vernalisation, as *vrn*-family mutants cannot reduce *FLC* levels in response to vernalisation (Sheldon et al, 2000). *VRN2* encodes for a zinc finger protein which is involved in the maintenance of *FLC* downregulation needed to induce flowering as a result of vernalisation. Expression levels of *VRN2* are not altered by vernalisation and *VRN2* is not required to establish *FLC* repression. The function of *VRN2* is instead to maintain the repression and so retain the memory of cold experienced by the plant through reduced *FLC* expression (Gendall et al, 2001).

Levels of *VIN3* mRNA increase over periods of prolonged cold, leading to repression of *FLC* (Sung and Amasino, 2004; De Lucia et al, 2008; Finnegan et al, 2011). For *VIN3* to be consistently upregulated in response to temperature, constant low temperatures which do not rise above a 15°C warm threshold are required, as *VIN3* is rapidly downregulated in warm temperatures (Hepworth et al, 2018).

FLC activity is also regulated through the *FRIGIDA* pathway. *FRI* promotes *FLC* expression and contributes to flowering time through the formation of a *FRI-C* transcription activator complex (Johanson et al, 2000; Choi et al, 2011). *FLC* expression is also upregulated by other transcriptional activators, such as the PAF1 complex (He et al, 2004) and *ATX1* (Pien et al, 2008).

The impact of *FLC* can be seen through its effects on downstream genes, such as *SOC1* and *FT*. *FLC* represses flowering through the repression of floral activators, including *FT* through the formation of a heterodimer with *SVP* which acts to repress *FT* (Mateos et al, 2015). Both *FT* and *SOC1* can be repressed through the action of the PRC2 complex (Del Olmo et al, 2016). In seed dormancy however, *FT* suppresses *FLC* mRNA expression through the activation of *COOLAIR* transcription, and so in this context *FLC* acts downstream of *FT* (Chen and Penfield, 2018).

Variation in growth environments impact the genetic control of vernalisation (Hepworth et al, 2020; Zhao et al, 2021). In Arabidopsis, vernalisation responses differ in efficiency across different optimal temperature ranges, based on underlying *FLC* variation between accessions (Duncan et al, 2015). Plants show a graded response to cold at intermediate temperatures, with vernalisation occurring more weakly as temperatures approach the threshold for vernalisation (Wollenberg and Amasino, 2012). The *VIN3*-independent pathway represses *FLC* at low temperatures, with lower temperatures resulting in more repression (Hepworth et al, 2018; Antoniou-Kourounioti et al, 2018). *VIN3*-dependent regulation of *FLC* in response to temperature occurs over long-term, short-term, and immediate timescales, allowing the integration of a range of different temperature signals (Antoniou-Kourounioti et al, 2018). Differences in autumn temperatures affect the timing of *VIN3* activity as well as the rate of *VIN3*-independent *FLC* silencing, with warmer autumn temperatures delaying the upregulation of *VIN3* and slowing the initial downregulation of *FLC* (Hepworth et al, 2018).

1.3.2 FLC genes in Brassica napus

Much of our knowledge about *FLC* comes from research carried out in *Arabidopsis thaliana*, a relative of *Brassica napus*. However, there has also been research into *FLC* in crop species of the *Brassicaceae* family, where *FLC* behaves as a flowering regulator

analogous to its function in Arabidopsis (Tadege et al, 2001). *Brassica napus* is an allotetraploid, with an A-genome derived from *Brassica rapa* and a C-genome derived from *Brassica oleracea*. Whole genome duplication has given rise to multiple copies of *FLC* within the *Brassica napus* genome; additionally, *BnaFLC.A03b*, *BnaFLC.C03b, and BnaFLC.C09a* are believed to have been generated by smaller duplication events (Calderwood et al, 2021; Cai et al, 2014). The *Brassica napus* genome contains up to nine copies of *FLC* in total (Zou et al, 2012).

Variations in *FLC* genes have been identified in different crop types of *Brassica napus* with different vernalisation requirements (Hou et al, 2012; Wu et al, 2019). Total *FLC* dynamics, rather than the dynamics of individual copies of *FLC*, are most related to the difference in vernalisation requirement and flowering habits between crop types of *Brassica napus* (Calderwood et al, 2021). Spring OSR shows little to no cold requirement, and levels of *FLC* are low from early in development. Semi-winter OSR varieties have been bred for mild winters and display a low vernalisation requirement with *FLC* levels decreasing rapidly following cold exposure. Winter OSR varieties generally have the strongest vernalisation requirement and require a long period of cold to vernalize, as *FLC* levels drop slowly.

Polyploidy in *Brassica napus* has led to gene duplication, causing loss of function or sub-functionalization in some genes through mechanisms such as changes in gene regulation or protein function (Osborn et al, 2003). This is seen in *Brassica napus* flowering time genes including *TFL1* (Jones et al, 2018) and *FLC* (Zou et al, 2012).

Not all *FLC* copies present in *Brassica napus* have the same functions or response to low temperatures (Schiessel, 2020). *BnaFLC.A10* is strongly associated with the vernalisation requirement in *Brassica napus* (Tadege et al, 2001; Hou et al, 2012). *BnaFLC.C03b* is a pseudogene which is non-functional due to the presence of stop codons (Zou et al, 2012). A lack of cold response is seen in three *Brassica napus FLC* genes*: BnaFLC.C09a*, *BnaFLC.C09b*, *BnaFLC.C03a* (Schiessel et al, 2019).

1.3.3. *FLC* in seeds

As the process of vernalisation epigenetically silences *FLC* through H3K27 trimethylation, *FLC* expression must be reactivated during seed development to ensure the next generation of plants also has a vernalisation requirement. A variety of genes and gene complexes have been shown to impact *FLC* reactivation. It has been shown that for the full reactivation of *FLC* to occur during embryogenesis, a protein complex containing *FRIGIDA* (*FRI*) is required, as well as the *ELF7* and *VIP4* genes which are homologs of members of the yeast Paf1 complex (Yun et al, 2011). Reducing the activity of *ELF6*, a gene which functions as a H3K27me3 demethylase, impairs the reactivation of *FLC* in seeds. *ELF6* activity was also found to increase *FLC* expression through the progression of embryo development (Crevillén et al, 2014).

The parental vernalised state at *FLC* is reset by several proteins in addition to *ELF6*. *LEC1* is a seed-specific NF-YB transcription factor which has a role in reversing *FLC* silencing that is inherited from gametes. *LEC1* acts to establish an active *FLC* chromatin state in the seed (Tao et al, 2017). *LEC1* has been found to directly re-activate *FLC* expression in early embryogenesis, within 1 day after pollination (Tao et al, 2017).

Resetting of *FLC* occurs differently in the male and female germlines. *FLC* is not expressed in gametophytes and is reactivated after fertilization in the embryo but not the endosperm (Sheldon et al, 2008; Choi et al, 2009). It has been shown that although both the maternal and paternal copies of *FLC* are reset in the new generation, the paternal copy of *FLC* is reactivated in the fertilized embryo while the maternal *FLC* copy is reactivated later, in the early-globular stage of the embryo (Sheldon et al, 2008). FLC is silenced by DNA methylation using H3K27me3 markers. DNA methylation markers are not globally erased from DNA in the sperm of flowering plants (Calarco et al, 2012). In Arabidopsis, sperm chromatin contains a sperm-specific histone variant H3.10 which is not compatible with K27 methylation (Ingouff et al, 2010; Borg et al, 2015; Borg et al, 2020). Its incorporation in sperm would therefore cause a loss of H2K27me3 markers. This contributes to resetting of methylation in the male germline. *FLC* in the male germline was found to be primed with H3K4me3 markers which encourages transcription in early embryogenesis (Borg et al, 2020).

This system of gene resetting in the paternal germline is in contrast to the maternal germline, where PRC2 is expressed in the central cell and so maintains H3K27me3 markers (Grossniklaus and Paro, 2014). *FLC* remains maternally repressed through fertilization; resetting of methylation at maternal copies of *FLC* instead occurs later in embryo development, with longer developmental time required to reset *FLC* expression in plants which were more strongly vernalised (Luo et al, 2020). As the vernalisation state in early embryo development is inherited maternally and not paternally, these differences in regulation can explain the maternal *FLC* effect seen in Arabidopsis.

The downstream targets of *FLC* include *SOC1* and *AP1*; in addition to their role in flowering time, these genes contribute to the regulation of seed germination by *FLC* through downstream integration with ABA and GA (Chiang et al, 2009). High *FLC* expression during seed maturation influences gene expression in germinating seeds and is associated with increased germination at cooler temperatures (Chiang et al, 2009). *FLC* has also been implicated in variation in germination and seedling vigour in *Brassica rapa* (Basnet et al, 2015).

1.3.4 Temperature effects on seeds

In a variety of plant species, temperature during the vegetative stage of maternal development, before seeds are produced, has been shown to impact seed development (Roach and Wulff, 1987; Kegode and Pearse, 1998; Andalo et al, 1999; Donohue, 2009). A number of studies have focused on how maternal environment impacts seed germination and dormancy. Seed dormancy refers to the state in which the seed will not germinate in conditions which are favourable for germination (Bewley, 1997).

Low temperatures experienced by the mother plant have been shown to result in increased levels of dormancy in *Arabidopsis* mature seeds (Kendall et al, 2011), and further research has shown that imprinting is used as a mechanism for the maternal inheritance of seed dormancy (Piskurewicz et al, 2016). This process involved the preferential expression of maternal alleles in the endosperm for genes which could be associated with seed dormancy after imbibition (including *CYP1, CYSTEINE*

PROTEASE 1) allowing the maternal genotype to be reflected in seed dormancy levels. In *Wahlenbergia tumidifructa*, seed longevity was also found to be affected by the pre-zygotic growth environment, with cooler temperatures resulting in a greater variation in seed longevity (Kochanek et al, 2010). In *Arabidopsis thaliana*, seed production at high temperatures was found to be improved through three generations produced from plants which had experienced heat treatments (Whittle et al, 2009).

The temperature experienced by the plant during the processes of seed development and seed maturation is also known to affect seed traits (Fenner, 1991). The maturation temperature seeds are exposed to is associated with variability in seed performance (He et al, 2014). In Arabidopsis, the seed maturation temperature has been shown to have a strong effect on seed performance, with low temperatures reducing seed longevity (He et al, 2016).

Seed dormancy in Arabidopsis is also sensitive to environmental effects including temperature during the seed maturation stage (reviewed in Penfield and MacGregor, 2017). Low seed maturation temperatures in Arabidopsis lead to deep primary dormancy, which is induced via mechanisms relying on *DOG1* and gibberellin and abscisic acid signalling (He et al, 2016; Kendall et al, 2011). In contrast, higher seed maturation temperatures lead to shallower dormancy (MacGregor et al, 2015; Chiang et al, 2011; Kendall et al, 2011). Seed dormancy is controlled via a range of systems, with effects on fruit development, the endosperm and the seed coat having effects on the final dormancy level of the seed (Penfield and MacGregor, 2017; Bethke et al, 2007; MacGregor et al, 2015). Effects of temperature on any of these processes could go on to impact seed dormancy.

1.4 Temperature impacts on crop development

1.4.1 Impact of vernalisation temperature on crops

A general plant response to vernalisation occurs at temperatures below 17°C and at temperatures higher than this many plants, including oilseed rape, will no longer be able to vernalise conditions (Tommey and Evans; 1991; Porter and Gawaith, 1999; Penfield, 2008; Duncan et al, 2015; Penfield et al, 2021). The environment experienced by the parent plant before seed set has occurred has been shown to affect the seed in a number of crop species, including wheat and *Brassica rapa* (Sanhewe et al, 1996; Sinniah et al, 1998).

Previous work (O'Neill et al, 2019) has shown that winter OSR varieties of *Brassica napus* complete vernalisation during October and undergo the floral transition shortly after in November. The subsequent floral primordia are therefore exposed to cooler winter temperatures. As winter temperature has an effect on the final overall crop yield (Brown et al, 2019), this suggests there may be a mechanism whereby the winter temperature experienced by the floral meristem has effects later in development. In the field warming trial described in O'Neill et al (2019), warming the plots delayed the floral transition, although this did not delay final flowering time. This suggests that vernalisation temperature is having an effect on the timing of the floral transition, which in turn delays development of the floral primordia.

Previous work in the Penfield group (Figure 1.2, unpublished data) has also shown that vernalisation has an effect on *Brassica napus* seed size in a Winter x Semi-winter OSR DH population, created by crossing the Winter OSR variety Tapidor with the Semi-winter OSR variety Ningyou. The lines produced from this cross showed range of responses to temperature. Notably, plants which were vernalised in the lab before being transplanted to the field produced seeds which were always equal to or larger in size than seeds produced by plants which did not undergo a vernalisation treatment.

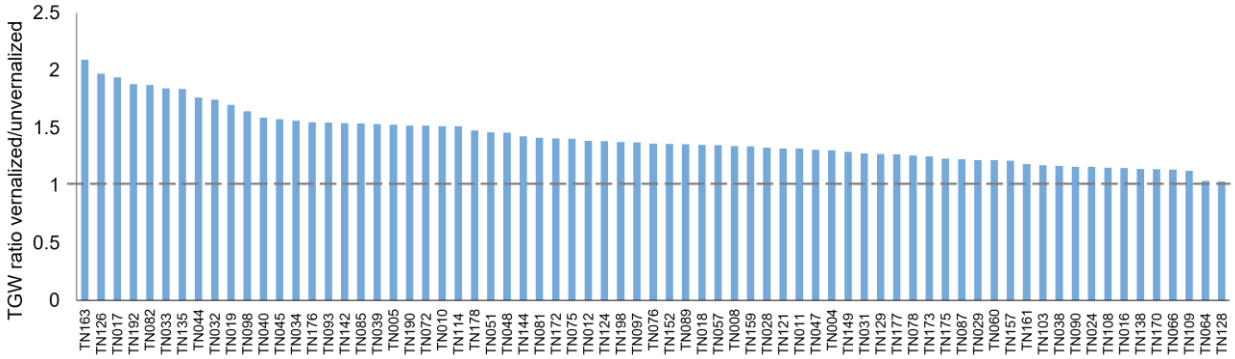

Figure 1.2: TGW ratio of a Tapidor x Ningyou population, vernalised/unvernalised treatments. Ratios of greater than 1 (grey dashed line) indicate seeds were larger in plants given a vernalisation treatment.

1.4.2 Impact of seed maturation temperature on crops

Seed properties are not only affected by vernalisation; other environmental factors such as photoperiod, light and nutrient levels impact the quality of seeds (Fenner, 1991; Imaizumi et al, 2017). In *Brassica napus*, increased temperatures throughout development have been shown to negatively affect reproductive development. These temperature effects impact a range of processes throughout development required for the plant to produce high-quality seeds. For example, high temperatures can perturb ovule development and microsporogenesis; the resulting abnormal development of floral organs leads to the production of sterile flowers and a loss of yield (Polowick and Sawhney, 1988). Alongside floral sterility, high levels of heat stress during flowering can impact the number and size of seeds produced (Morrison and Stewart, 2002).

Crops are affected by high temperatures during seed development. For example, heat stress reduces seed size in rice by a PRC2 based mechanism (called FIE1 in rice) (Folsom et al, 2014). Heat stress early in seed maturation has been shown to negatively reduce *Brassica napus* seed quality (Magno Massuia de Almeida et al, 2021), as well as reducing seed weight when heat stress occurs during seed filling (Rashid et al, 2018).

High temperatures during seed development and maturation also affect dormancy and germinability in seeds: lower seed maturation temperatures generally increasing seed

dormancy (Fenner, 1991). In *Arabidopsis*, increasing temperature during seed set by 1°C from 14°C to 15°C significantly impacted seed dormancy. Specifically, seeds which developed at a higher temperature were shown to germinate more readily (Springthorpe and Penfield, 2015). The effect of increased temperatures on promoting seed germination appears to be regulated within the endosperm (Chen et al, 2021).

1.4.3 Impacts of temperature on seed size and development

Auxin is a key phytohormone that determines seed size and weight. It is involved in development of the embryo, endosperm, and seed coat (reviewed in Cao et al, 2020). Several genes have been identified which affect seed size in response to auxin. In Arabidopsis, the seed size regulator gene *LEC2* influences embryo development by regulating the supply of auxin to the embryo (Stone et al, 2008; Wójcikowska et al, 2013). In *Brassica napus*, the gene *CYP78A9,* which encodes a cytochrome P450 enzyme, contributes to seed size by influencing auxin metabolism (Shi et al, 2019). In Arabidopsis seeds, auxin accumulates from the heart embryo stage through to the development of cotyledons (Ni et al, 2001).

Auxin contributes to various developmental processes related to seed size, across a range of species. Auxin is involved in the regulation of endosperm development in maize (Lur et al, 1993), and is required for normal seed development in peas (McAdam et al, 2017). Auxin signalling and seed size are also related in rice, where a loss of function in rice cytochrome P450 gene leads to lower IAA levels and increased embryo size (Chen et al, 2015). Auxin production also affects endosperm development in Arabidopsis: increased auxin levels are linked to endosperm cellularization defects (Batista et al, 2019).

Auxin signalling pathways can be affected by temperature. Growth at high temperatures leads to increased levels of auxin (Gray et al, 1998; Sun et al, 2012). High temperatures influence auxin signalling in determination of microspore polarity and resulting embryo

development (Dubas et al, 2014). As such, auxin signalling pathways could be a linking mechanism between temperature responses and resulting effects on seed size.

Abscisic acid (ABA) is important for seed development. ABA regulates seed traits, including lipid synthesis and seed dormancy (reviewed in Finkelstein et al, 2002). Levels of ABA in early seed development contribute to seed size; Arabidopsis mutants with lower ABA levels show delayed endosperm cellularization and increased seed size (Cheng et al, 2014).

ABA is also well known to have a role in seed maturation and the development of dormancy (Kanno et al, 2010). Loss of seed desiccation tolerance and premature seed germination can be observed in a range of Arabidopsis mutants with both ABA-dependent and ABA-independent responses. This includes *abi3* mutants, which have reduced ABA sensitivity in the seeds (Koornneef et al, 1984; Nambara et al, 2000; To et al, 2006). Plants with reduced ABA signalling have early germinating phenotypes and strongly reduced requirement for seed dormancy. High levels of ABA are associated with strong dormancy in seeds and repression of germination (Koornneef et al, 1984; Bewley, 1997; Graeber et al, 2012).

Previous studies have also shown that ABA signalling pathways are responsive to environmental stresses, including temperature (Daie and Campbell, 1981; Walker-Simmons, 1988; Larkindale and Knight, 2002; Lim et al, 2013; Chiu et al, 2016). As ABA levels are responsive to temperature and involved in control of seed development, ABA pathways could form part of a link between temperature response and seed traits. The effect of maturation temperature on ABA in seeds has been shown; low seed maturation temperatures lead to increased levels of ABA in the endosperm in *Brassica oleracea* and warm maturation temperatures lead to higher germination (Chen et al, 2021). The effect of seed maturation temperature on germinability is a feature of plant development seen in a wide range of species (Penfield, 2017) including Arabidopsis (Chen et al, 2014; Springthorpe and Penfield, 2015), lettuce (Gray et al, 1998), soybean (Egli et al, 2005) and pea (Lamichaney et al, 2021).

1.5 Overall Aims

The overall aim of the PhD project presented in this thesis is to learn more about how temperature during development impacts seeds in *Brassica napus*, with a view to improving OSR yield, and finding out what genetic pathways are involved in temperature responses as well as the mechanism by which vernalisation affects seed size.

This has been done by growing and phenotyping a range of *Brassica napus* varieties and using this phenotyping data as a basis for identifying genetic differences underlying temperature responses. This was to test the hypothesis that temperature during vernalisation and maturation would have impacts of seed properties which contribute to overall yield. Results will be described from studying the properties of the seeds such as seed size, seed number, oil content and germination. The results of analysing genetic differences between varieties with contrasting phenotypes will be described, exploring the hypothesis that specific genes and genetic pathways which may underlie temperature effects on seeds in *Brassica napus* can be identified using this dataset.

Some varieties of *Brassica napus* have also been tested for differences in gene expression resulting from changes in vernalisation temperature, exploring the hypothesis that vernalisation temperature will affect gene expression patterns throughout development and in the seed. Combining phenotyping and transcriptomics with a range of environmental conditions is intended to reveal more about how *Brassica napus* responds to different temperatures on a gene expression level. Exploring these factors can lead to a greater understanding of how developmental pathways in the seed respond to temperature and affect seed traits. These results will be considered to test the hypothesis that gene expression changes can be linked to identified phenotypes in the seed.

Chapter 2 - Exploring the effect of temperature on *Brassica napus* yield

2.1 Introduction

As previously described in the introduction, *Brassica napus* is grown as an oilseed crop with seed yield being a vital factor in overall crop yield. Two major contributing factors to overall seed yield are the number of seeds produced by each pod (seeds per pod, or SPP) and the seed size, measured by thousand grain weight (TGW). As temperature has been correlated with yield in *Brassica napus* under field growth conditions (Brown et al, 2019), this chapter explores the effect of temperature on these two yield parameters.

When phenotyping accessions of *Brassica napus* in this project, a core set of 96 accessions from the Diversity Fixed Foundation Set (DFFS) was used (listed in Appendix A). The DFFS was developed as part of the OReGIN project (Harpet et al, 2012). This set contained a number of different crop types, including Winter OSR, Spring OSR and Semi-winter OSR varieties alongside swedes, kales, exotics and fodder crops. As *Brassica napus* crop types have different vernalisation requirements, this population enables us to investigate the impact of variation in vernalisation requirement on seed development.

When studying crops in the field, it can be difficult to separate the knock-on effect of the timing of developmental transitions from the direct effects of temperature on specific growth stages. For example, a warm vernalisation temperature can delay flowering, which may then cause seed development to occur later during the growing season. At this point it can become difficult to tell if effects on seed development were caused directly by a warm vernalisation temperature, by a difference in seed maturation

temperature or if changes in yield were due to other confounding factors between growing seasons. When designing phenotyping experiments to explore temperature effects on yield, it was important to be able to separate the effects of vernalisation temperature and seed maturation temperature in a way that may be obscured when growing plants in field conditions.

2.2 Hypotheses & Aims

As winter temperature has been previously correlated with yield (Brown et al, 2019), the first hypothesis was that the temperature during vernalisation treatment is correlated with seed yield parameters (SPP and TGW).

As vernalisation is a response to extended periods of cold, length of vernalisation treatment was also considered. The second hypothesis was that the duration of vernalisation is correlated with seed yield parameters (SPP and TGW).

Maturation temperature has also been shown to be important for seed development and yield (Rashid et al, 2018), so the third hypothesis tested in this chapter is that maturation temperature is correlated with seed yield parameters (SPP and TGW).

As crop type is a factor contributing to adaptation of *Brassica napus* to different environments, the fourth hypothesis explored in this chapter is that temperature responses would vary between crop types.

After completing the large phenotyping experiment, this chapter addresses the aim to find out if the results are sufficiently robust that a subset the results can be replicated for continued experiments.

2.3 Results

2.3.1 Experimental design of large phenotyping experiment

For this phenotyping experiment, 96 accessions of *Brassica napus* from the Diversity Fixed Foundation Set (DFFS) were included (Harper et al, 2012).

In total, twelve different temperature treatments were selected for the large phenotyping experiment (summarized in Figure 2.1). Plants would undergo one combination of: vernalisation at either 5°C, 10°C, or 15°C; for either 6 weeks or 12 weeks; and then seed maturation at either 18°C or 24°C. The timing of sowing was staggered so that plants which underwent 6 weeks of vernalisation were sown later and all plants would be removed from the vernalisation treatment at the same time, to ensure growth conditions were consistent between treatments.

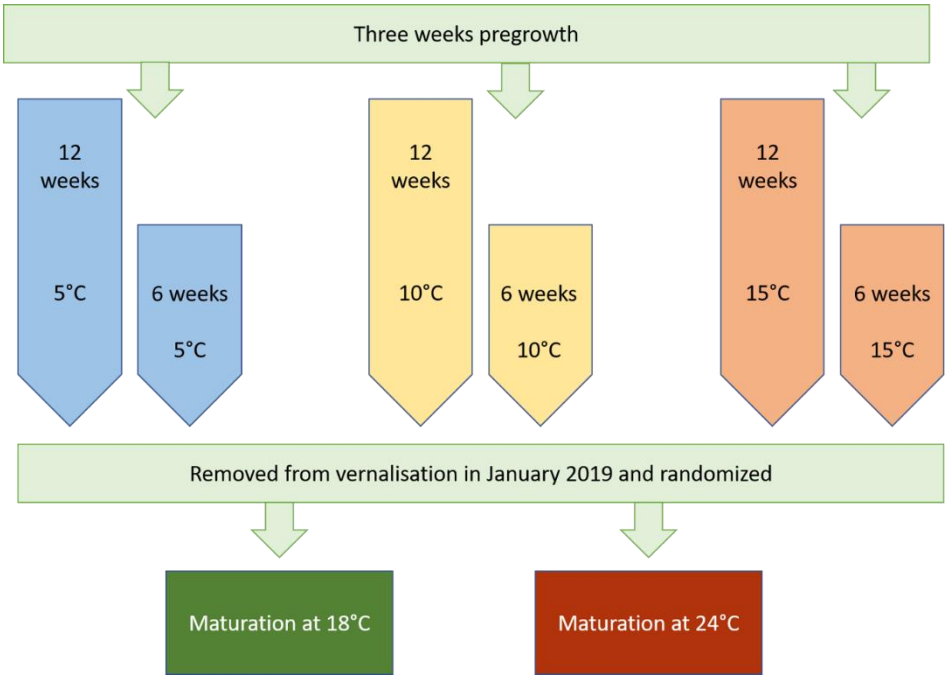

Figure 2.1: Experimental design of phenotyping experiment to test effect of temperature on seeds.

Experimental groups consisted of differences in treatment based on vernalisation temperature, vernalisation duration and maturation temperature. Each of the twelve experimental groups contained 288 plants consisting of three biological replicates of each of the 96 lines included in the experiment. To reduce the confounding effects of factors contributing to differences in growth, all plants were grown in the same soil type and size of pot. The plants were all transferred to larger growth pots at the same timepoint, after being removed from the vernalisation treatment. Watering was carried out using an automated irrigation system to ensure consistent watering across all treatments. To control for variation across the experiment, an alpha randomized blocking design was used with replicates for vernalisation temperature split across blocks (detailed further in section 2.3.4.1).

2.3.1.2 Overview of flowering time data

To understand effects of vernalisation on seed traits it was necessary to understand the variety flowering time responses to the six different vernalisation treatments. Flowering

time was examined in the plants from this phenotyping experiment by the BRAVO team at John Innes Centre in 2019. The effect of different vernalisation temperatures on each crop type is seen in Figure 2 (from Rachel Wells and Steve Penfield, unpublished data).

Plants vernalised for 12 weeks had a longer mean day number to flowering compared to plants vernalised for 6 weeks (Figure 2.2a). This effect was least pronounced in Semi-winter and Spring OSR varieties, which were the fastest flowering. Swede and Winter OSR varieties showed a larger effect, with some swede varieties vernalised for 12 weeks not flowering. Plants vernalised at 15°C had the highest mean days to flowering. Particularly in Winter OSR and swede varieties, increased vernalisation temperature was associated with taking a longer time to flower, with many swedes not flowering after a vernalisation treatment above 5°C (Figure 2.2b). In Semi-winter and Spring OSR varieties, there was little difference in time taken to flower between vernalisation treatments. These varieties also flowered most quickly.

This data shows that swede varieties took longer than other crop types to flower under similar conditions, while Spring and Semi-winter OSR had the fastest time to flowering after being removed from vernalisation and showed only small decreases in flowering time when given stronger vernalisation treatments.

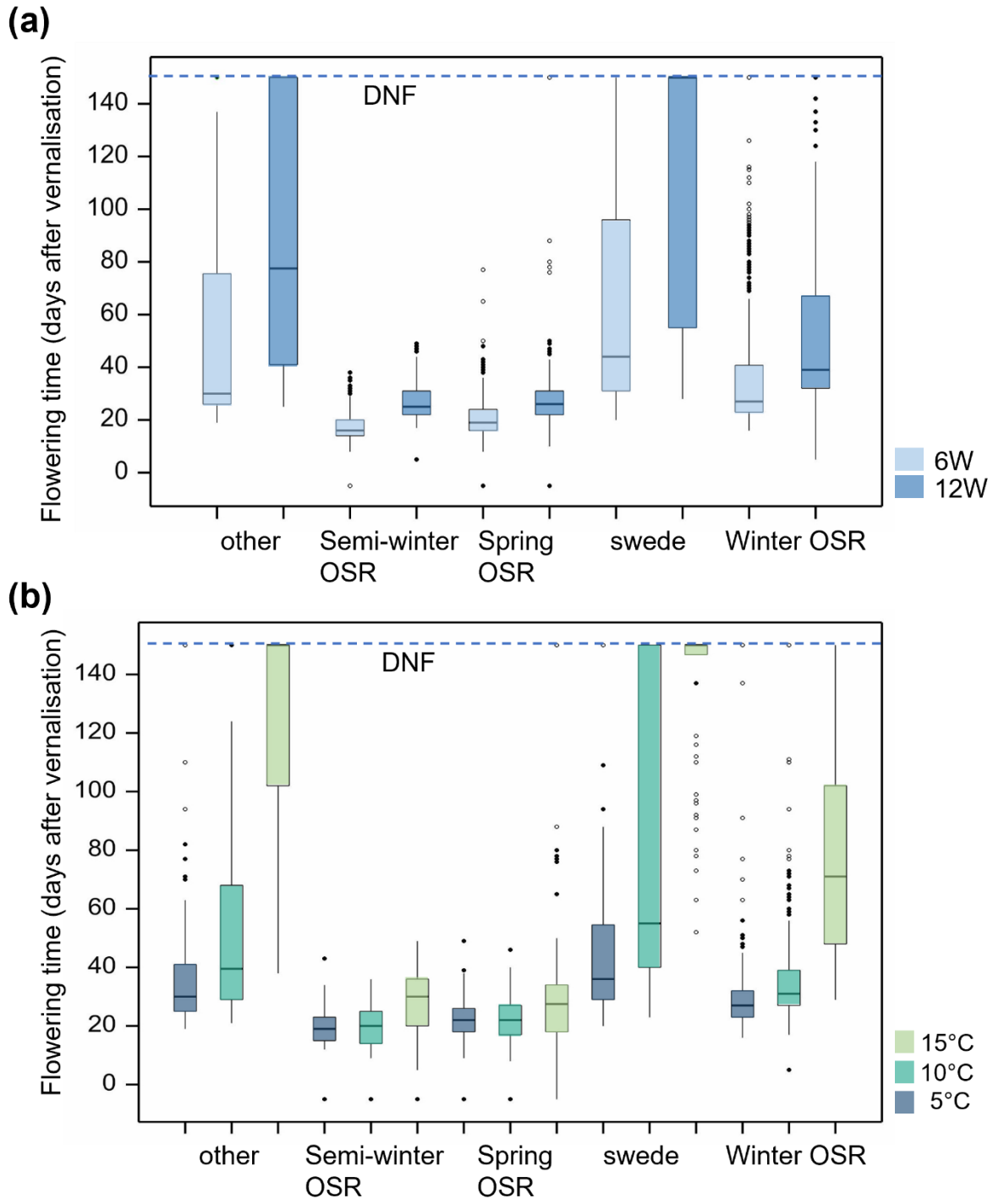

Figure 2.2: Flowering time for each crop type by (a) vernalisation duration (b) vernalisation temperature. DNF = did not flower. Data provided by Steven Penfield.

2.3.2 Comparison of seed trait data before and after cleaning

Seed size was measured from 20 pods collected from the main raceme. The phenotyping experiment was originally designed to have three replicates from each accession for each of the temperature treatments. However, not all harvested plants produced usable results which necessitated data cleaning before the full analysis could be carried out. This was due to a combination of failure to flower, failure to produce enough seed pods, disease, pre-harvest sprouting or limited fruit development (see Table 2.1).

Table 2.1: Number of samples remaining in each temperature treatment after data cleaning. The final number of plants used in the analysis was 2596 (75% of total possible samples).

Temperature Treatment	Samples remaining after data cleaning
5°C , 6 weeks vernalisation, 18°C maturation	258
5°C , 12 weeks vernalisation, 18°C maturation	261
5°C , 6 weeks vernalisation, 24°C maturation	232
5°C , 12 weeks vernalisation, 24°C maturation	244
10°C , 6 weeks vernalisation, 18°C maturation	245
10°C , 12 weeks vernalisation, 18°C maturation	263
10°C , 6 weeks vernalisation, 24°C maturation	221
10°C , 12 weeks vernalisation, 24°C maturation	240
15°C , 6 weeks vernalisation, 18°C maturation	127
15°C , 12 weeks vernalisation, 18°C maturation	211
15°C , 6 weeks vernalisation, 24°C maturation	134
15°C , 12 weeks vernalisation, 24°C maturation	160

Firstly, if any variety had less than two replicates for a given temperature treatment, that treatment was removed from the analysis for that variety. The experiment also aimed to compare the effect of temperature on seed yield parameters for seeds which had properly developed; however, some plants did not set sufficient numbers of seeds to provide reliable information. Any samples which had a TGW of less than 1.5 g were removed from the analysis as previous work had shown that seed lots with low TGW

contained high numbers of aberrant seeds. Additionally, plants where 18 or less pods had been harvested were removed from the analysis to ensure that results were taken from healthy plants which consistently produced seeds.

The difference in data distribution can be seen in Figure 2.3. Before data cleaning, for SPP (Figure 2.3a) and TGW (Figure 2.3b) the lowest bins have a higher frequency when compared to the same traits after cleaning (Figure 2.3c and 2.3d). The overall impact of the data cleaning is that major outliers towards the lower end are removed. For TGW, as the data was filtered directly on this parameter it can be seen that the small number of samples where TGW was below 1.5 g are removed (Figure 2.3d). For SPP, as the results were filtered on the overall number of pods harvested and not directly on SPP itself, the range of bins remains the same although those bins containing extremely few seeds have a reduced frequency after cleaning the data (Figure 2.3c).

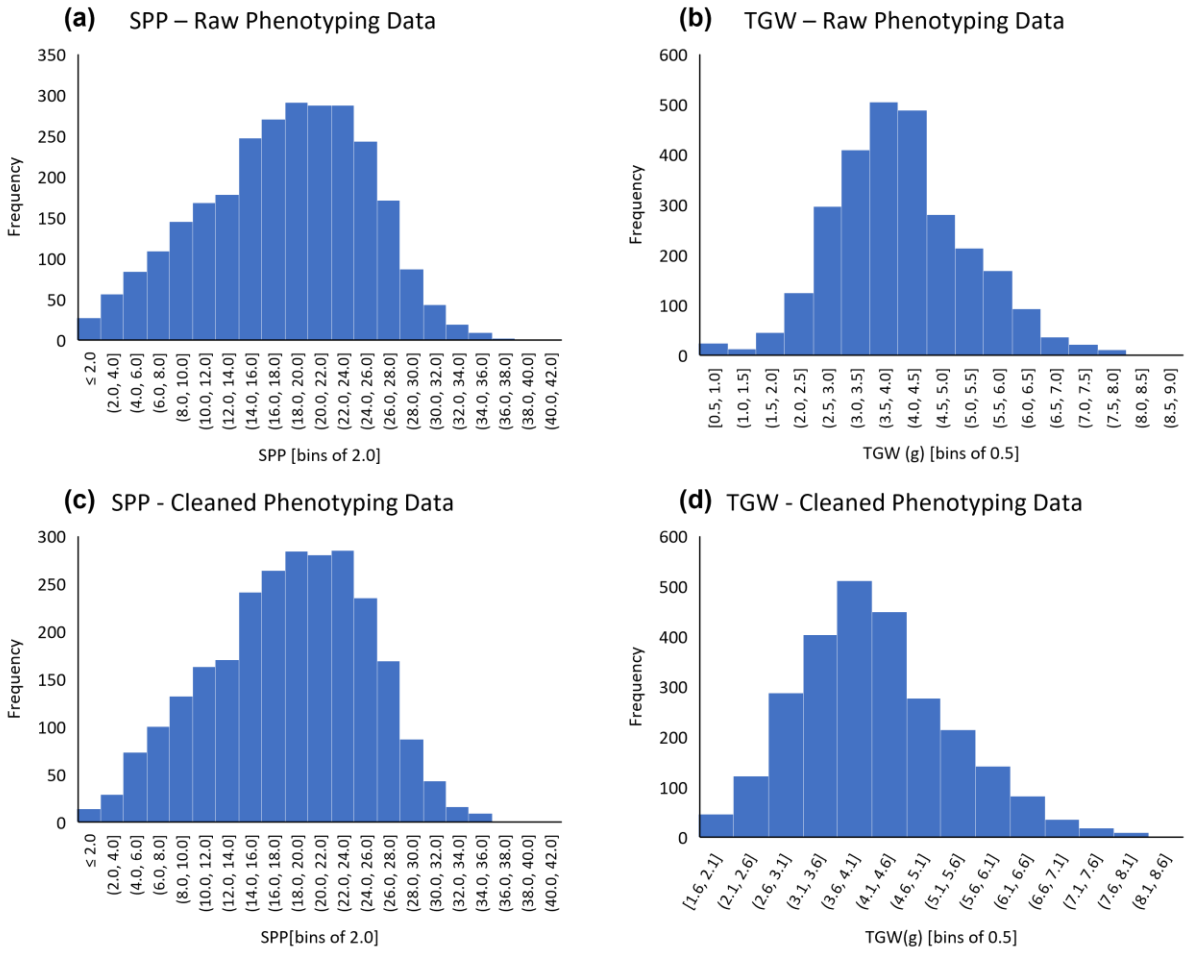

Figure 2.3: Histograms showing the distribution of data from the phenotyping experiment, before and after data cleaning. (a) Raw data for SPP. (b) Raw data for TGW. (c) Cleaned data for SPP (Pod number < 19). (d) Cleaned data for TGW (TGW < 1.5).

As can be seen in Table 2.1, the two temperature treatments with the fewest samples remaining for reliable analysis were the combination of 15°C vernalisation for 6 weeks with both 18°C and 24°C maturation temperature. This suggests that this vernalisation treatment (6 weeks at 15°C) was the least favourable for producing high-yielding plants. Conversely, the temperature treatments with the most remaining samples were plants which were vernalised at either 5°C or 10°C for 12 weeks and then matured at 18°C.

2.3.3 Preliminary overview of effects of temperature treatments on cleaned seed trait data

2.3.3.1 Effect of Vernalisation Temperature on TGW and SPP

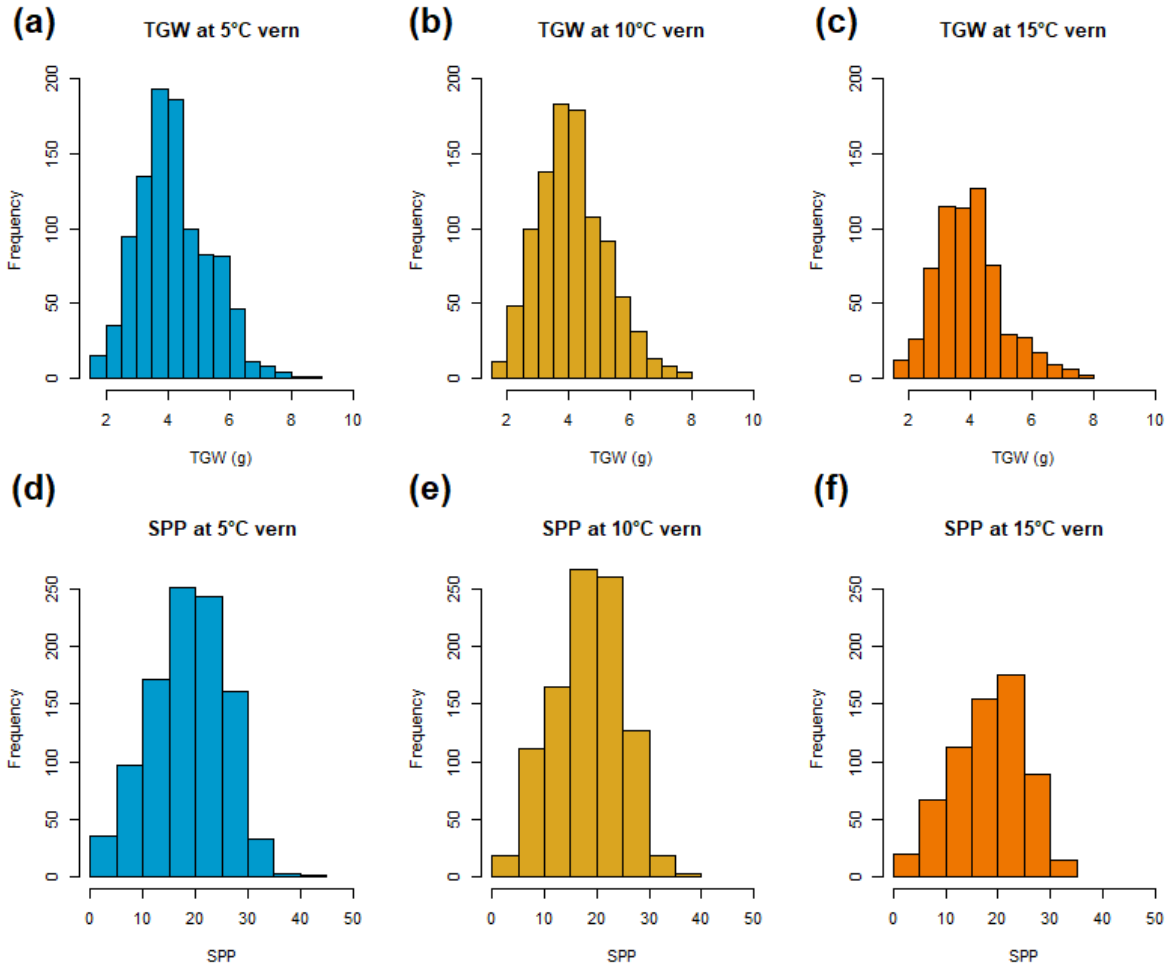

Figure 2.4: Histograms showing the distribution of TGW (a - c) and SPP (d - f) over three different vernalisation temperatures with incremental bin values of 0.5 for TGW and 5 for SPP.

In Figure 2.4 the distributions of TGW and SPP from the cleaned data across the three vernalisation temperatures can be seen. Figure 2.4a shows that TGW at 5°C from the cleaned data extended to a maximum of 8.5 – 9.0 g, with the greatest frequency in the 3.5 – 4.0 g.0 g range. Figure 2.4b shows that TGW at 10°C had a maximum range of 7.5

- 8.0 g, with the greatest frequency in the 3.5 – 4.0 g range. Figure 2.4c shows that TGW at 15°C had a maximum range of 7.5 – 8.0 g, with the greatest frequency in the 4.0 – 4.5 g range. At 5°C and 10°C, the maximum frequency for a given range approaches 200, while at 15°C the maximum frequency is below 150.

Figure 2.4d shows that SPP at 5°C from the cleaned data extended to a maximum of 40 – 45 seeds, with the greatest frequency in the range 15 – 20 seeds. Figure 2.4e shows that SPP at 10°C had a maximum range of 35 – 40 seeds, with the greatest frequency in the range 15 – 20 seeds. Figure 2.4f shows that SPP at 15°C had a maximum range of 30 – 35 seeds, with the greatest frequency in the range 20 – 25 seeds. At 5°C, the maximum frequency for a given range approaches 250, and at 10°C the maximum frequency exceeds 250, while at 15°C the maximum frequency is below 200.

Mean TGW and SPP across the treatments, calculated from the raw data, indicates that the highest values for both TGW and SPP were found in plants vernalised at 5°C, while the lowest values for TGW and SPP were found in plants vernalised at 15°C (Table 2.2). However, the difference in SPP across vernalisation treatments was minimal.

Table 2.2: Mean TGW and SPP calculated from the cleaned data.

	Mean TGW (g)	**Mean SPP (g)**
Vernalisation Temperature		
5°C	4.197	18.589
10°C	4.102	18.230
15°C	3.987	18.107
Vernalisation Duration		
6 weeks	4.123	18.233
12 weeks	4.100	18.429
Maturation Temperature		
18°C	4.324	19.880
24°C	3.875	16.627

2.3.3.2 Effect of vernalisation duration on TGW and SPP

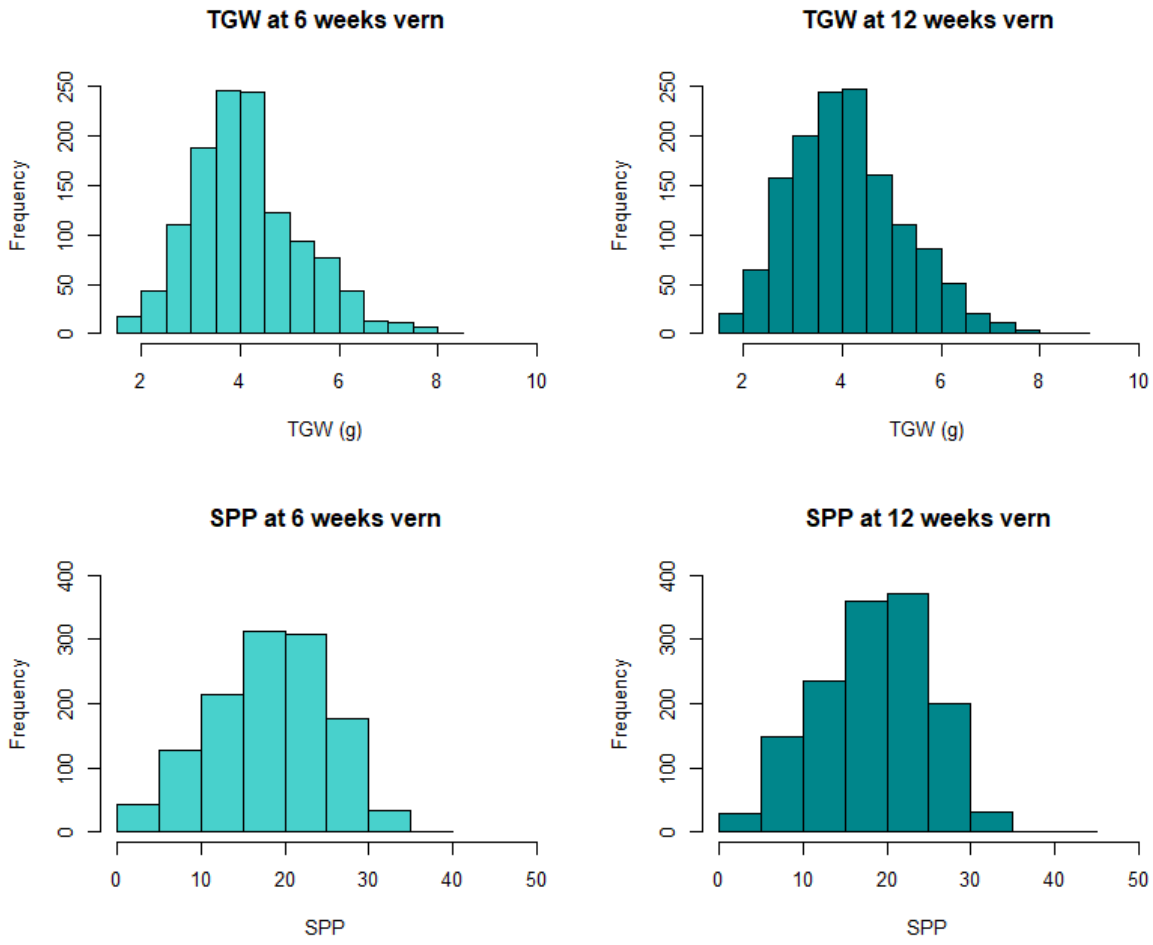

Figure 2.5: Histograms showing the distribution of TGW (a, b) and SPP (c, d) over two different vernalisation lengths.

In Figure 2.5 the distributions of TGW and SPP from the cleaned data across the two vernalisation durations can be seen. Figure 2.5a shows that TGW at 6 weeks from the cleaned data extended to a maximum of 8.0 – 8.5 g, with the greatest frequency in the 3.5 – 4.0 g range. Figure 2.5b shows that TGW at 12 weeks had a maximum range of 8.5 – 9.0 g, with the greatest frequency in the 4.0 – 4.5 g range.

Figure 2.5c shows that SPP at 6 weeks from the cleaned data extended to a maximum range of 35 - 40 seeds, with the greatest frequency in the range 15 - 20 seeds. Figure 2.5d shows that SPP at 12 weeks had a maximum range of 40 - 45 seeds, with the greatest frequency in the range 20 - 25 seeds. At 6 weeks, the maximum frequency for a given range is just over 300, while at 12 weeks the maximum frequency approaches 400.

Mean TGW and SPP across the treatments, calculated from the raw data, indicates that the highest values for TGW were found in plants vernalised for 6 weeks, while the highest values for SPP were found in plants vernalised for 12 weeks (Table 2.2).

2.3.3.3 Effect of Maturation Temperature on TGW and SPP

In Figure 2.6 the distributions of TGW and SPP from the cleaned data across the three vernalisation temperatures can be seen. Figure 2.6a shows that TGW at 18°C from the cleaned data extended to a maximum range of 8.5 – 9.0 g, with the greatest frequency in the 4.0 – 4.5 g range. Figure 2.6b shows that TGW at 24°C had a maximum range of 7.5 – 8.0 g, with the greatest frequency in the 3.5 – 4.0 g range. At 18°C, the maximum frequency for a given range exceeds 250, while at 24°C the maximum frequency is below 250.

Figure 2.6c shows that SPP at 18°C from the cleaned data extended to a maximum range of 35 - 40 seeds, with the greatest frequency in the range 20 - 25 seeds. Figure 2.6d shows that SPP at 24°C had a maximum range of 40 - 45 seeds, with the greatest frequency in the range 15-20 seeds. At 18°C, the maximum frequency for a given range exceeds 400, while at 24°C the maximum frequency is below 350.

Mean TGW and SPP across the treatments, calculated from the raw data, indicates that the highest values for both TGW and SPP were found in plants grown at 18°C, while the lowest values for TGW and SPP were found in plants grown at 24°C (Table 2.2).

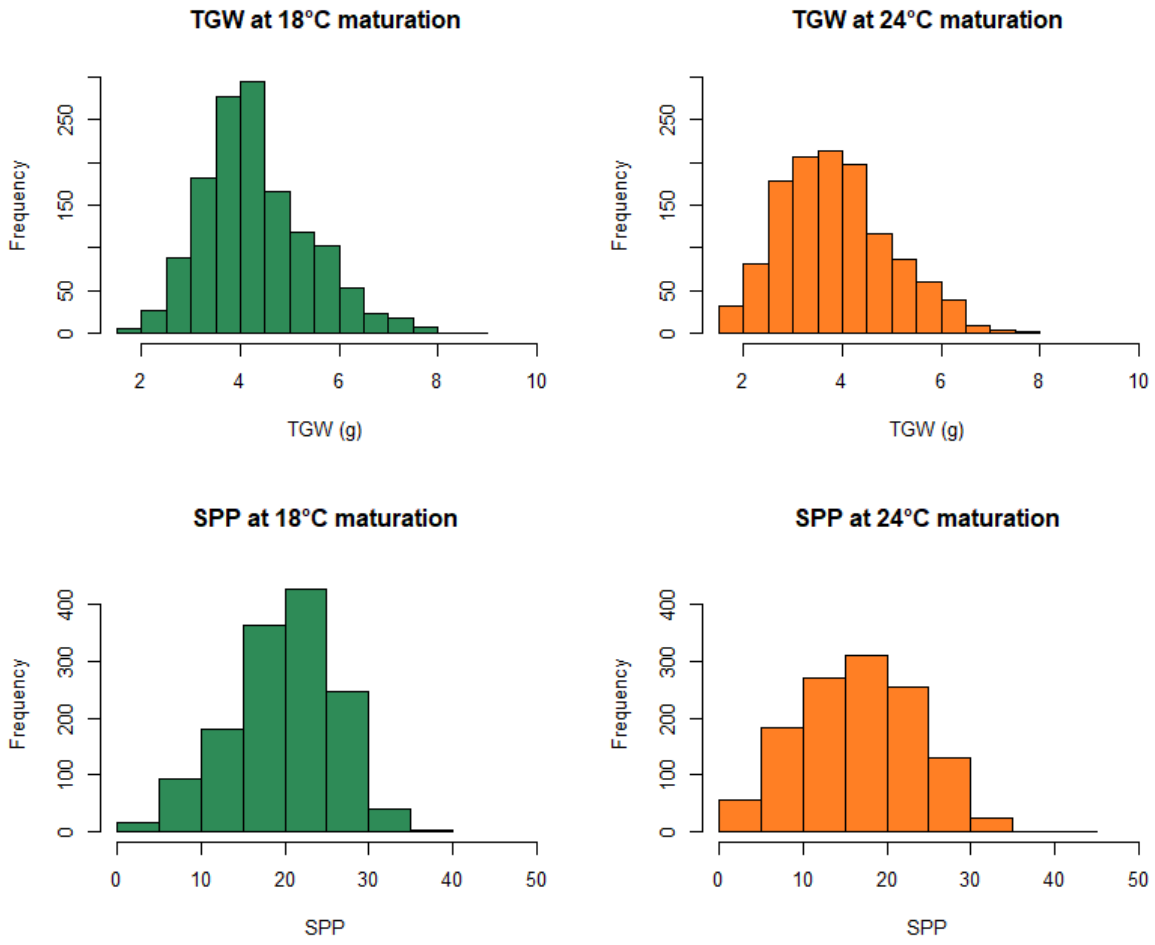

Figure 2.6: Histograms comparing the distribution of TGW (a,b) and SPP (c,d) over two different seed maturation temperatures.

Figure 2.7 shows histograms of TGW for each individual treatment. Comparing these graphs to the histograms for combined treatments can begin to indicate in greater detail how the combination of treatments affects TGW. For example, looking at TGW across all 18°C matured plants (Figure 2.6a) the maximum range is 8.5 – 9.0 g. Looking at the individual treatments matured at 18°C, it can be seen that TGW of 8.5 – 9.0 g is seen only in plants vernalised at 5°C, for either 6 or 12 weeks. Similarly, the maximum range

of TGW across all plants matured at 24°C was 7.5 – 8.0 g (Figure 2.6b). Plants contributing to this maximum range were only found in the treatments 5°C, 6 weeks, 24°C and 10°C, 12 weeks, 24°C. From Figure 2.7, the reduced number of plants which produced seeds in some treatments can be seen from the low total frequency, such as in the 15°C, 6 weeks, 24°C treatment.

Figure 2.8 shows histograms of SPP for each individual treatments. Comparing these graphs to the histograms for combined treatments can begin to indicate in greater detail how the combination of treatments affects SPP. When looking at all plants matured at 24°C, SPP at 24°C had a maximum range of 40 - 45 seeds (Figure 2.6d). From Figure 2.8, it can be seen that this range was only seen in plants grown at the 5°C, 12 weeks, 24°C treatment; for all treatments, the maximum range was 30 – 35 seeds.

In both Figures 2.7 and 2.8, the combination of vernalisation at 15°C (for either 6 or 12 weeks) and seed maturation at 24°C displayed the lowest range for most common frequency, with a TGW of 3.0 – 3.5 g and SPP of 10 – 15 seeds. This may be an indicator that combinations of higher temperatures throughout development are detrimental to traits contributing to increased yield. However, a more robust analysis is required to fully understand the temperature effects.

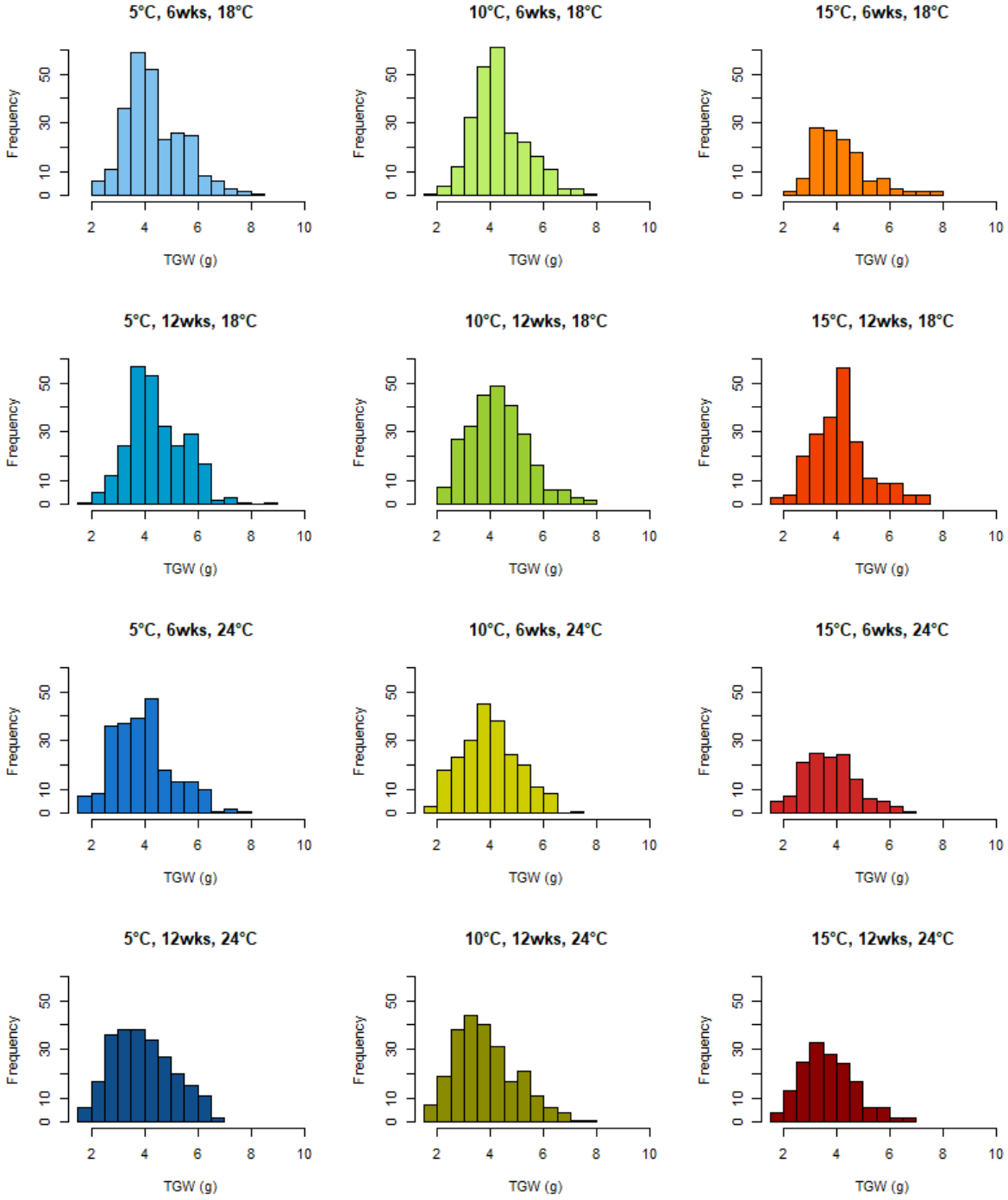

Figure 2.7: Histograms comparing TGW across all 12 temperature treatments used in the phenotyping experiment.

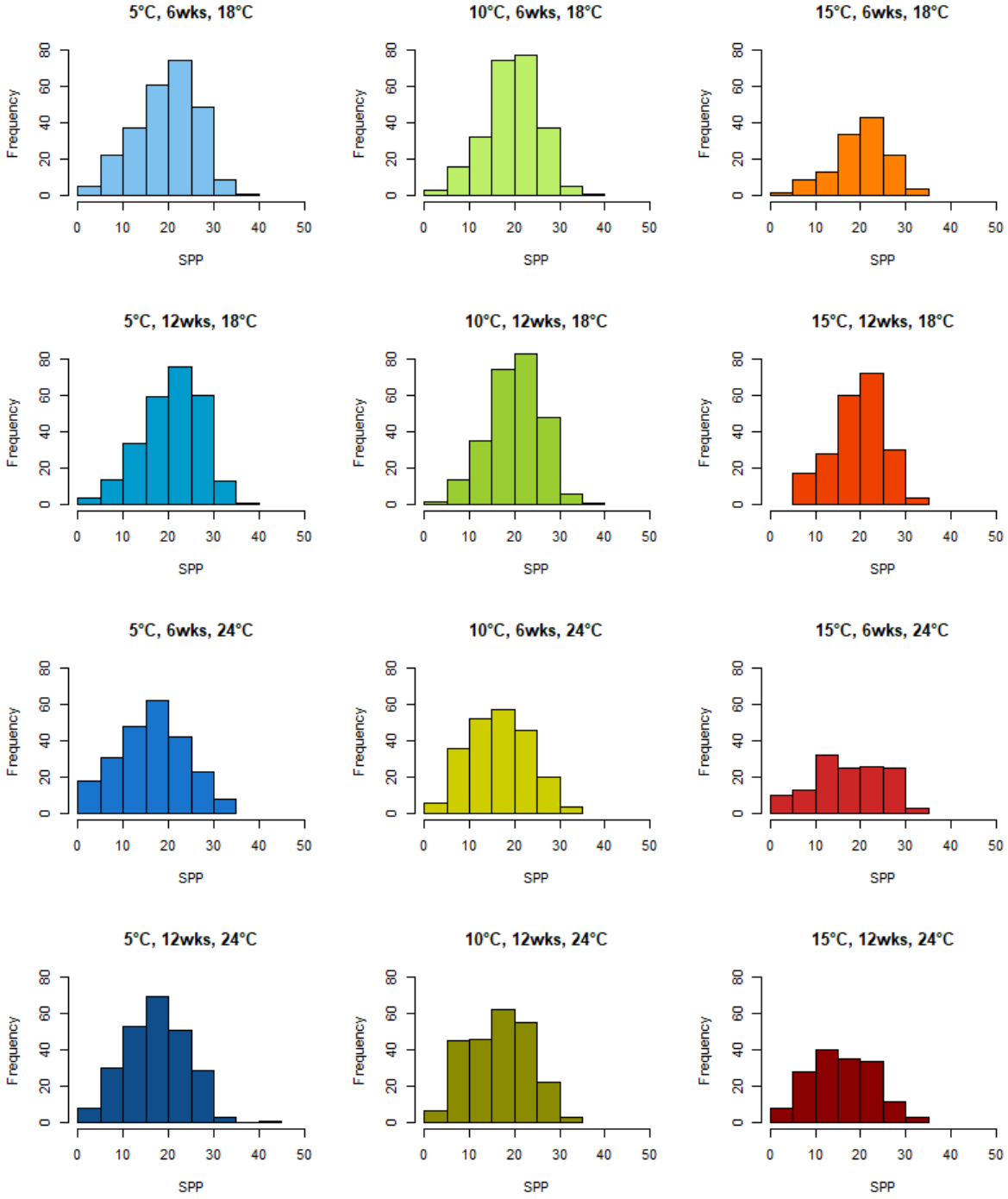

Figure 2.8: Histograms comparing SPP across all 12 temperature treatments used in the phenotyping experiment.

2.3.4 Linear modelling analysis of phenotyping results

In order to fully understand the interplay of vernalisation temperature, vernalisation duration and seed maturation temperature on seed yield parameters, a mixed linear model was used to estimate the overall impact of temperature.

2.3.4.1 Effect of blocking in phenotyping experiment

During this experiment plants were grown in an alpha-randomized incomplete block design. Statistical analysis was first used to examine the effect of block on TGW and SPP.

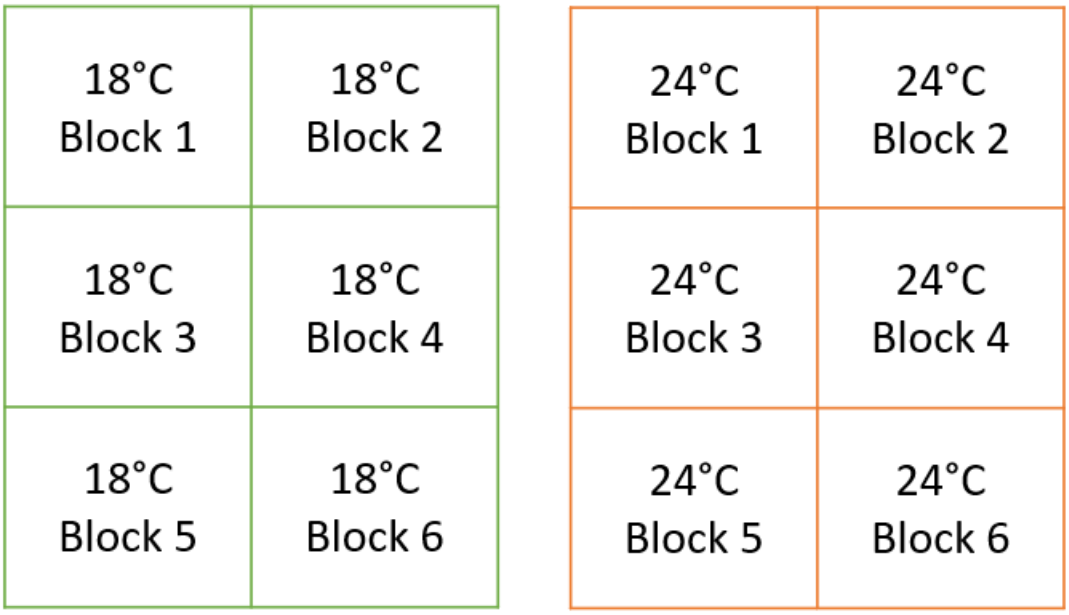

Figure 2.9: Blocking layout of glasshouse phenotype experiment.

The blocking pattern used in this experiment consisted of 12 blocks in total, with six in the 18°C glasshouse and six in the 24°C glasshouse (Figure 2.9). As a result of this, each block contained only a single maturation temperature. Plants were also split between blocks by vernalisation duration, with blocks 1, 3 and 5 containing plants vernalised for six weeks and blocks 2, 4 and 6 containing plants vernalised for twelve weeks. Each block then contained one of the three replicate plants from all 96 varieties from each of

the three vernalisation temperatures to a total of 288 plants per block, and the location of each plant was randomized within the block.

The effect of block on TGW can be seen in Figure 2.9a and 2.9b. The effect of block was considered using ANOVA. This ANOVA model estimated TGW or SPP (with trait referring to either TGW or SPP separately) in block (B) i, with the residual term e_i:

$$Trait_i = \mu + B_i + e_i$$

This analysis showed that block location had no impact on TGW at 18°C (Figure 2.10a) but did have an impact on TGW at 24°C, where it can be seen that Block 2 has the lowest mean TGW while Block 3 has the highest (Figure 2.10b). To account for this effect of location, block was included in further models for estimating TGW. There was no effect of block on SPP at either temperature (Figure 2.10c and 2.10d), so block was not included in further models estimating SPP.

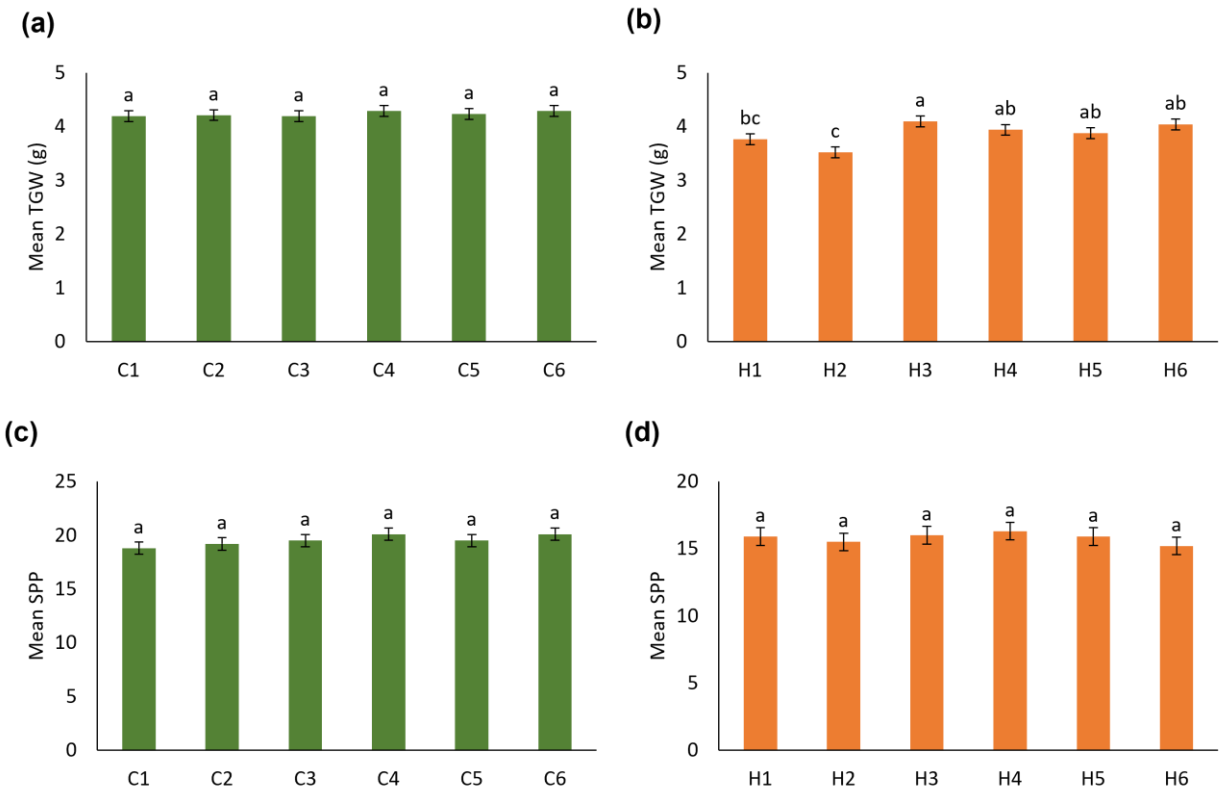

Figure 2.10: Effect of block on (a) TGW at 18°C (b) TGW at 24°C (c) SPP at 18°C (d) SPP at 24°C. Significance produced by ANOVA with Tukey post-hoc test, alpha = 0.05. In (a) and (c), C1-C6 indicates "cool glasshouse" blocks at 18°C; in (b) and (d), H1-H6 indicates "hot glasshouse" blocks at 24°C.

2.3.4.2 Optimizing the linear model

Linear modelling was carried out to test the effects of each factor in the treatments used in the phenotyping experiment on TGW and SPP. The linear models included maturation temperature, vernalisation temperature and vernalisation duration as fixed effects, as these were the factors being tested in the phenotyping experiment. In order to assess the effects of these treatments as accurately as possible, the linear models used to estimate TGW and SPP for each treatment were optimized.

2.3.4.2.1 Optimizing fixed effects terms used to estimate TGW and SPP

The formula for the mixed linear model used to estimate TGW was:

$$TGW_{ijklm} = \mu + MT_j + VT_k + VD_l + B_m + v_i + e_{ijklm}$$

This model predicts the TGW of variety (v) i in maturation temperature (MT) j in vernalisation temperature (VT) k and vernalisation duration (VD) l and block (B) m, where e_{ijklm} is the residual term. In this model the fixed terms are *MT*, *VT,* and *VD*, with *B* as a random variable and v as a random factor.

The formula for the mixed linear model used to estimate SPP was:

$$SPP_{ijkl} = \mu + MT_j + VT_k + VD_l + v_i + e_{ijkl}$$

This model predicts the SPP of variety (v) i in maturation temperature (MT) j in vernalisation temperature (VT) k and vernalisation duration (VD) l, where e_{ijkl} is the residual term. In this model the fixed terms are *MT*, *VT*, and *VD*, with v as a random factor.

Table 2.3: Results of linear model estimating TGW showing effects of each treatment. DF = degrees of freedom. DenDF = denominator degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Maturation Temperature	10.328	10.3275	1	8.99	20.4423	0.001447
Vernalisation Temperature	41.624	20.8119	2	2489.18	41.195	< 2.2e-16
Vernalisation Duration	0.09	0.0897	1	8.99	0.1776	0.6833

The results of the linear model detailed in Table 2.3 show that TGW across the diversity set is affected by vernalisation temperature (p < 2.2E-16). TGW is also affected by seed maturation temperature (p = 0.001447). However, vernalisation duration was not found to have a significant effect on TGW (p = 0.6833).

Table 2.4: Results of linear model estimating SPP showing effects of each treatment. DF = degrees of freedom. DenDF = denominator degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Maturation Temperature	9389.8	9389.8	1	2500.8	536.9984	< 2.2e-16
Vernalisation Temperature	615.7	307.8	2	2501.2	17.605	2.56E-08
Vernalisation Duration	38.5	38.5	1	2500.8	2.2018	0.138

The results of the linear model detailed in Table 2.4 show that SPP across the diversity set is affected by vernalisation temperature (p = 2.56E-8). TGW is also affected by seed maturation temperature (p < 2.2E-16). However, vernalisation duration was not found to have a significant effect on TGW (p = 0.138). As vernalisation duration was found to have no impact on TGW or SPP, it was removed from the models before further analysis.

2.3.4.2.2 Interaction effects when estimating TGW

Interactions between the fixed effects of the model estimating TGW were examined. This was done using the model:

$$TGW_{ijkl} = \mu + MT_j \times VT_k \times v_i + B_l + e_{ijkl}$$

This model predicts the TGW of variety (*v) i* in maturation temperature (*MT) j* in vernalisation temperature (*VT*) *k* and block (*B*) *l,* where e_{ijkl} is the residual term. In this model the fixed terms are *v, MT,* and *VT*, with *B* as a random variable and includes interactions between *v*, *MT,* and *VD*.

The results of this analysis showed that the fixed effects had significant interactions (Table 2.5). These interactions exist between maturation temperature and variety, vernalisation temperature ad variety, and maturation temperature and vernalisation temperature. A significant three-way interaction term was also identified between variety, maturation temperature and vernalisation temperature. As these interaction effects were identified as significant, they were included in the model for further analysis.

Table 2.5: Linear model output testing interactions between fixed effects when estimating TGW. DenDF = denominator degrees of freedom.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Variety x MT	215.51	2.2685	95	2050.7	5.9743	< 2.2e-16
Variety x VT	152.08	0.8544	178	2050.62	2.25	< 2.2e-16
MT x VT	2.93	1.467	2	2053.29	3.8635	0.021147
MT x VT x Variety	100.22	0.6303	159	2050.59	1.6599	1.21E-06

2.3.4.2.3 Interaction effects when estimating SPP

Interactions between the fixed effects of the model estimating SPP were examined. This was done using the model:

$$SPP_{ijk} = \mu + MT_j \times VT_k \times v_i + e_{ijk}$$

This model predicts the TGW of variety (v) i in maturation temperature (MT) j in vernalisation temperature (VT) k, where e_{ijk} is the residual term. In this model all terms are fixed and includes interactions between MT and VT.

The results of this analysis (Table 2.6) showed significant interactions between maturation temperature and variety as well as between vernalisation temperature and variety. No significant interaction was identified between maturation temperature and vernalisation temperature. Interactions identified as significant were included in model for further analysis.

Table 2.6: Linear model output testing interactions between fixed effects when estimating SPP. DF = degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Variety	95	65348	687.9	46.1525	< 2.2e-16
Maturation Temperature (MT)	1	9401	9401	630.751	< 2.2e-16
Vernalisation Temperature (VT)	2	611	305.6	20.5065	1.52E-09
Variety x MT	95	5622	59.2	3.9702	< 2.2e-16
Variety x VT	178	4594	25.8	1.7317	3.33E-08
MT x VT	2	16	8.1	0.5437	0.5807
Variety x MT x VT	159	2703	17	1.1407	0.1179
Residuals	2063	30748	14.9		

2.3.4.2.4 Assessing optimized models used to estimate TGW and SPP

The optimized formula for the mixed linear model used to estimate TGW was:

$$TGW_{ijkl} = \mu + MT_j + VT_k + v_i + (MT_j \times VT_k) + (MT_j \times v_i) + (VT_k \times v_i) + (MT_j \times VT_k \times v_i) + B_l + e_{ijkl}$$

This model predicts the TGW of variety (v) i in maturation temperature (MT) j in vernalisation temperature (VT) k and block (B) l, where e_{ijkl} is the residual term. In this model the fixed terms are v, MT, and VT, with B as a random variable. Interactions are included between MT and VT, MT and v, and VT and v, as well as between MT, VT, and v.

The results of the updated linear model detailed in Table 2.7 show that TGW across the diversity set is affected by vernalisation temperature ($p < 2.2E-16$). TGW is also affected by seed maturation temperature ($p = 0.000843$). Using the Nakagawa and Schielzeth method (2013) of calculating R-squared statistics for mixed linear models suggested that this model explains 71.6% of variation in TGW. Figure 2.11 shows a scatterplot of the values for TGW predicted by the linear model against the actual TGW values measured. The model-predicted values align generally with the observed values, although at higher values the model more often underestimates TGW.

Table 2.7: Results of updated linear model estimating TGW showing effects of each factor. DF = degrees of freedom. DenDF = denominator degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method.

	Sum Sq	Mean Sq	DF	DenDF	F value	Pr(>F)
Variety	1494.57	15.7323	95	2050.73	41.4315	< 2.2e-16
Maturation Temperature (MT)	9.3	9.3009	1	10.63	24.4942	0.000483
Vernalisation Temperature (VT)	29.98	14.9922	2	2053.17	39.4824	< 2.2e-16
Variety x MT	215.51	2.2685	95	2050.7	5.9743	< 2.2e-16
Variety x VT	152.08	0.8544	178	2050.62	2.25	< 2.2e-

						16
MT x VT	2.93	1.467	2	2053.29	3.8635	0.021147
MT x VT x Variety	100.22	0.6303	159	2050.59	1.6599	1.21E-06

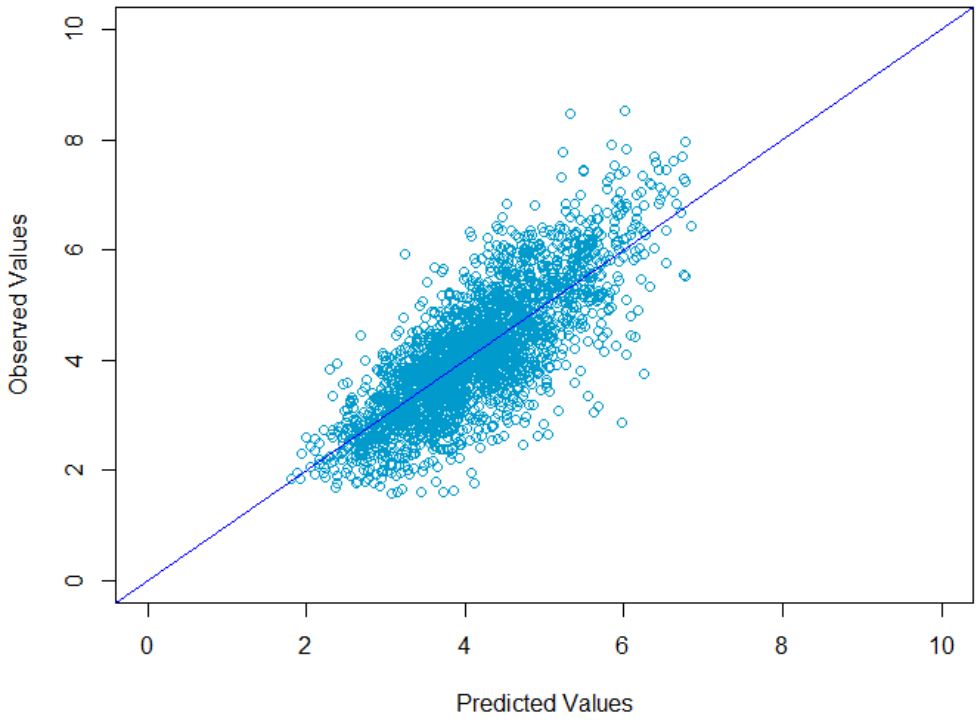

Figure 2.11: Predicted vs actual values for the updated linear model used to predict TGW.

The optimized formula for the general linear model used to estimate SPP was:

$$SPP_{ijk} = \mu + MT_j + VT_k + v_i + (MT_j \times v_i) + (VT_j \times v_i) + e_{ijk}$$

This model predicts the SPP of variety (v) i in maturation temperature (MT) j in vernalisation temperature (VT) k where $e_{ijk\ is}$ the residual term. In this model all terms are fixed and includes interactions between MT and v, and VT and v.

Table 2.8: Results of updated linear model estimating SPP showing effects of each factor. DF = degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Variety	95	65348	687.9	45.7113	< 2.2e-16
Maturation Temperature	1	9401	9401	624.7211	< 2.2e-16
Vernalisation Temperature	2	611	305.6	20.3104	1.82E-09
Variety x MT	95	5622	59.2	3.9323	< 2.2e-16
Variety x VT	178	4594	25.8	1.7151	5.17E-08
Residuals	2224	33467	15		

The results of the updated linear model detailed in Table 2.8 show that SPP across the diversity set is affected by vernalisation temperature (p = 2.92E-08). SPP is also affected by seed maturation temperature (p <2.2E-16). Using the Nakagawa and Schielzeth method (2013) of calculating R-squared statistics for mixed linear models suggested that this model explains 68.7% of variation in SPP. Figure 2.12 shows a scatterplot of the values for SPP predicted by the linear model against the actual SPP values measured. The model-predicted values align generally with the observed values, although the values have a wider spread at lower ranges of SPP, with the model overestimating some SPP values which were found to be low in actual measurements.

These models show that on average for the 96 lines included in the study, both TGW and SPP were found to be affected by seed maturation temperature and vernalisation temperature; however, vernalisation duration was not found to have an effect. These predictions made by the models for TGW and SPP will be explored in further sections.

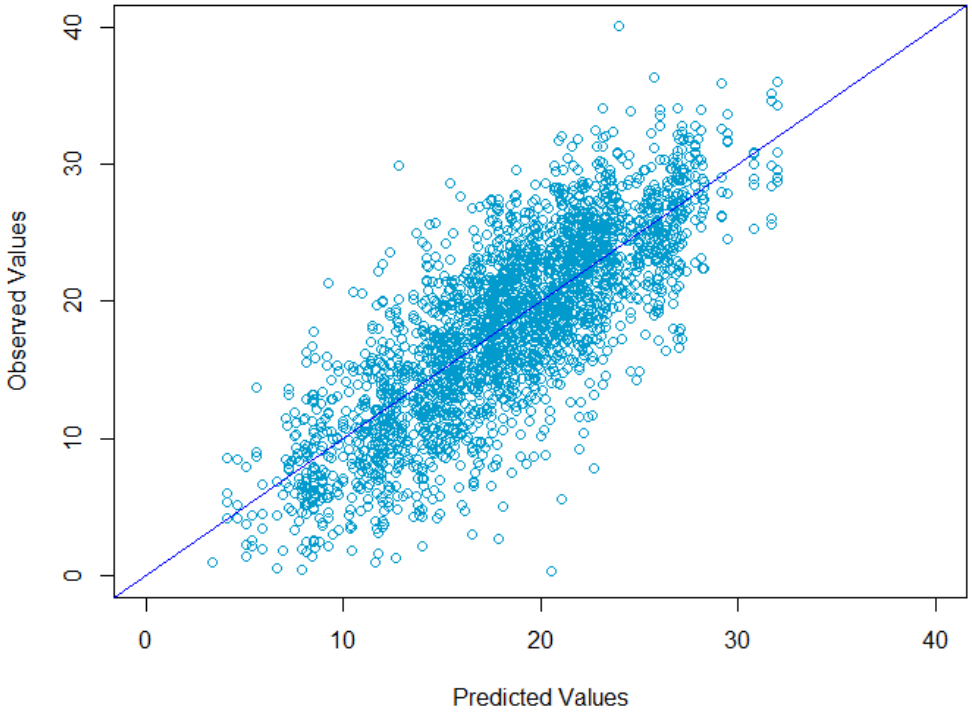

Figure 2.12: Predicted vs actual values from the updated linear model used to predict SPP.

2.3.4.3 Duration of vernalisation treatment does not affect seed traits

To test if vernalisation duration has an impact on yield components, two vernalisation treatments were compared between 6 weeks and 12 weeks. When considering the effects of vernalisation temperature and maturation temperature using a mixed linear model, vernalisation length had no significant effect on TGW or SPP across the diversity set (Figure 2.13).

In Table 2.9, it can be seen that the mean values for SPP and TGW were extremely similar and had no statistically significant difference between six and twelve weeks of

vernalisation. As vernalisation duration had no significant effect on yield components it was removed from the models.

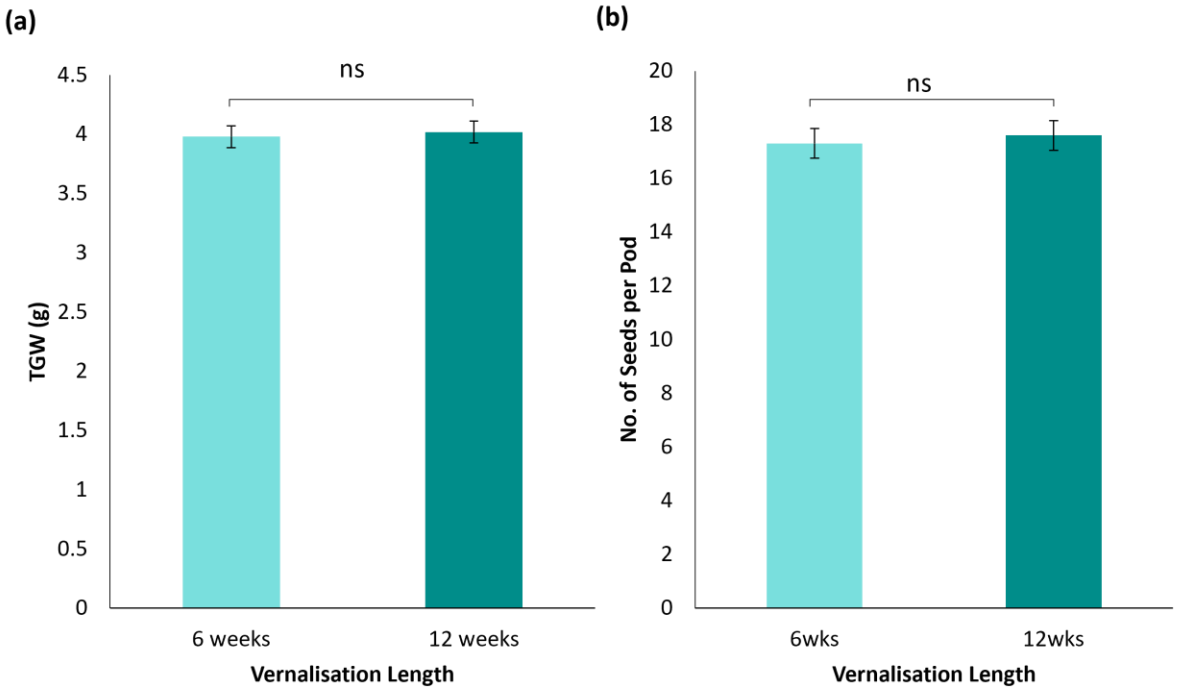

Figure 2.13: Estimated means from linear model of(a) TGW and (b) SPP, by vernalisation length. Error bars show SE and annotations show significant differences and p-values producing using Welch's t-test; ns = not significant.

Table 2.9: Details of estimated means for TGW and SPP from linear model. All values shown to 3 significant figures. Data for TGW is measured in g. Data for SPP is count data. SE = standard error. df = degrees of freedom. CL = 95% confidence limit.

	SPP					***TGW***				
Vernalisation Duration	**Mean**	**SE**	**df**	**Lower CL**	**Upper CL**	**Mean**	**SE**	**df**	**Lower CL**	**Upper CL**
6 weeks	17.3	0.551	101	16.2	18.4	4.02	0.11	56.5	3.8	4.24
12 weeks	17.6	0.548	99	16.5	18.7	3.98	0.11	57.4	3.76	4.2

2.3.4 Impact of vernalisation temperature on seeds

Three vernalisation temperatures were used in this experiment to explore the impact of vernalisation temperature on seed yield parameters. Using a mixed linear model, it was possible to see the effect of each of these vernalisation temperatures on yield parameters across the 96 *Brassica napus* accessions studied, independently of the other features of the treatment undergone by specific plants. The estimated mean TGW for each vernalisation temperature is shown in Figure 2.14 and Table 2.10.

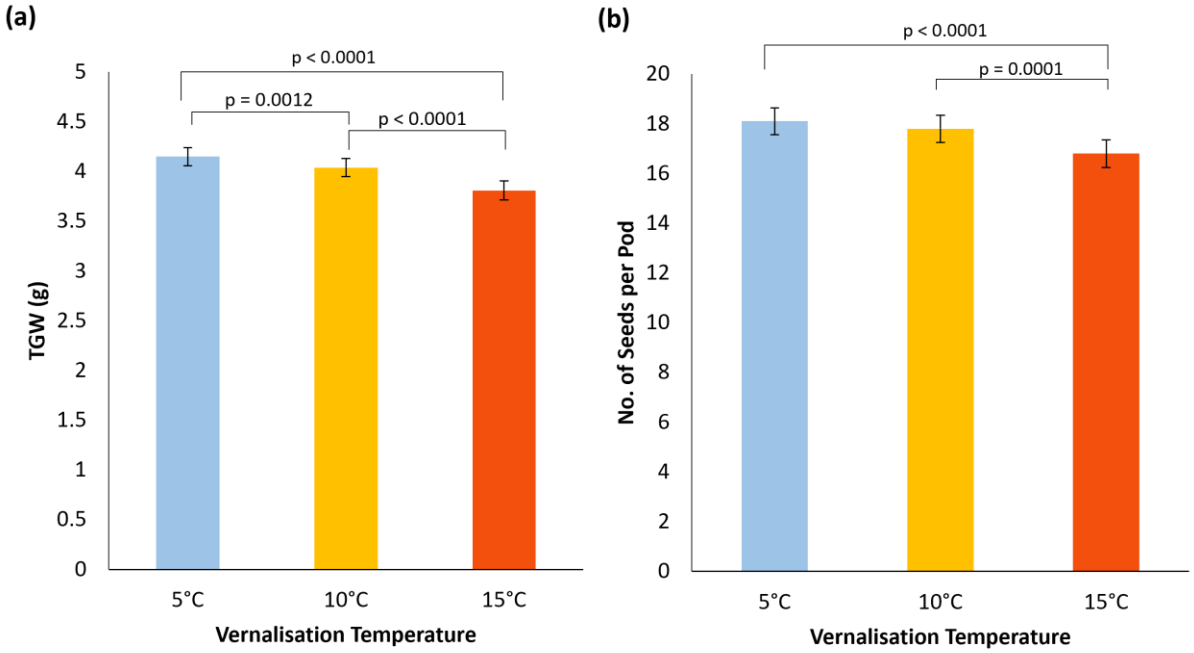

Figure 2.14: Estimated means from linear model for (a) TGW and (b) man SPP, by vernalisation temperature. Error bars show SE and annotations show significant differences and p-values calculated using Welch's t-test.

Table 2.10: Details of linear model output for TGW and SPP. All values shown to 3 significant figures. Data for TGW is measured in g. Data for SPP is count data. SE = standard error. df = degrees of freedom. CL = 95% confidence limit.

	SPP					***TGW***				
Vernalisation Temperature	**Mean**	**SE**	**df**	**Lower CL**	**Upper CL**	**Mean**	**SE**	**df**	**Lower CL**	**Upper CL**
5°C	18.1	0.543	102	17	19.2	4.15	0.101	97.7	3.95	4.35
10°C	17.8	0.544	103	16.7	18.9	4.04	0.101	98.1	3.84	4.24
15°C	16.8	0.555	111	15.8	17.9	3.82	0.103	105.1	3.61	4.02

As can be seen in Table 2.10, increasing vernalisation temperature from 5°C to 15°C was associated with a reduction in TGW of 0.34g (8.2%) and 1.3 fewer seeds per pod (7.2%). Higher temperatures during vernalisation led to smaller seeds in terms of TGW (Figure 14a) and fewer seeds per pod (Figure 14b). Although the overall effect is significant, the effect size across the diversity set is small. The effect of increasing temperature from 5°C to 10°C is extremely limited, with SPP falling by 0.3 (1.7%) and TGW falling by 0.11g (2.7%). The negative impact becomes greater when vernalisation temperature is increased above 10°C. When vernalisation temperature was increased from 10°C to 15°C, plants produced seeds which were 0.23g smaller (5.7%) and pods had 1 seed less on average (5.6%). As a result, I concluded that plants which experienced cooler temperatures produced both more seeds per pod and seeds with a greater TGW, and that the negative effect of warmer vernalisation is increased at temperatures above 10°C.

2.3.4.1 Effects of vernalisation temperature for individual varieties

As well as estimating the effect size across the full *Brassica napus* population used in this phenotyping experiment, the differences in estimated values for TGW and SPP between vernalisation temperatures can be considered for individual varieties. Although the overall effect size of vernalisation temperature on TGW and SPP was small, the impact varied between individual varieties. Some varieties experienced little or no

effects, while others showed much larger effects than the overall estimate. While the overall trend across the diversity set suggests that vernalizing *Brassica napus* at 5°C compared to 15°C results in plants producing on average 1 more seed per pod and that seeds are 0.23g larger, some lines showed little to no differences between vernalisation temperature while other showed much greater differences, or produced more or larger seeds at 15°C compared to 5°C.

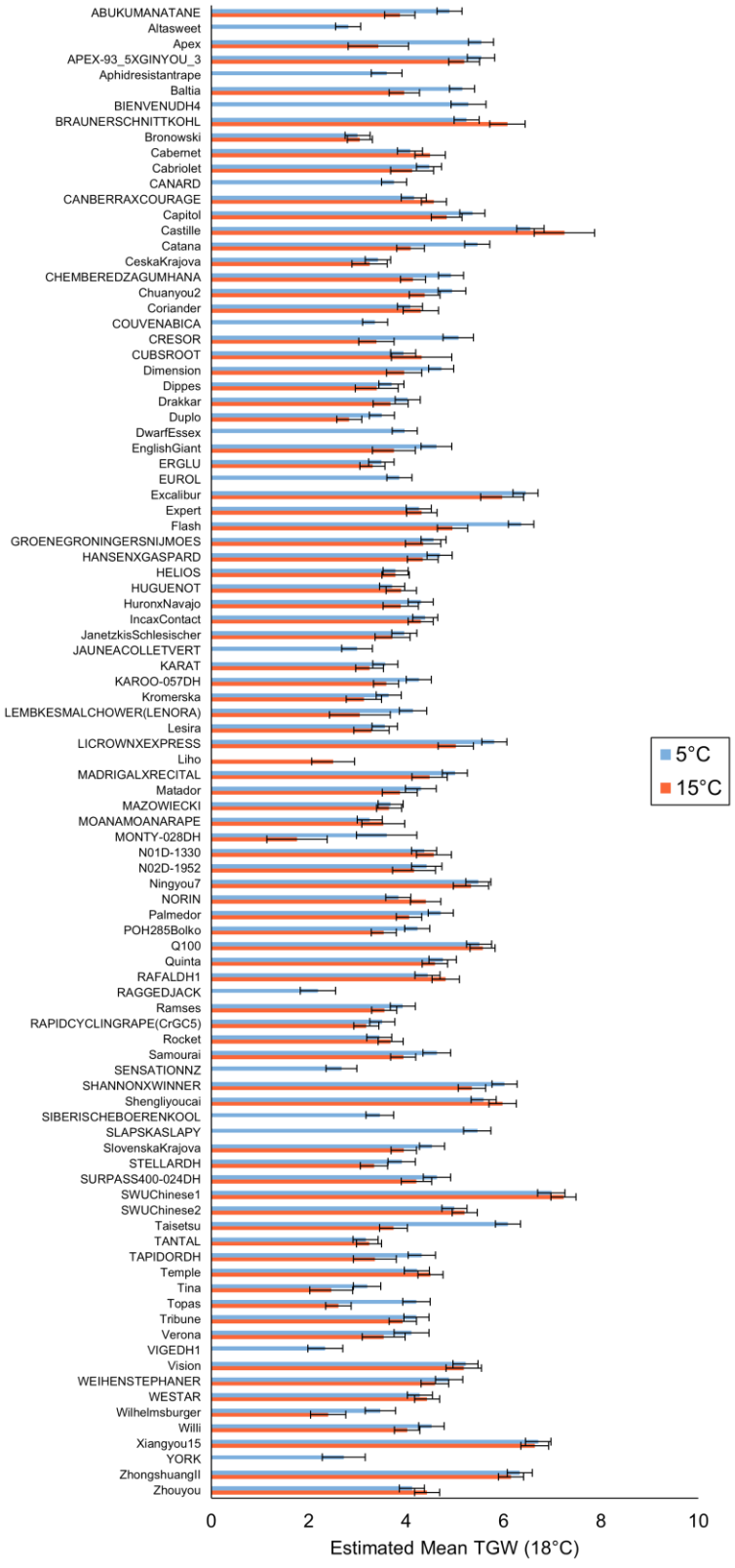

Figure 2.15: Estimated mean TGW from linear model for each variety at 18°C maturation temperature, comparing vernalisation treatments at 5°C and 15°C. Error bars show standard error.

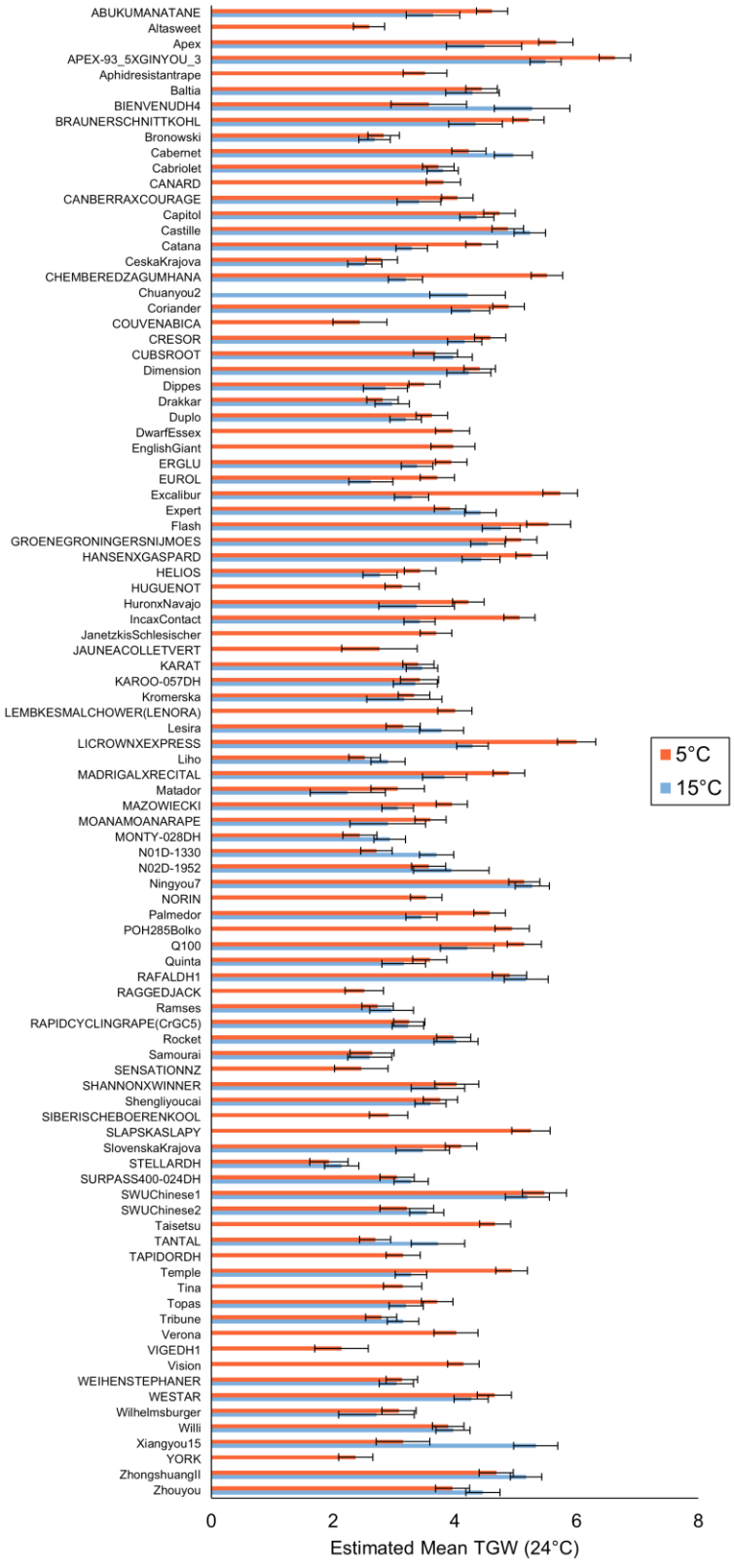

Figure 2.16: Estimated mean TGW from linear model for each variety at 24°C maturation temperature, comparing vernalisation treatments at 5°C and 15°C. Error bars show standard error.

Due to missing data, means could not be estimated for every variety under all treatment conditions examined here. While these varieties are still included in the graphical summary, they were not considered when identifying varieties with the greatest differences between treatments.

Figures 2.15 and 12.6 show the range of TGW across all 96 varieties as estimated by the linear model, compared between vernalisation temperatures, for plants grown in both maturation temperature treatments. At 18°C, there are some lines which have larger seeds when vernalised at 15°C such as the exotic variety Brauner Schnittkohl, which showed the largest increase at the warmer vernalisation temperature being 0.84g. Comparatively at 24°C, lines such as the Semi-winter OSR variety Xiangyou 15 showed a larger increase when vernalised at 15°C, with a 2.18g increase over the 5°C vernalisation treatment. A number of lines showed very small differences in TGW between vernalisation temperatures, including a number of Spring OSR varieties such as Bronowski and Mazowiecki. At 24°C, the winter variety Rocket was also not strongly affected by vernalisation temperature, indicating that there is variation in the effect of vernalisation on TGW within crop types.

Figures 2.17 and 2.18 show the range of SPP across all 96 varieties as estimated by the linear model, compared between vernalisation temperatures, for plants grown in both maturation temperature treatments. The varieties with the largest increase in SPP at 15°C vernalisation compared to 15°C were the Spring OSR N02D-1952 and the Semi-winter OSR Xiangyou 15, with an estimated increase of 4 – 5 seeds per pod. The effect size was greater when considering the variety with the largest increase associated with 5°C vernalisation was the Winter OSR variety Samourai, with an expected increase of 9 seeds per pod.

These results indicate that while there was an overall small effect of vernalisation temperature on TGW and SPP, the response of individual varieties varied. Particularly

for TGW these effects varied based on seed maturation temperature due to the interaction of vernalisation and maturation temperature, suggesting that seed weight is sensitive to temperatures throughout development. This suggests that the genetic diversity within the DFFS set used here represents a range of temperature responses which can be further explored.

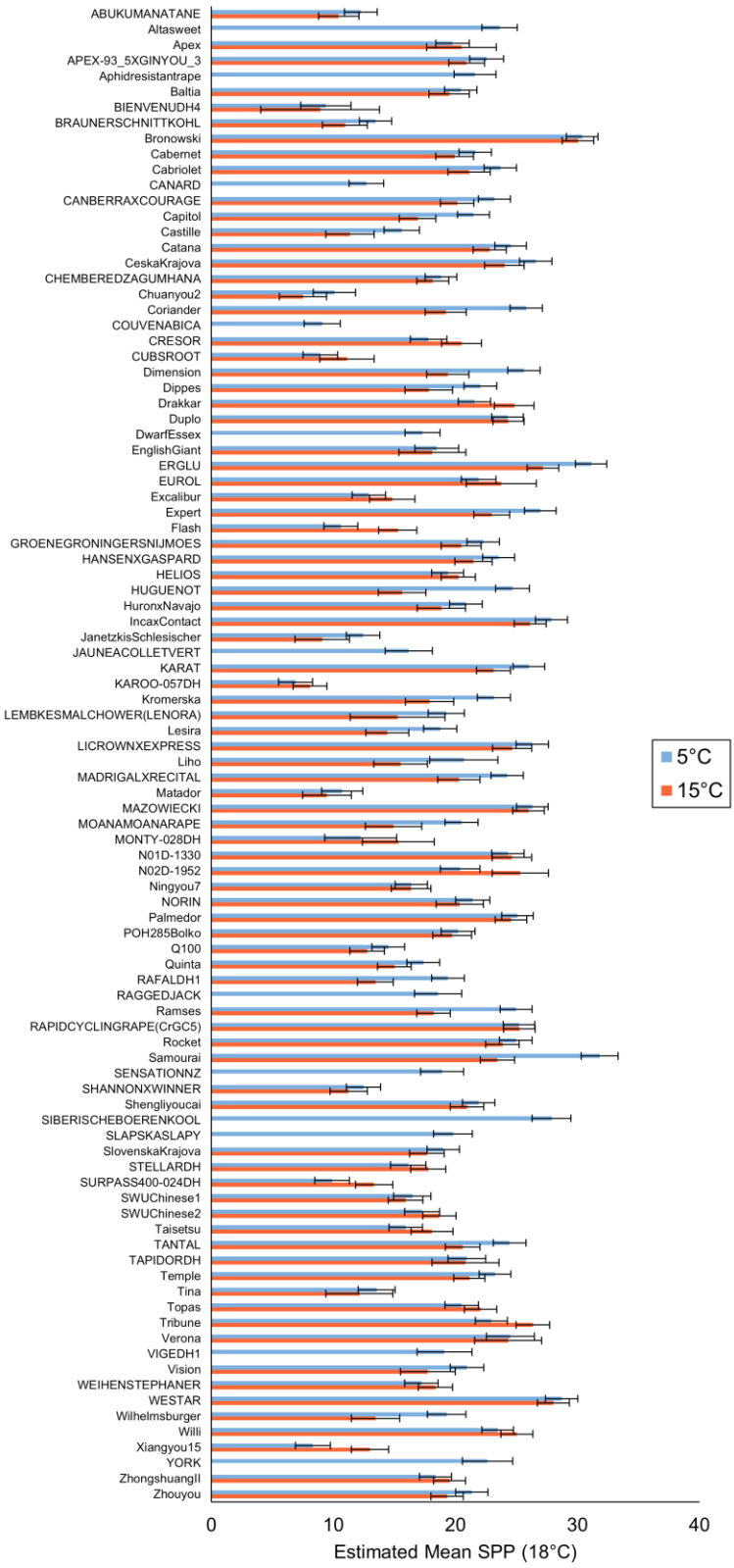

Figure 2.17: Estimated mean SPP from linear model for each variety at 18°C maturation temperature, comparing vernalisation treatments at 5°C and 15°C. Error bars show standard error.

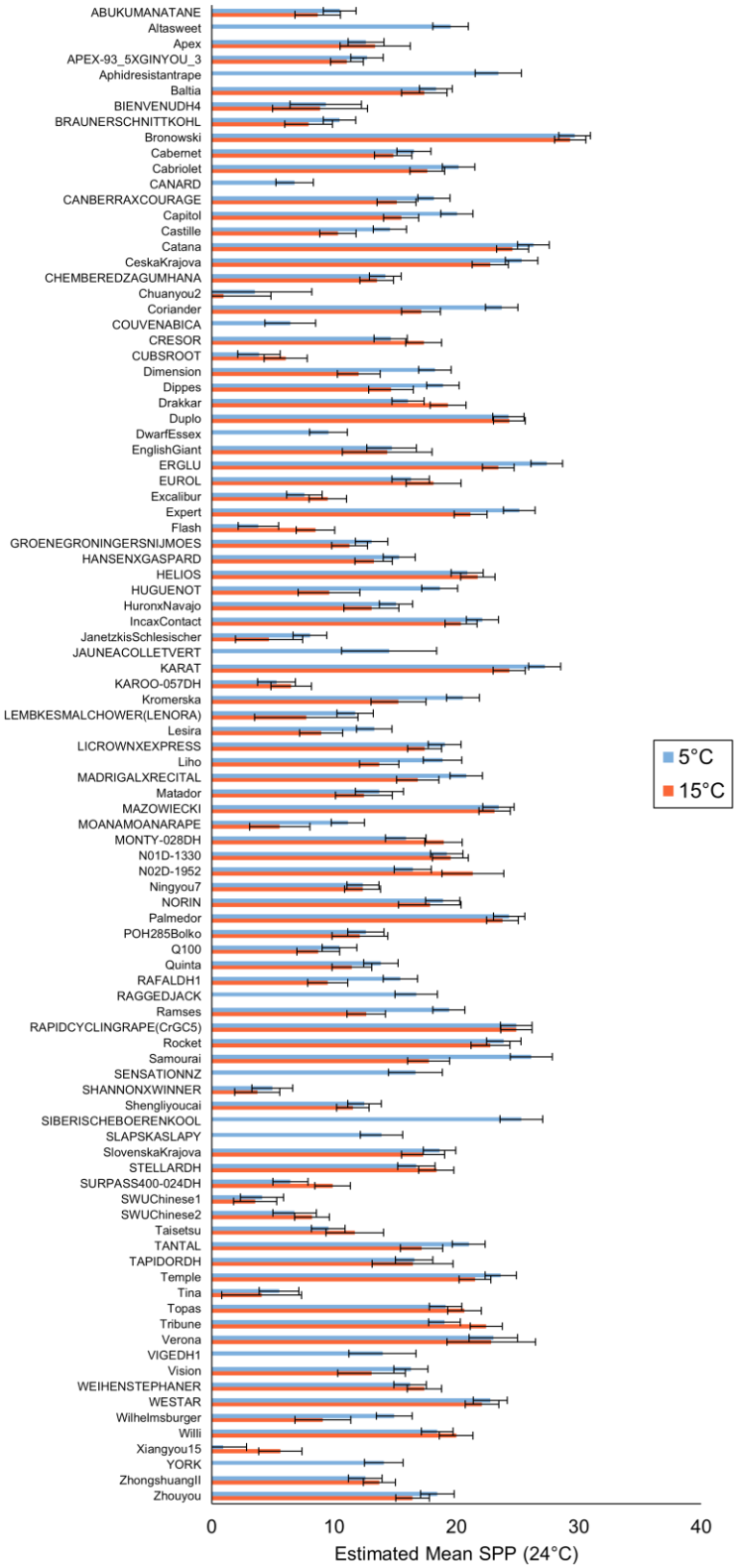

Figure 2.18: Estimated mean SPP from linear model for each variety at 24°C maturation temperature, comparing vernalisation treatments at 5°C and 15°C. Error bars show standard error.

2.3.5 Impact of maturation temperature on seeds

Two seed maturation temperatures were compared in this experiment – a cooler temperature of 18°C and a warm temperature of 24°C. The effects of seed maturation temperature were explored independently of vernalisation treatment by using a mixed linear model. Seed maturation temperature here refers to the temperature plants were moved to after the vernalisation treatment was completed, and so also refers to temperature during seed set.

Considering the estimated means, increasing seed maturation temperature from 18°C to 24°C caused a drop in TGW of 0.41g (10%) and was associated with an average of 3.9 (20%) fewer seeds per pod (Table 2.11). Higher temperatures during seed maturation were associated with a reduction in both yield parameters studied in this phenotyping experiment (Figure 2.19).

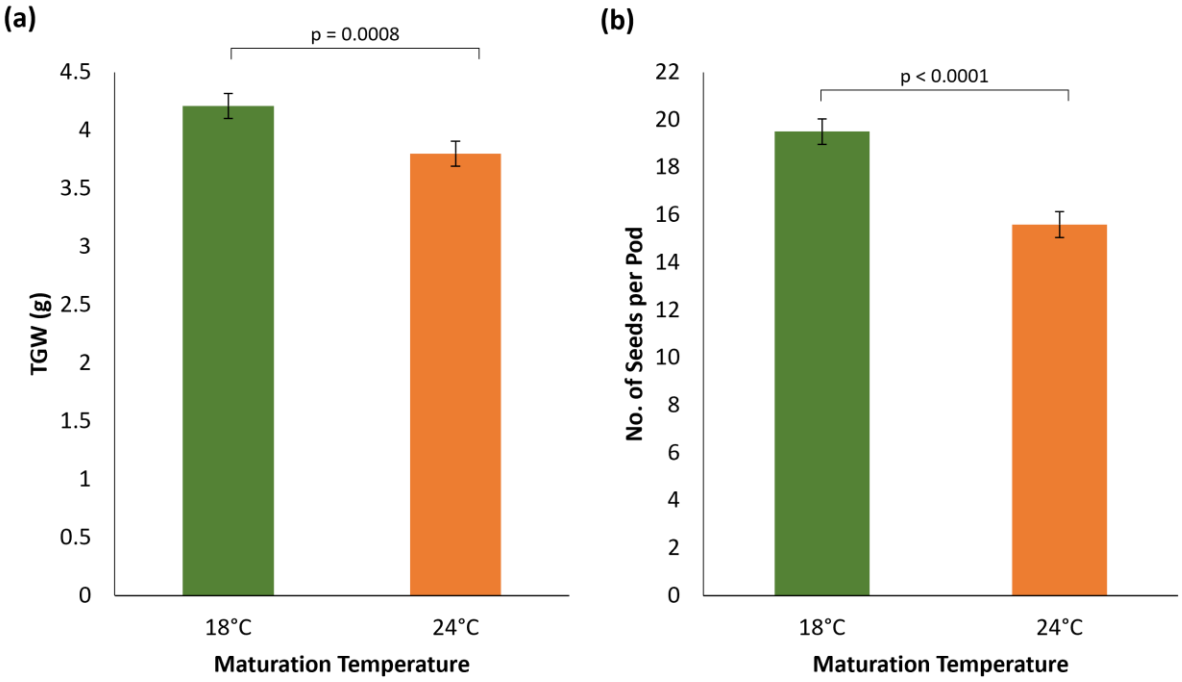

Figure 2.19: Estimated means of (a) TGW and (b) SPP, by seed maturation temperature. Error bars show SE and annotations show significant differences and p-values as calculated using Welch's t-test.

Table 2.11: Details of estimated means for TGW and SPP from linear model. All values shown to 3 significant figures. Data for TGW is measured in g. Data for SPP is count data. SE = standard error. df = degrees of freedom. 95% CL = 95% confidence limit.

	SPP					***TGW***				
Maturation Temperature	**Mean**	**SE**	**df**	**Lower CL**	**Upper CL**	**Mean**	**SE**	**df**	**Lower CL**	**Upper CL**
18°C	19.5	0.539	99.4	18.4	20.6	4.21	0.108	65	3.99	4.42
24°C	15.6	0.541	100.7	14.6	16.7	3.8	0.108	65.6	3.58	4.01

2.3.4.1 Effects of maturation temperature for individual varieties

As well as estimating the effect size across the full *Brassica napus* population used in this phenotyping experiment, the differences in estimated values for TGW and SPP between maturation temperatures can be considered for individual varieties. Due to limitations of the model in in accounting for data which could not be collected from the phenotyping experiment, means could not be estimated for every variety under all treatment conditions examined. As a result, these varieties were not considered when identifying varieties with the greatest differences between treatments. When evaluating the effects of maturation temperature on TGW and SPP, the effect size was larger than that seen by changing vernalisation temperature.

The general estimate across the diversity set indicated that a *Brassica napus* plant grown at 18°C would have on average 3.9 more seeds per pod and that those seeds would be 0.41g heavier. However, the effects varied between individual varieties, with some experiencing only small effects and some varieties experiencing much larger effects than the overall estimate. Additionally, some varieties were estimated to produce more seeds per pod or heavier seeds at 24°C compared to 18°C in contrast to the overall trend.

Figures 2.20 - 2.22 show the range of TGW across all 96 varieties for each maturation temperature, for plants grown in all three vernalisation temperature treatments.

While the overall model predicted that TGW would be greater in plants matured at 18°C, across all vernalisation temperatures some varieties such as Apex X Ginyou, Apex, Coriander and Hansen X Gaspard – all Winter OSR varieties – were estimated to produce heavier seeds at 24°C. Varieties with the greatest estimated increase in seed weight at the lower maturation temperature of 18°C included Winter and Semi-winter OSR varieties such as Shengliyoucai, Xiangyou 15, Samourai and Excalibur. These results indicate a large degree of phenotypic diversity even within the crop types assessed in this experiment, as well as across the diversity set.

Figures 2.23 - 2.25 show the range of SPP across all 96 varieties as estimated by the linear model, compared between maturation temperatures, for plants grown in all three vernalisation temperature treatments.

Spring OSR variety Monty was estimated to produce 3 – 4 additional seeds per pod at 24°C compared to 18°C. In contrast, the Semi-winter varieties SWU Chinese 1 and SWU Chinese 2 were strongly affected by maturation temperature, being estimated to produce 10 – 12 more seeds when matured at 18°C compared to 24°C.

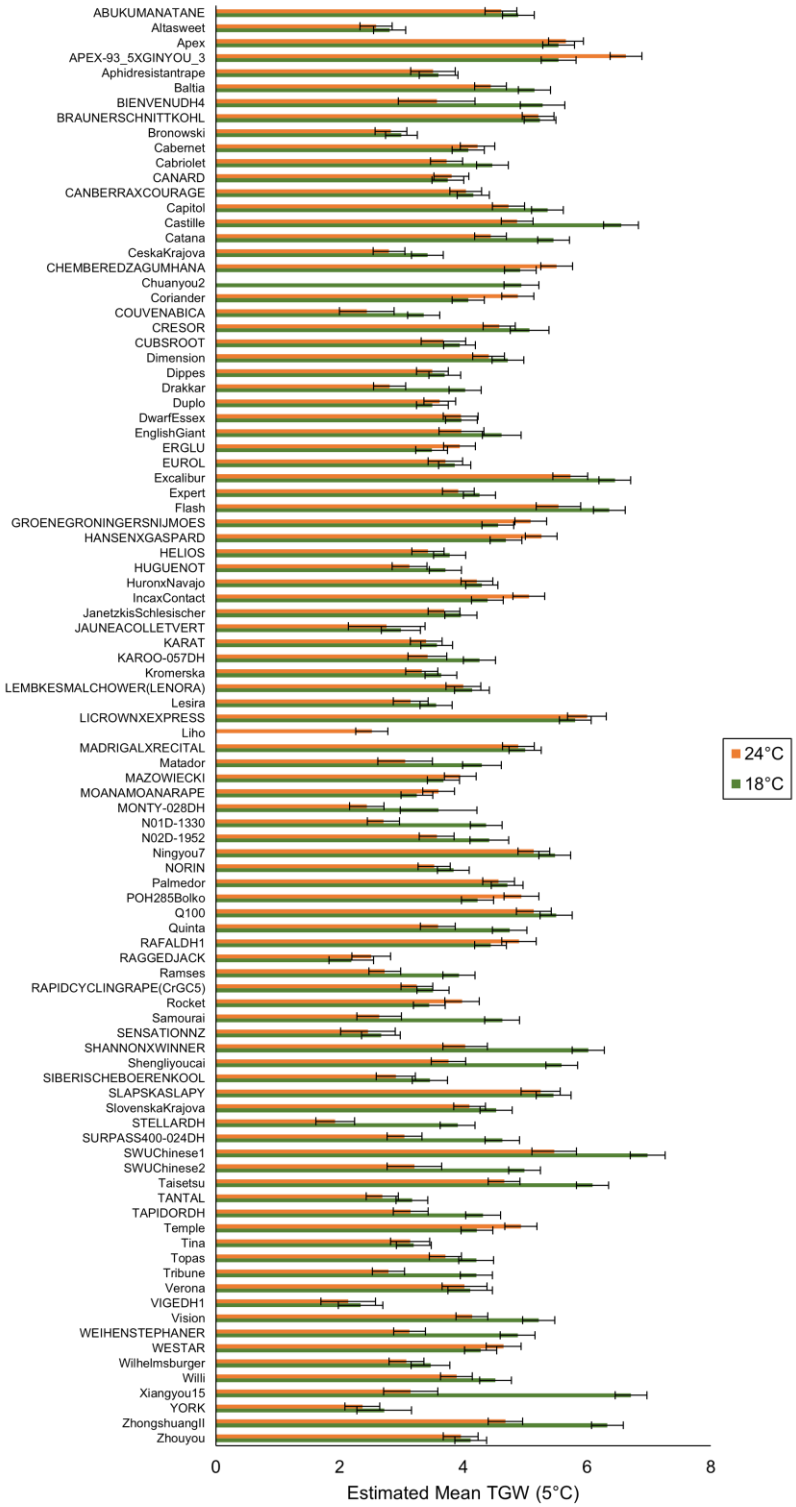

Figure 2.20: Estimated mean TGW from linear model for each variety at 5°C vernalisation temperature, comparing maturation treatments at 18°C and 24°C. Error bars show standard error.

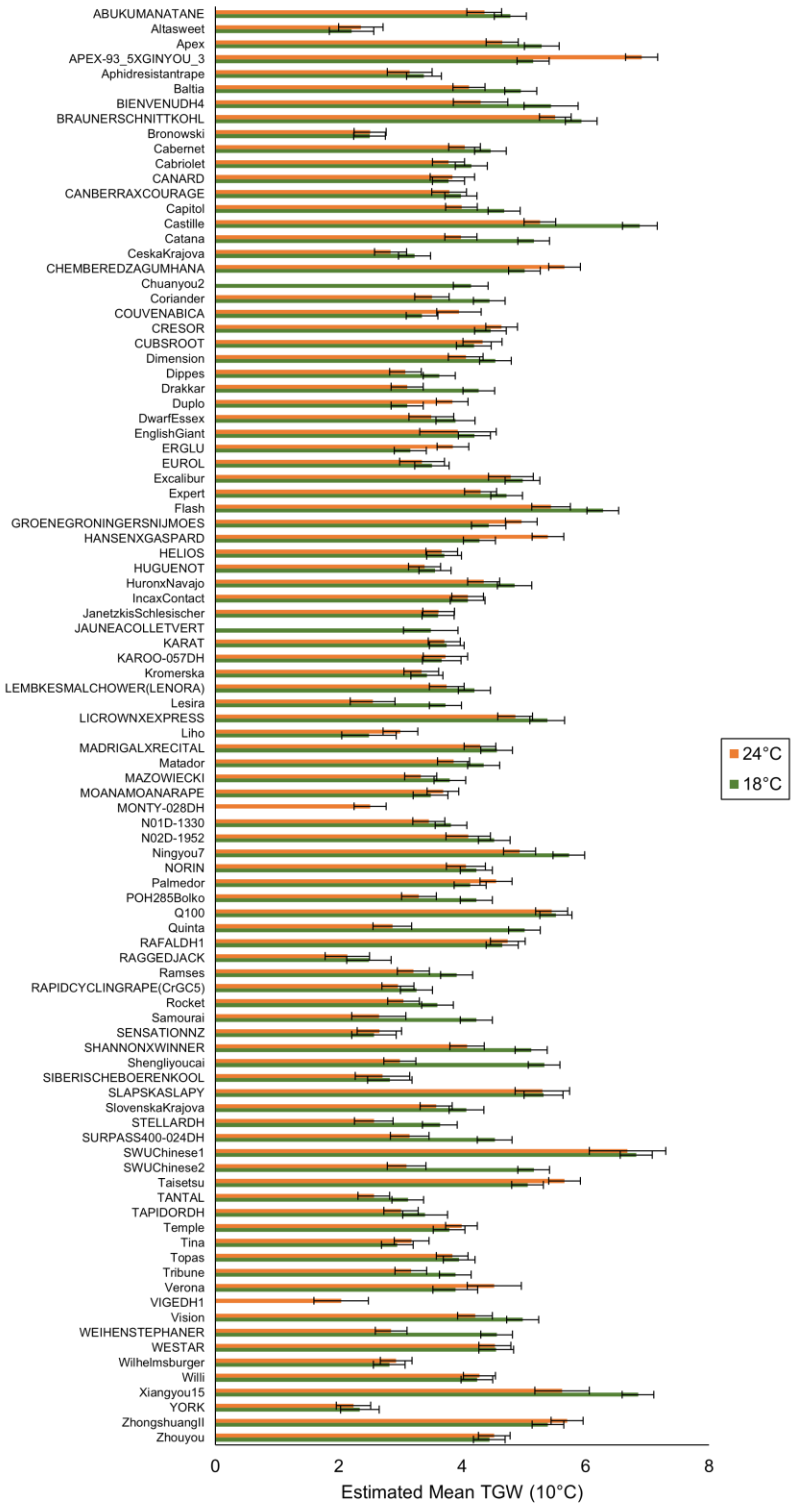

Figure 2.21: Estimated mean TGW from linear model for each variety at 10°C vernalisation temperature, comparing maturation treatments at 18°C and 24°C. Error bars show standard error.

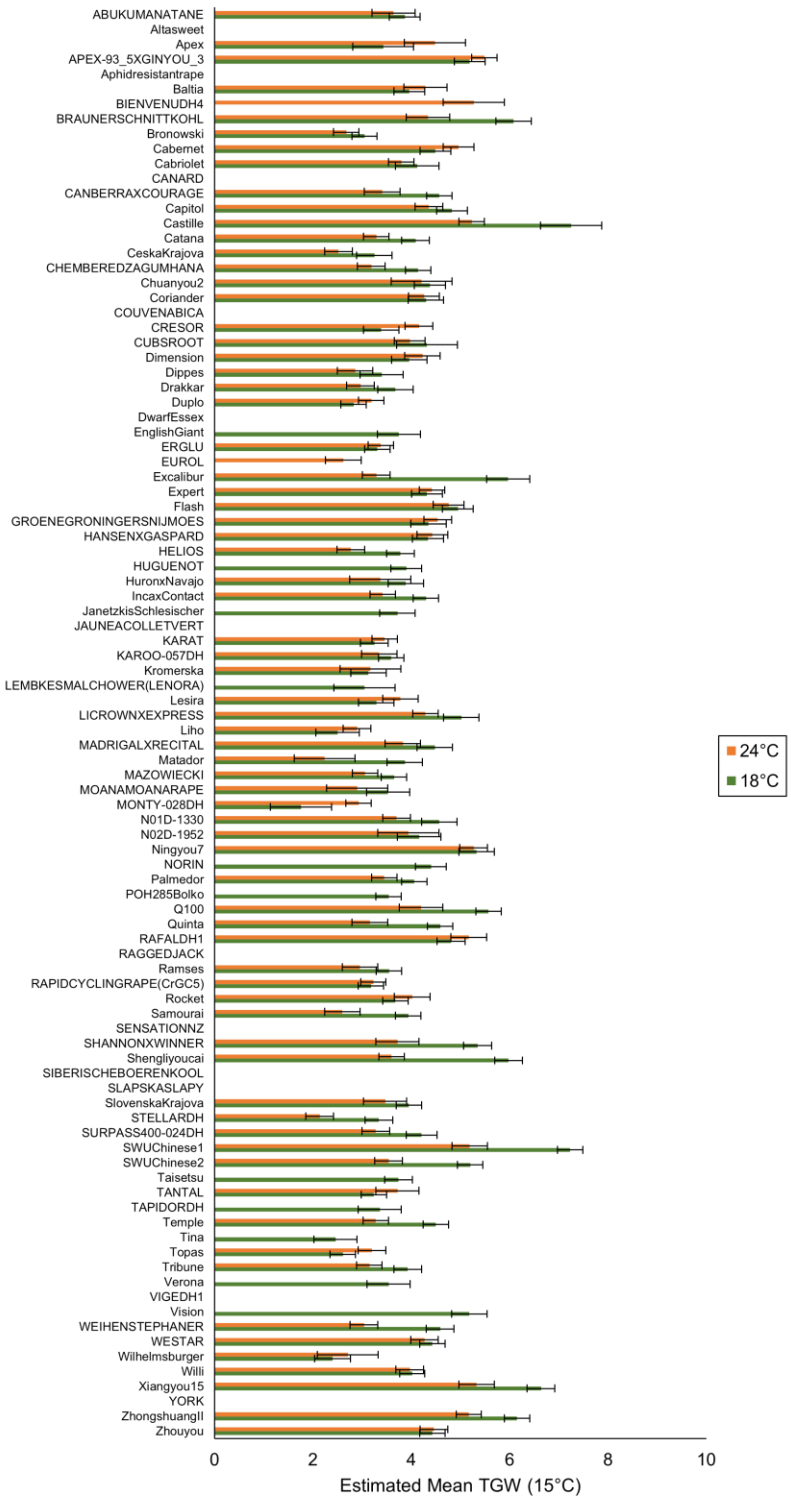

Figure 2.22: Estimated mean TGW from linear model for each variety at 15°C vernalisation temperature, comparing maturation treatments at 18°C and 24°C. Error bars show standard error.

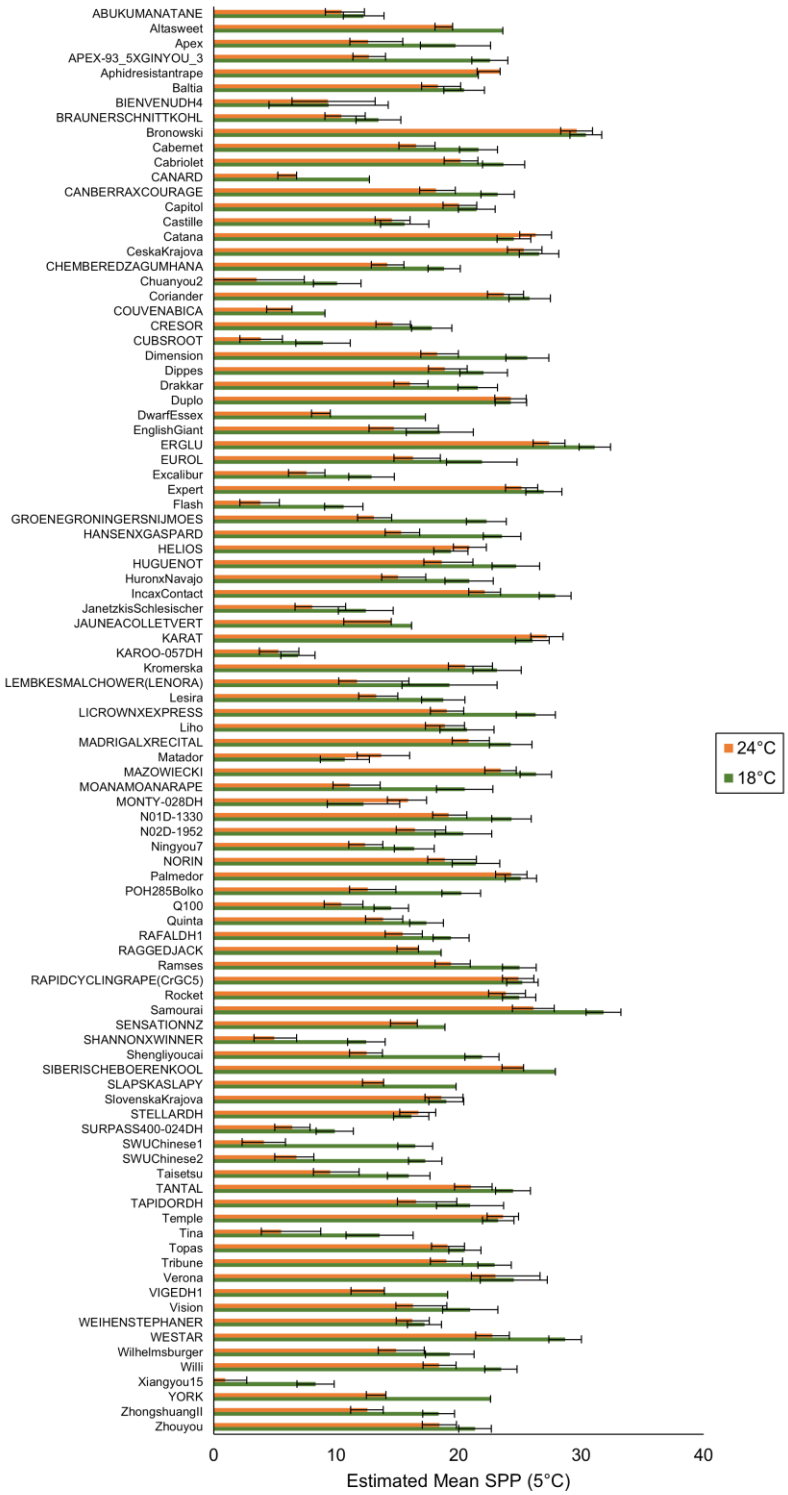

Figure 2.23: Estimated mean SPP from linear model for each variety at 5°C vernalisation temperature, comparing maturation treatments at 18°C and 24°C. Error bars show standard error.

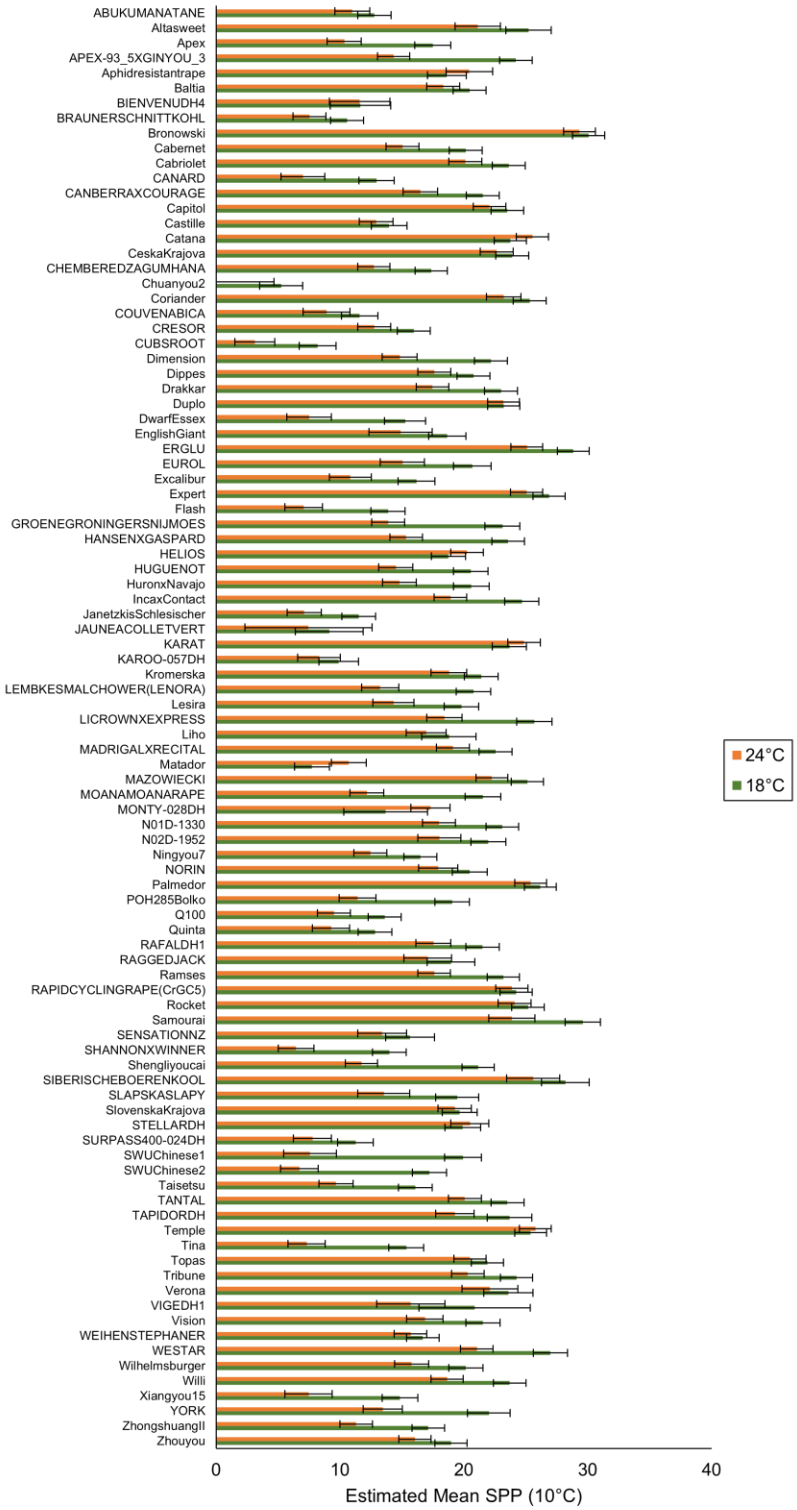

Figure 2.24: Estimated mean SPP from linear model for each variety at 10°C vernalisation temperature, comparing maturation treatments at 18°C and 24°C. Error bars show standard error.

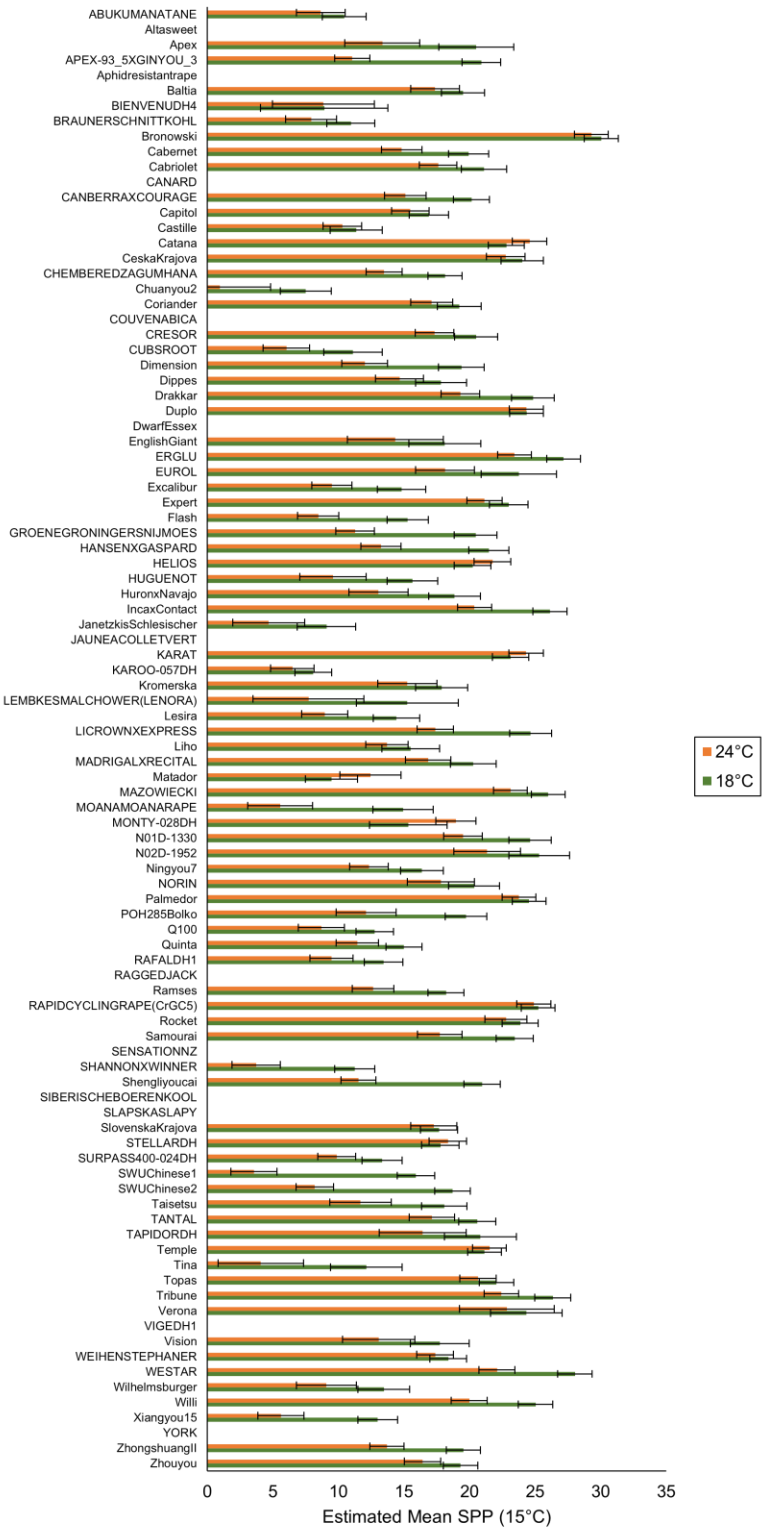

Figure 2.25: Estimated mean SPP from linear model for each variety at 15°C vernalisation temperature, comparing maturation treatments at 18°C and 24°C. Error bars show standard error.

2.3.6 Impact of temperature treatment on seed yield

Although total plant yield was not directly measured in this experiment, TGW and SPP can be used as indicators of the overall yield of the plant by computing mean seed weight per pod:

$$Yield = \frac{TGW}{1000} \times SPP$$

This new estimation of yield was then used with mixed linear model to produce an estimation of yield for each temperature treatment.

2.3.6.1 Optimizing a linear model to estimate yield

When estimating TGW, there was found to be a block effect which impacted TGW. As TGW was part of the input used to estimate yield, it was important to first check if block was also affecting the yield estimate. The effect of block on yield was explored using ANOVA to estimate yield (Y) in block (B) i, with the residual term e_i:

$$Y_i = \mu + B_i + e_i$$

This determined that there was an effect of block on yield which needed to be accounted for (Table 2.12).

Table 2.12: ANOVA results indicating effect of block on yield.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Block	11	0.3241	0.02946	36.36	<2e-16
Residuals	2581	2.0912	0.00081		

The twelve treatments used in the phenotyping experiment consisted of variations in vernalisation temperature, vernalisation duration and maturation temperature. These three components were included in the linear model to assess their effects on yield:

$$Y_{ijklm} = \mu + MT_j + VT_k + VD_l + B_m + v_i + e_{ijklm}$$

This model estimates yield Y of variety (v) i in maturation temperature (MT) j, vernalisation temperature (VT) k, vernalisation duration (VD) l and block (B) m with the residual term e_{ijklm}. This is a mixed linear model with MT, VT and VD as fixed effects and v and B as random effects.

Table 2.13: Linear model results of treatment effects on yield. DenDF = denominator degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Maturation temperature	0.043004	0.043004	1	9.01	107.3371	2.64E-06
Vernalisation Temperature	0.051152	0.025576	2	2490.86	63.8378	< 2.2e-16
Vernalisation duration	0.000323	0.000323	1	9.01	0.8062	0.3926

As previously found with TGW and SPP, the duration of the vernalisation treatment was not found to have a significant effect on yield across the diversity set used in the experiment (Table 2.13).

Interactions in this model between the fixed effects terms MT, VT and VD were also explored. While the duration of vernalisation was not found to have an effect on yield

individually, it was found to have a significant interaction with vernalisation temperature (Table 2.14).

Table 2.14: Interactions of fixed terms in linear model estimating yield. DenDF = denominator degrees of freedom. Results displayed using Type III ANOVA with Satterthwaite's method.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Maturation Temperature (MT)	0.045916	0.045916	1	8.15	115.2471	4.34E-06
Vernalisation Temperature (VT)	0.046012	0.023006	2	2484.28	57.744	< 2.2e-16
Vernalisation Duration (VD)	0.00014	0.00014	1	8.13	0.3519	0.569204
MT : VT	0.000861	0.000431	2	2479.84	1.0811	0.339382
MT : VD	0.00069	0.00069	1	8.1	1.7314	0.224251
VT : VD	0.006931	0.003465	2	2479.94	8.6983	0.000172
MT : VT: VD	0.000065	0.000032	2	2480.07	0.0815	0.921725

2.3.6.2 Differences in estimated yield across treatments

To assess how estimated yield changes across the 12 treatments, the linear model was used:

$$Y_{ijklm} = \mu + MT_j + VT_k \times VD_l + B_m + v_i + e_{ijklm}$$

This model estimates yield (Y) of variety (v) i in maturation temperature (MT) j, vernalisation temperature (VT) k, vernalisation duration (VD) l and block (B) m with the residual term e_{ijklm}. This is a mixed linear model with MT, VT and VD as fixed effects, v

and B as random effects and an interaction term between VT and VD. Using the Nakagawa and Schielzeth method (2013) of calculating R-squared statistics for mixed linear models suggested that this model explains 61% of variation in TGW.

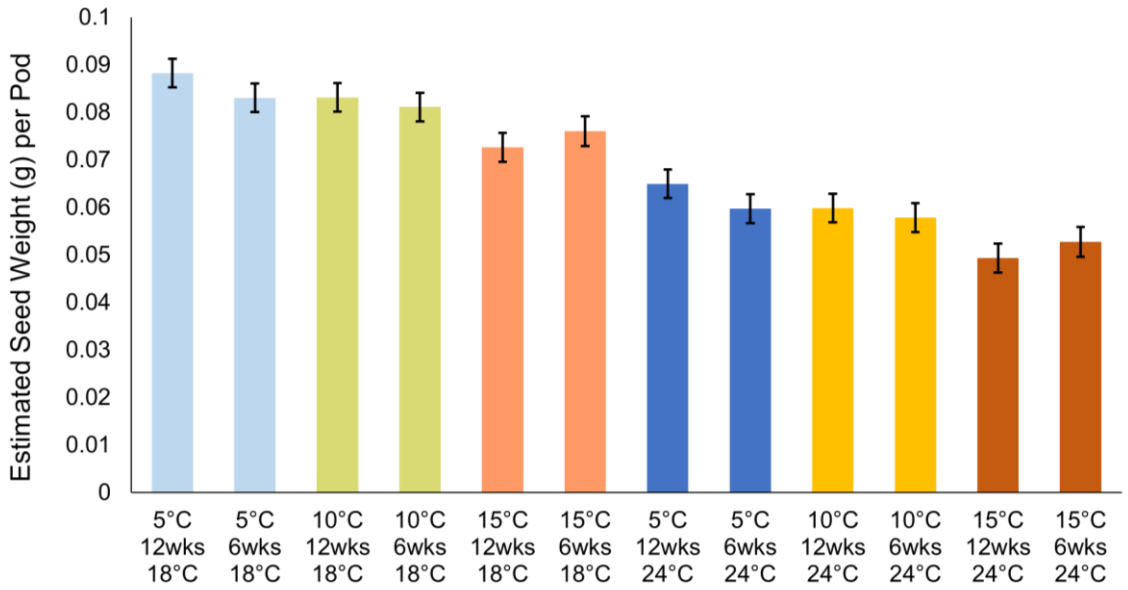

Figure 2.26: Estimated yield as seed weight per pod for each overall temperature treatment. Error bars show SE.

Table 2.15: Details of linear model estimated mean for yield, calculated as seed weight per pod. All values shown to 3 significant figures. SE = standard error. df = degrees of freedom. CL = 95% confidence limit.

Treatment	Mean Yield Estimate	SE	df	Lower CL	Upper CL
5°C 12wks 18°C	0.0882	0.00299	44	0.0822	0.0943
5°C 6wks 18°C	0.083	0.003	44.2	0.0769	0.089
10°C 12wks 18°C	0.0831	0.00299	44	0.0771	0.0892
10°C 6wks 18°C	0.0811	0.00301	44.8	0.075	0.0872
15°C 12wks 18°C	0.0726	0.00304	47.1	0.0665	0.0788
15°C 6wks 18°C	0.076	0.00313	52.8	0.0697	0.0823
5°C 12wks 24°C	0.0649	0.003	44.3	0.0589	0.071
5°C 6wks 24°C	0.0597	0.00301	44.7	0.0536	0.0657

10°C 12wks 24°C	0.0598	0.003	44.3	0.0538	0.0659
10°C 6wks 24°C	0.0578	0.00302	45.3	0.0517	0.0639
15°C 12wks 24°C	0.0493	0.00306	48.1	0.0432	0.0555
15°C 6wks 24°C	0.0527	0.00313	52.7	0.0464	0.0589

From the estimated means produced by the model, it can be seen that higher temperatures lead to an overall reduction in estimated yield (Figure 2.26). The highest yielding temperature treatment was an 18°C seed maturation temperature combined with a 5°C vernalisation treatment for twelve weeks; this was 0.0389g larger than the mean yield estimate for the lowest yielding treatment, 24°C seed maturation temperature combined with a 15°C vernalisation treatment for twelve weeks (Table 2.15).

2.3.7 Effect of crop type on temperature responses

Within the 96-accession diversity set studied in this phenotyping experiment, the accessions have been grouped into a variety of crop types including Winter OSR, Spring OSR, Semi-winter OSR, swedes and an exotic grouping which contains varieties such as Siberian kales. To fully understand the interactions between crop type and temperature treatments, linear modelling was used to compare TGW and SPP between different treatments.

2.3.7.1 Optimizing model for estimating TGW by crop type

Linear modelling was carried out using the lmerTest package in R. The formula for the general linear model used to estimate TGW was:

$$TGW_{ijklm} = \mu + MT_j \times VT_k \times c_i + B_m + e_{ijklm}$$

This model predicts the TGW of crop type (c) i in maturation temperature (MT) j in vernalisation temperature (VT) k and block (B) m, where e_{ijklm} is the residual term. Block was included due to the previously identified effect of block on TGW. Interactions between maturation temperature, vernalisation temperature and crop type were included in the model to assess which of these interactions were significant.

Table 2.16: Linear model output for estimating TGW by crop type. DenDF = denominator degrees of freedom. Results displayed using Type III Analysis of Variance Table with Satterthwaite's method.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Crop Type	515.66	85.944	6	2544.66	98.1751	< 2.2e-16
Maturation Temperature (MT)	5.42	5.423	1	45.12	6.1947	0.01657
Vernalisation Temperature (VT)	1.49	0.746	2	2544.84	0.8522	0.42658
Crop Type x MT	36.85	6.142	6	2544.69	7.0166	2.03E-07
Crop Type x VT	18.56	1.547	12	2543.18	1.7671	0.0481
MT x VT	1.09	0.543	2	2545.24	0.6204	0.53783
Crop Type x MT x VT	9.05	0.823	11	2543.27	0.9396	0.50069

This model (Table 2.16) confirms that there is an interaction between crop type and maturation temperature, as well as between crop type and vernalisation temperature. These terms were included as interaction terms for further assessment of TGW. This model also suggests that vernalisation temperature did not individually have an effect on TGW, but rather that its effect was due to interactions with crop type. Vernalisation temperature was included in further models in order to assess how crop type interacts with vernalisation temperature in regard to estimating TGW.

The optimized model for estimating TGW by crop type was:

$$TGW_{ijklm} = \mu + MT_j + VT_k + c_i + (MT_j \times c_i) + (VT_k \times c_i) + B_m + e_{ijklm}$$

This model predicts the TGW of crop type (c) i in maturation temperature (MT) j in vernalisation temperature (VT) k and block (B) m, where e_{ijklm} is the residual term. In this model, *MT, VT,* and c are fixed effects while B is a random effect. Block was included due to the previously identified effect of block on TGW. Interactions between maturation temperature and crop type as well as between vernalisation temperature and crop type were included (Table 2.17).

Table 2.17: Linear model results of optimized model for estimating TGW by crop type. Results displayed using Type III Analysis of Variance Table with Satterthwaite's method. DenDF = denominator degrees of freedom. DF = degrees of freedom.

	Sum Sq	**Mean Sq**	**DF**	**DenDF**	**F value**	**Pr(>F)**
Crop Type	568.92	94.821	6	2557.68	108.3638	< 2.2e-16
Maturation Temperature (MT)	4.37	4.368	1	34.59	4.9922	0.03203
Vernalisation Temperature (VT)	1.13	0.566	2	2557.32	0.6469	0.52377
Crop Type x MT	43.12	7.187	6	2557.3	8.214	7.98E-09
Crop Type x VT	17.96	1.497	12	2556.12	1.7108	0.0584

2.3.7.2 Optimizing model used for estimated SPP by crop type

The formula for the general linear model used to estimate SPP was:

$$SPP_{ijk} = \mu + MT_j \times VT_k \times c_i + e_{ijk}$$

This model predicts the SPP of crop type (c) i in maturation temperature (MT) j in vernalisation temperature (VT) k where $e_{ijk\ is}$ the residual term. Interactions between maturation temperature, vernalisation temperature and crop type were included in the model to assess which of these interactions were significant.

This model (shown in Table 2.18) identifies significant interactions between maturation temperature and crop type and vernalisation temperature and crop type. These interactions terms were included in further models for estimating SPP.

Table 2.18: Model output for estimating SPP and crop type. Results displayed using Type III Analysis of Variance Table with Satterthwaite's method. DF = degrees of freedom.

	DF	**Sum Sq**	**Mean Sq**	**F value**	**Pr(>F)**
Crop Type	6	10451	1741.9	45.0449	< 2e-16
Maturation Temperature (MT)	1	7795	7795.2	201.5871	< 2e-16
Vernalisation Temperature (VT)	2	271	135.5	3.5037	0.03023
Crop Type x MT	6	630	105	2.7147	0.01248
Crop Type x VT	12	859	71.6	1.8519	0.03565
MT x VT	2	4	2.1	0.0542	0.94725
Crop Type x MT x VT	11	233	21.2	0.5476	0.87155
Residuals	2555	98800	38.7		

The optimized model for estimating SPP by crop type was:

$$SPP_{ijk} = \mu + MT_j + VT_k + c_i + (MT_j \times c_i) + (VT_j \times c_i) + e_{ijk}$$

This model predicts the SPP of crop type (c) i in maturation temperature (MT) j in vernalisation temperature (VT) k where $e_{ijk\ is}$ the residual term. The model also contains interaction terms for maturation temperature and crop type as well as vernalisation temperature and crop type.

Table 2.19: Optimized general linear model output for estimating SPP by crop type. Results displayed using Type III Analysis of Variance Table with Satterthwaite's method. DF = degrees of freedom.

	DF	Sum Sq	Mean Sq	F value	Pr(>F)
Crop Type	6	10451	1741.9	45.1657	< 2e-16
Maturation Temperature (MT)	1	7795	7795.2	202.1277	< 2e-16
Vernalisation Temperature (VT)	2	271	135.5	3.5131	0.02995
Crop Type x MT	6	630	105	2.722	0.01227
Crop Type x VT	12	859	71.6	1.8569	0.03502
Residuals	2568	99037	38.6		

2.3.7.3 Effects of vernalisation on seed traits in different crop types

Various crop types respond differently to vernalisation temperature (Figure 2.27). As was the case across the diversity set as a whole, Winter OSR accessions are clearly affected by increasing vernalisation temperature (Figure 2.27a). In Winter OSR, increasing vernalisation temperature causes a reduction in TGW and SPP. Increasing vernalisation temperature from 5°C to 15°C caused a drop in TGW of 0.38g and a reduction in SPP of 1.9 (Table 2.20).

In spring OSR, the impact of vernalisation on yield parameters was only seen at 15°C vernalisation temperature. There was no significant difference between vernalisation temperatures of 5°C and 10°C; increasing vernalisation temperature from 10°C to 15°C caused a decrease in TGW of 0.23g, with no significant effect on the number seeds per pod.

Semi-winter OSR accessions had no significant response to vernalisation temperature, with no significant change in TGW or SPP between vernalisation temperatures. Swede varieties also displayed no significant effect on either TGW or SPP. Although the estimated mean SPP for swedes at 15°C was 5.8 lower than at 10°C, this result was not found here to be statistically significant ($p = 0.07$), likely due to a low number of swedes producing seeds at this vernalisation temperature (Figure 2.27b).

These results show that Semi-winter OSR varieties produce a lesser number of larger seeds, the size and number of which are unaffected by vernalisation temperature. Spring OSR varieties produce a larger number of smaller seeds, with seed size being slightly affected by vernalisation temperature. In Winter OSR, seed size and seed number are both negatively affected by increasing vernalisation temperature, with a 15°C vernalisation treatment associated with both smaller and fewer seeds produced.

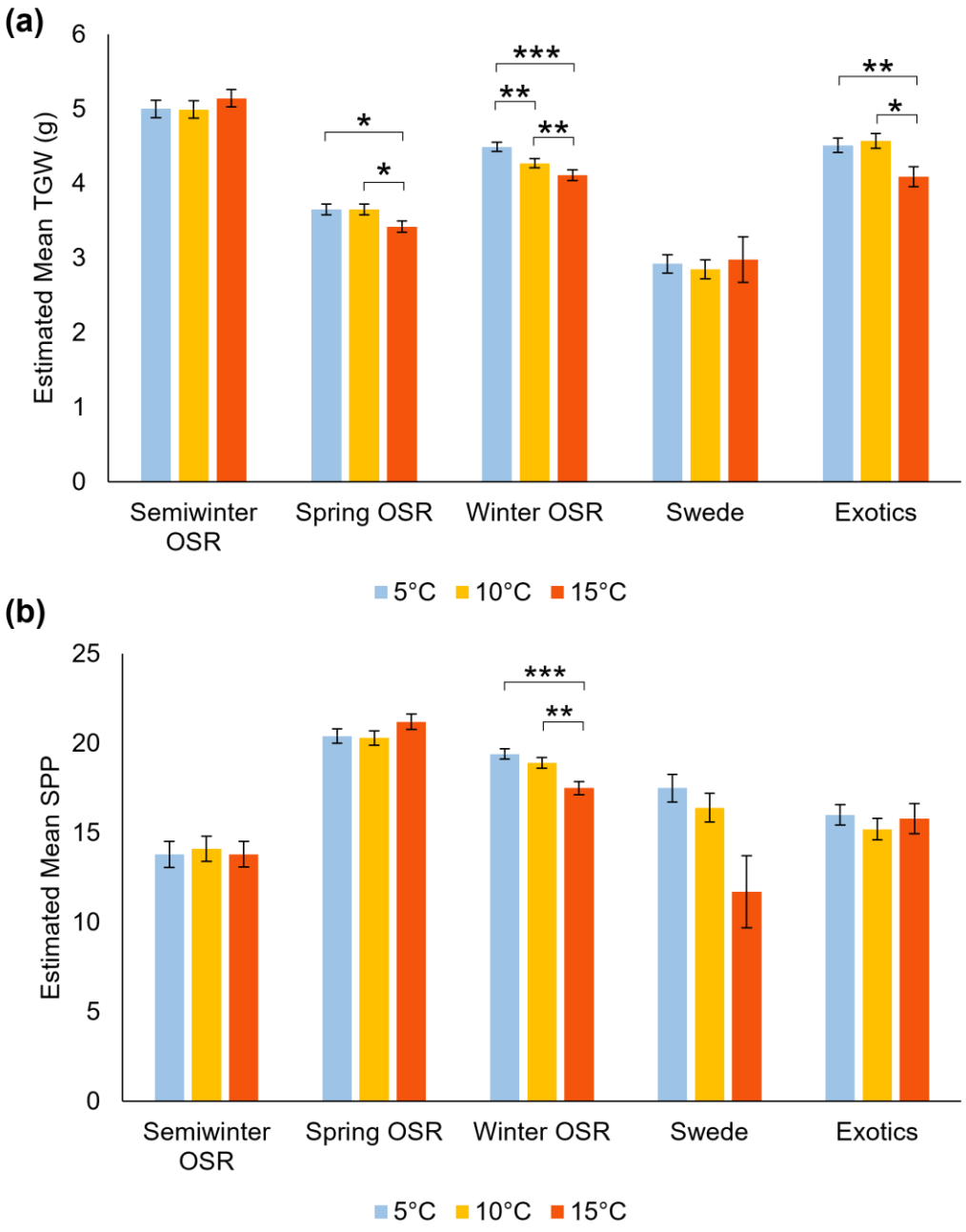

*Figure 2.27: Model-predicted mean values for (a) SPP and (b) TGW by crop type and vernalisation temperature. Error bars show SE and asterisks indicate p values where: p < 0.05 = *; p < 0.01 = **; p < 0.001 = ***.*

Table 2.20: Details of linear model output for TGW and SPP by crop type and vernalisation temperature, as generated by lmerTest package in R. Data for TGW is measured in g. Data for SPP is count data. SE = standard error. df = degrees of freedom. CL = confidence limit.

Crop Type	Vernalisation Temperature	Estimated Mean	SE	df	Lower CL	Upper CL
SPP						
Semi-winter	5°C	13.8	0.729	2568	12.42	15.3
Semi-winter	10°C	14.1	0.708	2568	12.67	15.4
Semi-winter	15°C	13.8	0.718	2568	12.39	15.2
Spring OSR	5°C	20.4	0.401	2568	19.61	21.2
Spring OSR	10°C	20.3	0.403	2568	19.54	21.1
Spring OSR	15°C	21.2	0.433	2568	20.3	22
Winter OSR	5°C	19.4	0.294	2568	18.8	19.9
Winter OSR	10°C	18.9	0.298	2568	18.34	19.5
Winter OSR	15°C	17.5	0.376	2568	16.78	18.3
Swede	5°C	17.5	0.777	2568	15.98	19
Swede	10°C	16.4	0.802	2568	14.86	18
Swede	15°C	11.7	2.013	2568	7.74	15.6
Exotics	5°C	16	0.577	2568	14.88	17.1
Exotics	10°C	15.2	0.598	2568	13.98	16.3
Exotics	15°C	15.8	0.841	2568	14.13	17.4
TGW						
Semi-winter	5°C	5	0.1173	408.3	4.77	5.23
Semi-winter	10°C	4.99	0.1143	374.2	4.76	5.21
Semi-winter	15°C	5.14	0.1157	390.1	4.91	5.36
Spring OSR	5°C	3.65	0.073	68.8	3.5	3.8
Spring OSR	10°C	3.65	0.0732	69.5	3.5	3.79
Spring OSR	15°C	3.42	0.077	84.9	3.27	3.58
Winter OSR	5°C	4.49	0.0604	32.6	4.37	4.62
Winter OSR	10°C	4.27	0.0607	33.2	4.15	4.4
Winter OSR	15°C	4.11	0.07	58.1	3.97	4.25
Swede	5°C	2.92	0.1241	494.3	2.67	3.16
Swede	10°C	2.85	0.1277	541.2	2.6	3.1
Swede	15°C	2.98	0.3065	2360.2	2.38	3.58
Exotics	5°C	4.51	0.0961	198.8	4.32	4.7
Exotics	10°C	4.57	0.099	222	4.38	4.77
Exotics	15°C	4.09	0.1333	619.7	3.83	4.35

Ratios of TGW and SPP between vernalisation temperatures of 5°C and 15°C were generated from the raw data from the phenotyping experiment (Figure 2.28). A TGW ratio of greater than 1 indicates that TGW was increased at 5°C compared to 15°C, while a SPP ratio of greater than 1 indicated that SPP was increased at 5°C compared to 15°C.

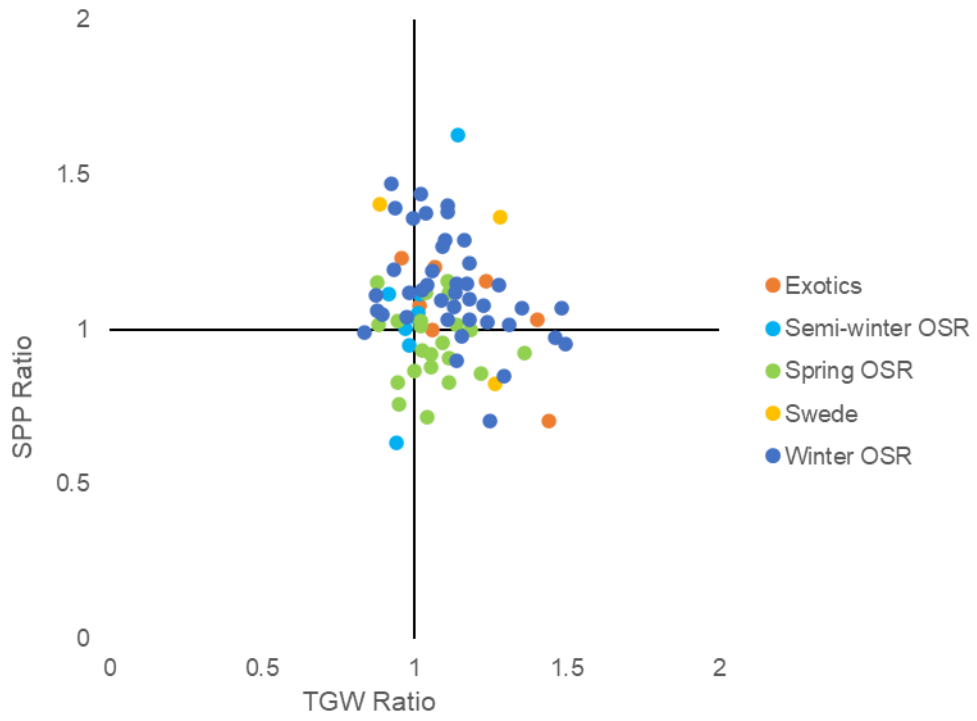

Figure 2.28: Scatter graph graph showing ratios of TGW and SPP by vernalisation temperature (5°C/15°C) using raw data from phenotyping. Crop types with less than three repeats were not included in the graph and varieties missing data for at least one temperature were excluded as ratio could not be calculated. Exotics (n = 7), Semi-winter OSR (n = 8), Spring OSR (n = 22), Swede (n = 3), Winter OSR (n = 41).

The graph indicates that TGW was increased at 5°C maturation temperatures for most lines while SPP was less strongly affected by changes in vernalisation temperature. Most Winter OSR varieties cluster in the top right quadrant of the graph, indicating that TGW and SPP is improved in these varieties at lower vernalisation temperatures. Spring and Semi-winter OSR varieties cluster closer to the centre-right of the graph, indicating a reduced effect of vernalisation temperature particularly with regards to SPP. From this graph it can be seen that the raw data generally corresponds to the predictions of the model regarding vernalisation temperature effects on different crop types.

2.3.7.4 Effects of maturation temperature on seed traits in different crop types

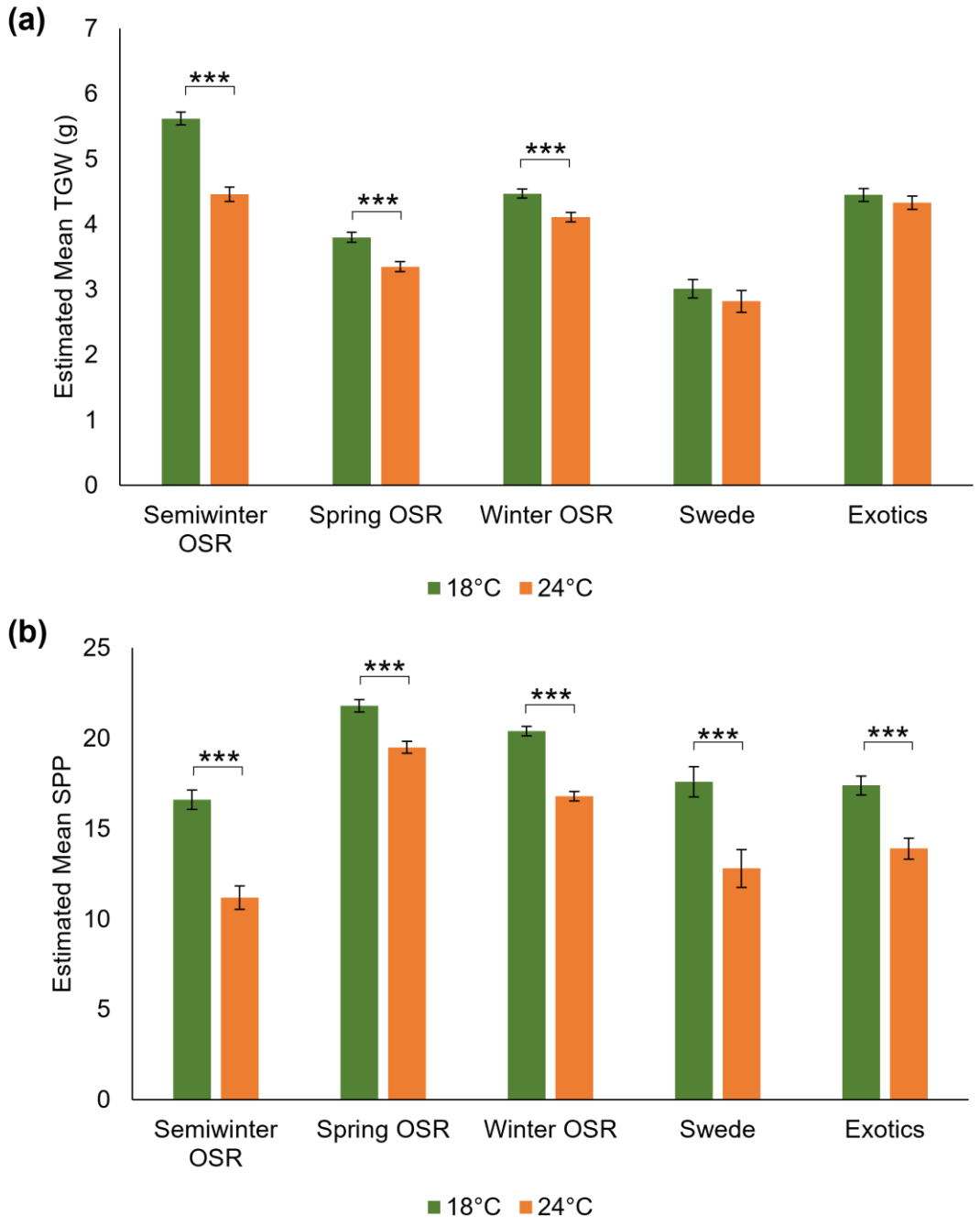

*Figure 2.29: Model-predicted mean values for (a) TGW and (b) SPP by crop type and maturation temperature. Error bars show SE and asterisks indicate p values where: p < 0.05 = *; p < 0.01 = **; p < 0.001 = ***.*

Table 2.21: Details of estimated means from linear model for TGW and SPP by crop type and maturation temperature. Data for TGW is measured in g. Data for SPP is count data. SE = standard error. df = degrees of freedom. CL = confidence limit.

Crop Type	Maturation Temperature	Estimated Mean	SE	df	Lower CL	Upper CL
SPP						
Semi-winter	18°C	16.6	0.537	2568	15.54	17.6
Semi-winter	24°C	11.2	0.645	2568	9.95	12.5
Spring OSR	18°C	21.8	0.339	2568	21.09	22.4
Spring OSR	24°C	19.5	0.333	2568	18.84	20.1
Winter OSR	18°C	20.4	0.254	2568	19.94	20.9
Winter OSR	24°C	16.8	0.27	2568	16.25	17.3
Swede	18°C	17.6	0.849	2568	15.96	19.3
Swede	24°C	12.8	1.034	2568	10.77	14.8
Exotics	18°C	17.4	0.517	2568	16.34	18.4
Exotics	24°C	13.9	0.569	2568	12.83	15.1
TGW						
Semi-winter	18°C	5.62	0.0996	59.6	5.42	5.82
Semi-winter	24°C	4.46	0.1131	98.4	4.24	4.69
Spring OSR	18°C	3.8	0.0773	21.9	3.64	3.96
Spring OSR	24°C	3.35	0.0767	21.1	3.19	3.5
Winter OSR	18°C	4.47	0.0695	14.3	4.32	4.62
Winter OSR	24°C	4.11	0.0708	15.4	3.96	4.27
Swede	18°C	3.01	0.1407	224.7	2.74	3.29
Swede	24°C	2.82	0.1664	411.2	2.49	3.15
Exotics	18°C	4.45	0.0971	54	4.25	4.64
Exotics	24°C	4.33	0.1036	69.7	4.12	4.54

Across different crop types, TGW was affected differently by maturation temperature (Figure 2.29a). An increased maturation temperature caused a significant decrease in

TGW in Semi-winter OSR, Spring OSR and Winter OSR. In Semi-winter OSR, TGW fell by 1.16g at 24°C compared to 18°C (Table 2.21). In Spring OSR, TGW fell by 0.5 while Winter OSR showed an estimated mean decrease of 0.36g.

When considering SPP, all crop types in this experiment were affected by maturation temperature, with increased maturation temperature being associated with significant reductions in SPP (Figure 2.29b). This effect was largest in Semi-winter OSR, with an estimated mean decrease of 5.39 seeds per pod, while the effect was smallest in Spring OSR which showed an estimated mean decrease of 2.26 seeds per pod at 24°C (Table 2.21).

These results show that Semi-winter OSR was the crop type most affected by increased maturation temperature, with the largest decreases in TGW and SPP. Varieties in the Exotics grouping were least affected by the increase in maturation temperature.

Ratios of TGW and SPP between maturation temperatures of 18°C and 24°C were generated from the raw data from the phenotyping experiment (Figure 2.30). A TGW ratio of greater than 1 indicates that TGW was increased at 18°C compared to 24°C, while a SPP ratio of greater than 1 indicated that SPP was increased at 18°C compared to 24°C.

The majority of varieties lie within the top right quadrant of the graph, indicating that 18°C temperatures were particularly favourable for both TGW and SPP across the varieties studied. A strong effect of maturation temperature can be seen in the large ratios of Semi-winter OSR in particular.

From Figure 2.30 it can be seen that the raw data generally corresponds to the predictions made by the model on the responses of different crop types under different maturation temperatures.

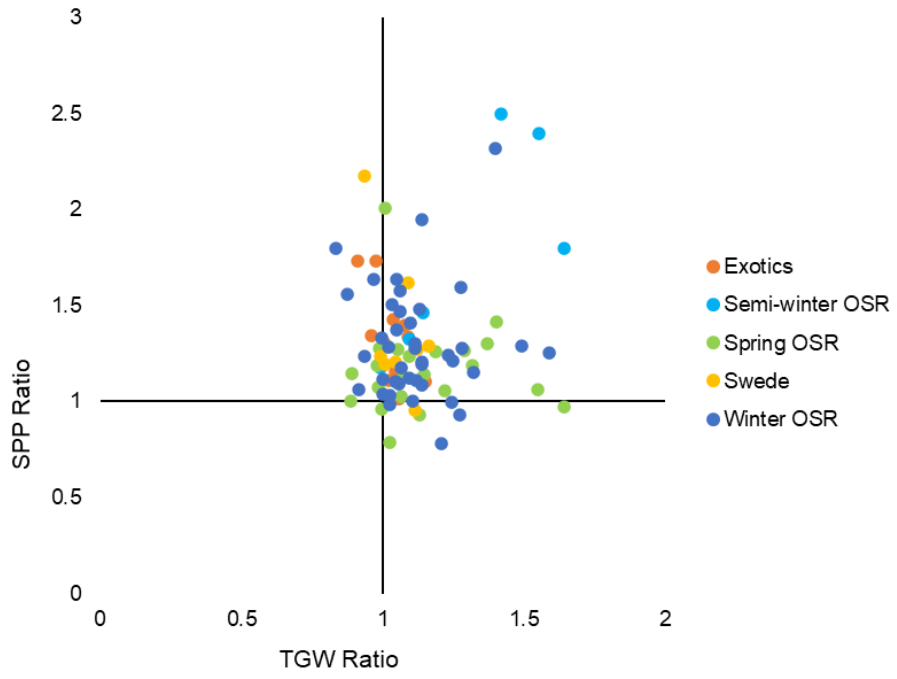

Figure 2.30: Scatter graph showing ratios of TGW and SPP by maturation temperature (5°C/15°C) using raw data from phenotyping. Crop types with less than three repeats were not included in the graph and two outlier varieties removed for readability (Chuanyou2: TGW Ratio = 8.0, SPP Ratio = 1.1; and SWU Chinese 1: TGW Ratio 4.0, SPP Ratio 1.3). Exotics (n = 11), Semi-winter OSR (n = 6), Spring OSR (n = 22), Swede (n = 8), Winter OSR (n = 41).

2.3.8 Reproducibility of results from the glasshouse for CER experiments

The results of the large phenotyping experiment provided a wealth of information on how *Brassica napus* responds to temperature. However, an experiment of this scale could not be easily repeated. A smaller scale experiment was planned to test if the TGW results from the larger experiment, carried out in a glasshouse, could be robustly reproduced in a controlled environment room (CER). Replicating the previously seen results of the effect of temperature in a different growth environment would be an indication that the effect seen in the glasshouse phenotyping experiment is robust.

To this end, four lines with contrasting temperature responses were selected from the results of the large phenotyping experiment: Palmedor, Castille, Catana, and Expert. Selecting four lines with different temperature responses meant the experimental design should confirm if all responses could be accurately repeated. Comparing TGW between the 5°C and 15°C vernalisation temperatures used in the original phenotyping experiment showed that in Catana and Palmedor, TGW was significantly lower in plants vernalised at 15°C with no vernalisation-associated change in TGW seen in Castille or Expert (Figure 2.31a). Comparing TGW between the 18°C and 24°C vernalisation temperatures used in the original phenotyping experiment showed that in Castille and Catana, TGW was significantly lower in plants matured at 24°C with no vernalisation-associated change in TGW seen in Palmedor or Expert (Figure 2.31b).

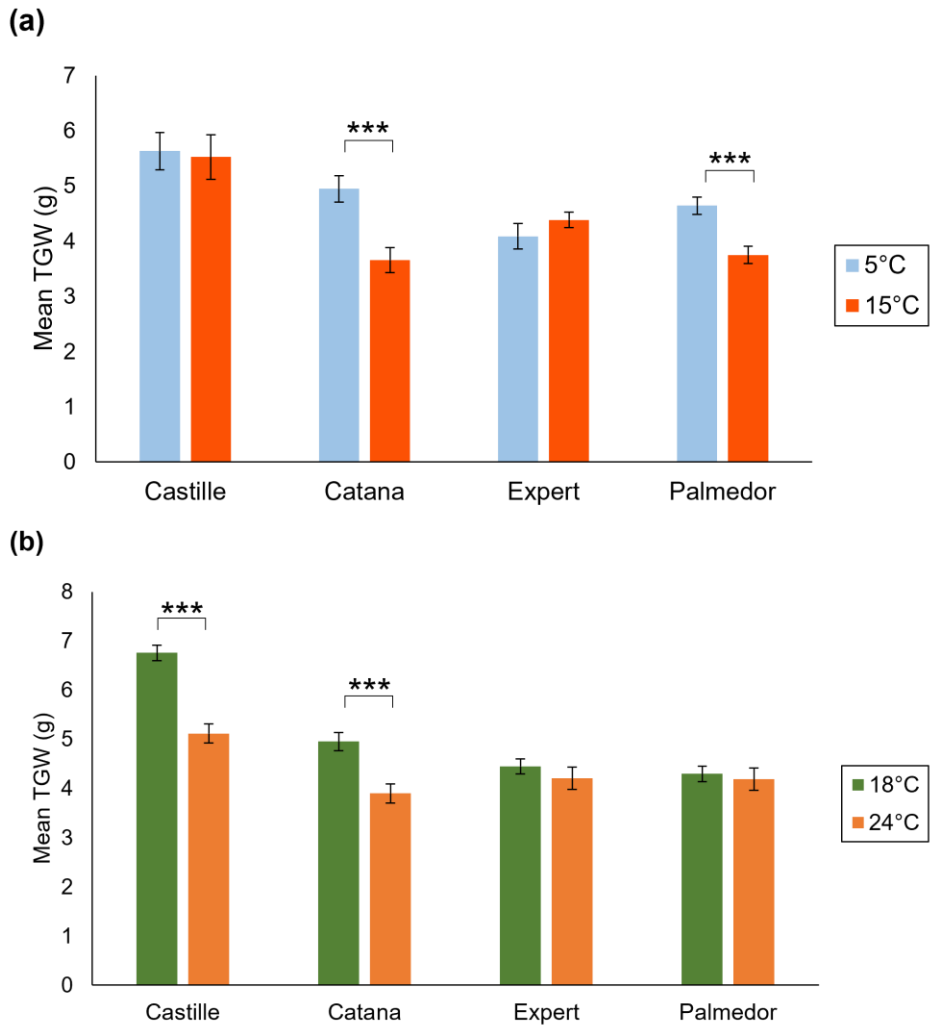

*Figure 2.31: TGW for four selected lines compared between (a) vernalisation temperature; (b) maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p < 0.01 is ** and p< 0.001 is ***, calculated using Student's t-test. Unmarked pairs of bars showed no significant difference. Error bars show SE.*

These four lines were then grown under a smaller version of the original phenotyping experiment (see Figure 2.32 for experimental design). As I found no impact of vernalisation duration across the diversity set, all lines were vernalised for 6 weeks. Vernalisation was carried out either at 5°C or 15°C, as comparing these temperature treatments gave the greatest contrast in responses.

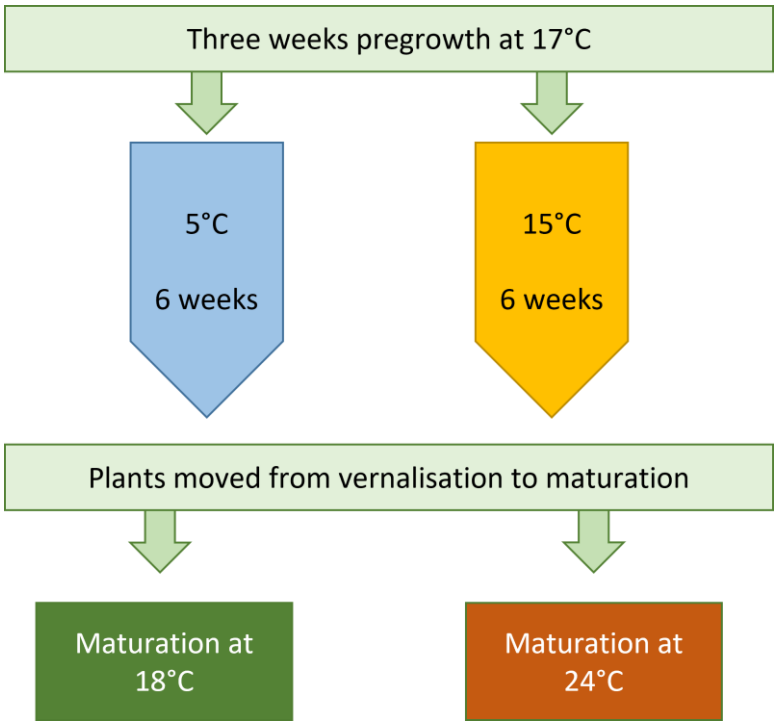

Figure 2.32: Experimental design of the repeat phenotyping experiment. Plants were given one of 4 temperature treatments.

When this experiment was completed and seeds were harvested, it was apparent that pre-harvest sprouting was extremely pronounced in plants grown at 24°C; some plants which were vernalised at 15°C also failed to flower. This resulted in a more limited scope for the analysis of this experiment, as not all accessions included in the experiment underwent sufficient seed development for a reliable analysis.

When considering the complete set of results, there were some notable differences when compared the original data seen in Figure 2.29. While a response to maturation temperature was previously only seen in Catana and Castille, all four lines now show a response (Figure 2.33). Additionally, responses to vernalisation were originally seen only in Catana and Palmedor; in this repeated experiment, Catana and Expert now show a vernalisation response while Palmedor does not. However, it is important to note that data is incomplete for several lines and the results from plants grown at 24°C may be unreliable due to pre-harvest sprouting.

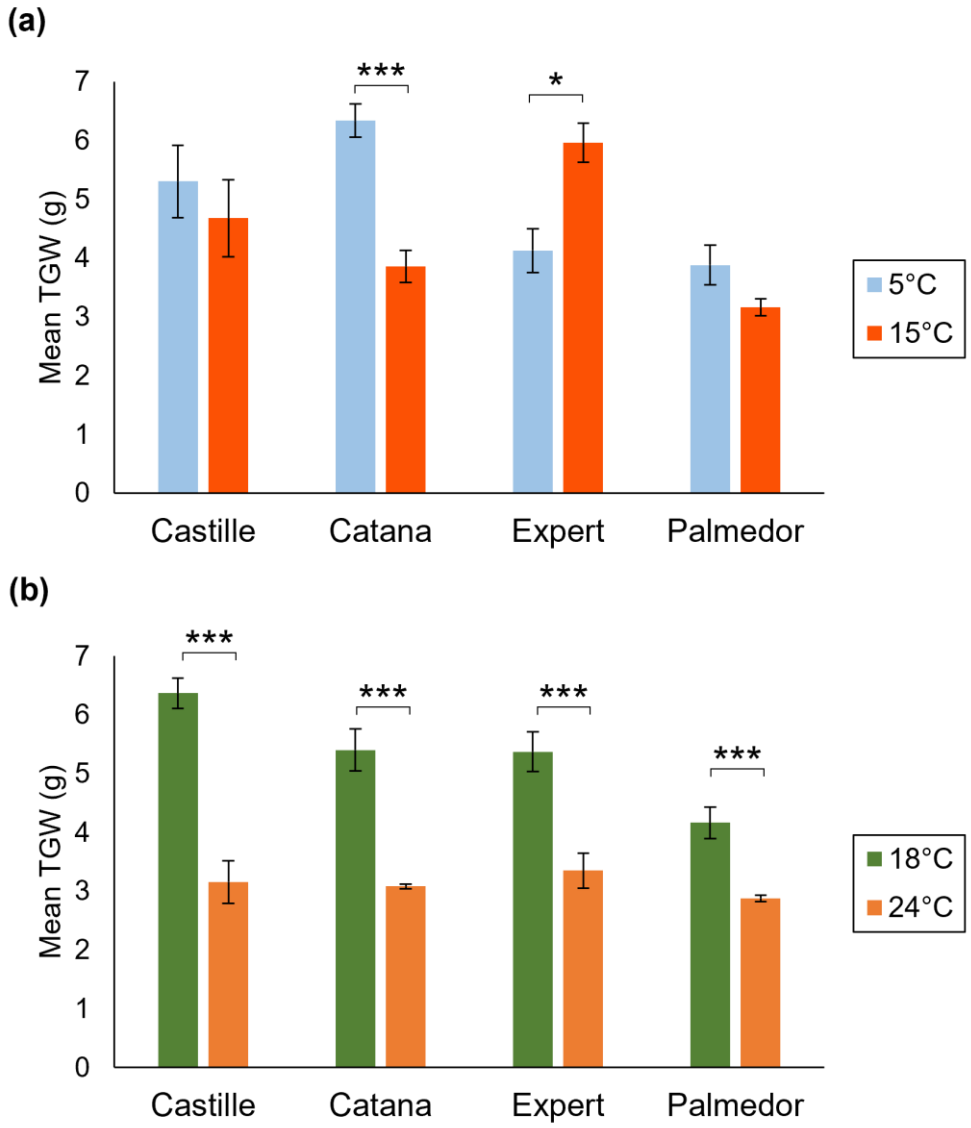

*Figure 2.33: TGW for four selected lines compared between (a) vernalisation temperature; (b) maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p < 0.01 is ** and p< 0.001 is ***, calculated using Student's t-test. Unmarked pairs of bars showed no significant difference. Error bars show SE.*

For the results to be more reliable, TGW was compared between the two vernalisation temperatures but only for a maturation temperature of 18°C (Figure 2.34). As in the original experiment, Expert shows no significant response to vernalisation temperature. Palmedor and Catana retain a significant effect of vernalisation temperature on TGW, and TGW in Castille was also affected by vernalisation. From these results I concluded

that the results from the glasshouse phenotyping experiment were robust and repeatable for some lines. The effect of vernalisation was similar in Palmedor, Catana and Expert across both experiments for a maturation temperature of 18°C.

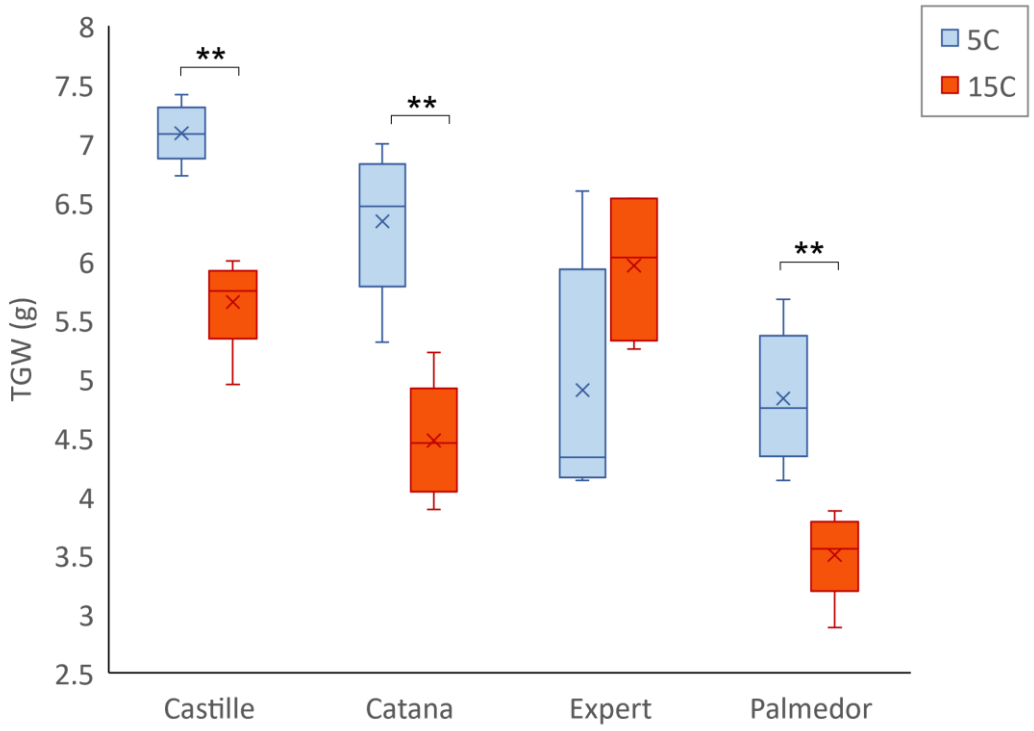

*Figure 2.34: Boxplot displaying the results of TGW from the repeated phenotyping experiment by vernalisation temperature, only for plants matured at 18°C. Asterisks used to indicate significance, where p ≤ 0.05 is *, p < 0.01 is ** and p< 0.001 is ***, calculated using Student's t-test.*

2.4 Conclusions & Discussion

The experiments described in this chapter aimed to clarify the effects of temperature on yield components in *Brassica napus*, using thousand grain weight (TGW) and number of seeds per pod (SPP) as a representation of total yield.

Results from the large phenotyping experiment showed that vernalisation temperature and seed maturation temperature had a significant effect on seed yield parameters across the *Brassica napus* diversity set of 96 accessions studied. Higher vernalisation

temperatures were associated with reduced TGW and SPP, particular at temperatures above 10°C. Two seed maturation temperatures were used in this experiment, with the warmer 24°C treatment being associated with lower TGW, and fewer SPP. Vernalisation length was not found to have an impact on TGW or SPP across the diversity set as a whole. Within the diversity set, different crop types were found to respond to temperature in different ways. Winter OSR varieties were affected strongly by vernalisation temperature and seed maturation temperature; however, spring and semi-winter OSR varieties displayed a reduced or absent effect of vernalisation temperature.

No effect of vernalisation length could be seen across the dataset, suggesting that effects of vernalisation treatment on overall seed yield are instead due to temperature. Within the diversity set, the recommended vernalisation length varies between accessions – for example, many will not require 6 weeks while some will need 12 weeks. Despite this, vernalisation length was not found to have a significant effect on TGW or SPP in any of the crop types. This may suggest that even if the vernalisation is not favourable, six weeks may be sufficient for complete reproductive development, at least in the accessions which were included in the analysis.

Plants which failed to flower or produced very few seed pods were not included in the analysis – hence it is possible that the effect of too little vernalisation had an impact on yield by impacting the ability of the plant to reproduce successfully. However, as this work focused on the effect of temperature on development, plants which failed to go through sufficient development were not included in order to ensure reliable results. It is possible that future studies could quantify the impact of vernalisation duration on overall quality of growth, including analyses of plants which produced little to no seed.

Across the entire dataset studied in this experiment, vernalisation temperatures above 10°C were particularly associated with reduced yield. However, this did not hold true for all crop types, as Semi-winter OSR accessions in particular were unaffected by

vernalisation temperature. The fact that vernalisation temperature has an effect on seed yield parameters, and that this differs between crop types which are known to have variation in *FLC* dynamics (Schiessl et al, 2019), suggests that it is possible the mechanism by which vernalisation temperature and yield are linked is through *FLC* gene expression and its downstream target genes.

In the field, it has been described that vernalisation in winter oilseed rape tends to occur through October when the temperature is approximately 10°C, with the floral meristem then exposed to lower temperatures (O'Neill et al, 2019). UKCP18 climate projections have suggested that the effects of climate change are likely to cause milder temperatures during autumn and winter in the UK. Warmer temperatures during a critical period of vernalisation for oilseed rape could negatively affect yield unless *Brassica napus* crops are bred to adapt to changing temperatures.

Seed maturation temperature was also found to affect TGW and SPP as studied in this experiment, an impact that was consistent across all crop types. As previously discussed in the introduction to this thesis, heat stress during seed development is known to have a negative impact on yield parameters such as seed weight (Rashid et al, 2018; Magno Massuia de Almeida et al, 2021). Previous work on this topic focused on high levels of heat stress; the results of this experiment indicate that the effects of high temperature are still significant when they occur over the full course of development even at 24°C. As temperatures are predicted to increase in many crop-growing regions (Rosenzweig et al, 2014), it is important for reliable crop yield in the future to produce *Brassica napus* varieties which can deliver reliable yield at higher temperatures.

The Winter OSR crop type contains lines with a strong vernalisation requirement and was the crop type found here to be most affected by vernalisation temperature. Both SPP and TGW were reduced more than the mean reduction across the diversity set predicted by mixed linear modelling. For spring varieties, vernalisation temperature had an effect

on TGW but not SPP; the reduction in TGW at higher vernalisation temperatures was also only seen at 15°C vernalisation. Spring OSR generally has no or a low vernalisation requirement and has low FLC levels before vernalisation occurs (Hou et al, 2012; Calderwood et al, 2021). In Semi-winter varieties, vernalisation temperature had no significant effect on either SPP or TGW. Although not significant, the mean TGW of Semi-winter OSR was higher for 15°C vernalised plants compared to 5°C. In Semi-winter OSR, warmer temperatures can be sufficient to pass the comparatively low threshold for vernalisation; higher temperatures may then encourage greater plant growth.

From this crop type analysis, there is also an implication of *FLC* dynamics being related to seed size. In spring OSR, which only has a low vernalisation requirement, TGW was affected by vernalisation temperature, but SPP was not. This suggests that TGW may be more sensitive to the effects of vernalisation temperature than seed number.

When considering the effects of increased vernalisation temperature and increased seed maturation temperature across the phenotyping experiment, one feature of note is that both TGW and SPP were found to decrease overall. In general, TGW and SPP tend to have an inverse relationship i.e., plants which produce smaller seeds will then produce a larger number of seeds (Cai et al, 2014). This could mean that even in plants where seed development is affected, overall yield may be stabilized by this compensation mechanism. In contrast to this idea, the results of this experiment indicate that if yield is considered as a result of TGW and SPP combined, overall yield is decreased at higher temperatures due to a decrease in both yield parameters. These results likely suggest that TGW and SPP are both affected by increased temperatures but respond through separate mechanisms.

When checking the reproducibility of the results from the large glasshouse phenotyping experiment in CER conditions, the final analysis had a limited scope due to the loss of

some treatments to high levels of pre-harvest sprouting in plants matured at 24°C. In the original glasshouse phenotyping experiment, pre-harvest sprouting was also more pronounced in plants matured at 24°C. This made it difficult to confirm the results at 24°C. When comparing vernalisation temperatures with a maturation temperature of 18°C, the effect of vernalisation temperature was confirmed to be reproducible in three of the four lines. As a result of this experiment, Catana was selected for further experiments to explore the impacts of vernalisation temperature. In the original phenotyping experiment, Catana responded to both maturation and vernalisation temperature, and these results were consistent when the experiment was repeated.

Previous work has shown that *FLC* is expressed in seeds and affects seed dormancy (Chiang et al, 2009). The effects of vernalisation temperature on seed yield components seen in this experiment suggests that the effects of *FLC* on seeds may also impact seed development and regulation of seed size. Vernalisation temperature also had an impact on the number of seeds per pod. It is possible that the results of this experiment indicate a link between *FLC* and mechanisms of fertility such as development of pollen and ovules, or by otherwise affecting the rate of fertilization in the plant.

Some anther-specific reactivation of *FLC* has been seen in Arabidopsis (Sheldon et al, 2008; Choi et al, 2009). Sheldon (2008) reports that *FLC* is reactivated in the pollen mother cells and somatic tissues of the developing anther before being repressed again in the mature anther; Choi (2009) reported *FLC* reactivation in somatic tissues of pollen sacs. The expression of *FLC* in pollen has also been studied in one of the parent species which give rise to *Brassica napus*. *Brassica rapa* contains three *FLC* genes. The promoter of one of these *FLC* genes was found to drive gene expression in pollen (Hong et al, 2011). Additionally in *Brassica oleracea*, *BoFLC4* was found to be expressed in pollen grains and pollen tubes (Lin et al, 2005).

Although the effect of *FLC* expression on pollen development in *Brassica napus* is not known, it is possible that differences in vernalisation state could affect pollen development via *FLC* re-activation in the male germline before fertilization has occurred. Changes in pollen development associated with vernalisation and *FLC* could explain the reduced number of seeds per pod produced at higher vernalisation temperatures seen in this experiment, in particular in Winter OSR.

The number of seeds per pod could also be affected by mechanisms acting on ovule development, such as changes in *FLC* level impacting the number of ovules or their growth. It has been previously shown that the vernalisation state of the maternal plant has important effects on reactivation of *FLC* in the embryo (Luo et al, 2020), meaning that *FLC* expression in maternal tissues could be important for later differences in the seed caused by vernalisation. In *Nicotiana tabacum*, a group of MADS-box transcription factors which were identified as being similar to *FLC* were found to be expressed in a number of reproductive tissues including the stigma, style and ovary (Bai et al, 2019). Expression of *FLC* has also been linked to ovary development in hazel (*Corylus spp.)* due to the expression of *FLC* and other vernalisation-related genes in the developing ovary (Cheng et al, 2018).

In Arabidopsis, *FLC* was not found to be expressed in the ovary of vernalised plants, although plants which had not been vernalised showed *FLC* expression in the integuments (Sheldon et al, 2008). The results of this study indicate an effect of the vernalisation state of *FLC* in the ovule. Another study in Arabidopsis found that reduced function mutants of *BRR2a* cannot splice *FLC* transcripts correctly, leading to low expression levels of correctly spliced *FLC*. These mutants show an early flowering time phenotype, and also have reduced seed set, less efficient ovule fertilization and defective development of the female gametophytes (Mahrez et al, 2016). The BRR2A protein is part of the spliceosome which is involved in processing the transcripts of flowering time genes; while its effects are not specific to *FLC*, the results of this study provide an interesting additional link between *FLC* expression and fertility which could be studied further.

In *Brassica rapa, FLC* promoters drive gene expression in the stigma for three *FLC* genes. In another *Brassica* species, *Brassica campestris*, an *FLC* related gene identified as *BcFLC1* was found to cause reduced fertility when expressed in Arabidopsis (Liu et al, 2012), through both shortening of anther filaments and the reduced expression of *SEP3*, a MADS-box transcription factor gene involved in the development of ovules (Favaro et al, 2003; Kaufmann et al, 2009). Although the effect of changes in *FLC* expression on ovule development and female fertility in *Brassica napus* are not known, future studies which examine this relationship could provide further insights into the link between vernalisation temperature and reduced seed number per pod.

Chapter 3 - Environmental effects on germination and vigour

3.1 Introduction

Seed vigour is a complex trait which is an important contributor to crop performance. Rather than being a single measurable trait, seed vigour is determined by a number of factors including seed germinability and early seedling growth which contribute to agricultural performance. Seedling vigour is generally accepted to describe the properties of the seed that determine how well the seed performs during germination as well as the rate of uniformity of seedling behaviour during emergence and early growth. Seedling vigour is further understood to determine not only how the seeds perform to these measures under laboratory conditions, but also under conditions of environmental stress and after time spent in storage before sowing (Finch-Savage and Bassel, 2016). Effective testing of seedling vigour is carried out using a low-temperature stress test (Filho, 2015).

Temperature during seed maturation has been previously shown to impact seed dormancy and rate of germination. Increased maturation temperature in *Arabidopsis* is associated with a decrease in seed dormancy, with seeds produced at a higher temperature germinating at a greater rate compared to those produced at cooler temperatures (Springthorpe and Penfield, 2015). High temperatures during seed development have been shown to affect germination and vigour in *Brassica* (Chen et al, 2021) as well as in most other plant species (Penfield, 2017).

Drought resistance is also an important trait for crop performance in *Brassica napus*. Oilseed rape crops are sensitive to drought, with low water conditions and dry weather associated with reduced overall yield (Wan et al, 2009). Drought stress affects both germination and the early growth of the seedling, with lower soil water content negatively affecting growth (Khan et al, 2019). The crop is sown in late August to early September; dry soil may increasingly become an issue due to the likelihood of droughts occurring for longer and more intense periods (Trenberth et al, 2014), and predicted

decreases in rainfall in the East of England under high emissions climate scenarios (Semenov, 2007).

The ability of the crop to successfully establish and grow in the field is key for the overall growth and final yield (Finch-Savage and Bassel, 2016). The speed of germination at sub-optimal temperatures is closely related to vigour (Bradford et al 1993; Reed et al, 2022), meaning this is an important trait for assessing how a crop may perform.

3.2 Aims & Hypothesis

As seed maturation temperature is known to affect seed properties, the first hypothesis of this chapter is that seed maturation temperature affects the time taken for seeds to germinate.

The environment of the germinating seed is impacted by the water content of the soil, and so the second hypothesis tested here is that soil water content affects the time taken for seeds to germinate.

Seedling vigour and emergence is also important for overall plant growth, so the third hypothesis tested was that seed maturation temperature affects early seedling growth.

As drought can impact crop growth, the fourth hypothesis tested was that in different varieties of *Brassica napus,* soil water content affects seedling growth differently.

Different crop types of *Brassica napus* have been bred for adaptation to different environments, so the fifth hypothesis tested here was that crop type affects seedling vigour, as well as variation in response to seed maturation temperature and drought.

Following previous experiments on seed size in this project, the sixth hypothesis tested in this chapter is that there is a relationship between seed size and seedling size.

3.3 Results

3.3.1 Experimental design

This seed germination experiment was carried out at the Institute of Biological, Environmental & Rural Sciences (IBERS) in Aberystwyth, Wales. Seeds used for germination were those produced in the large glasshouse phenotyping experiment details in Chapter 2. For the germination experiment, four seeds per pot were sown and randomized by seed lot within trays in the germination glasshouse; watering samples were paired and placed together to minimize environmental variation between treatments. Two runs were used to germinate the full collection of varieties and treatments, and two replicates pots of each combination of variety, watering treatment and maturation temperature were germinated on each belt run for a total of four replicates. The experiment included replicate seed lots produced from the original phenotyping experiment as biological replicates, and replicate pots of each seed lot as technical replicates.

The seed lots selected for this experiment were all vernalised for 12 weeks at 10°C, as this treatment produced the greatest number of seed lots with sufficient replicates for use in the germination experiment. Any lines which did not have at least two replicates for both maturation temperature treatments were not included in the experiment. This left 84 lines used to assess germination and seedling emergence.

3.3.2. Impact of maturation temperature and soil water content on germination

Days to germination was scored manually at IBERS using hypocotyl unfurling as a benchmark for germination. A mixed linear model was used to estimate the effects of seed maturation temperature on time taken for seeds to germinate.

The formula for the general linear model used was:

$$D_{ijk} = \mu + MT_j + W_k + v_i + e_{ijk}$$

This model estimates the days to germination (D) of variety (v) i in maturation temperature (MT) j with watering treatment (W) k and e_{ijklm} is the residual term. In this model, v was a random effect while v, MT and W were fixed effects.

Figure 3.1a shows that there is a statistically significant effect of maturation temperature on time taken for seeds to germinate, although this is effect is extremely small with seeds produced at both temperatures taking approximately 4 days to germinate. Similarly in Figure 3.1b the effect of watering treatment on the time taken for seeds to germinate is statistically significant but small. Table 3.1 details the overall differences in these effects. The time taken for seeds produced at 24°C is 0.07 days faster compared to at 18°C, equating to a difference of less than two hours between the temperatures. The effect of watering treatment is slightly larger, with a difference of 0.15 days between the two treatments and plants germinating at a 65% soil water content germinating faster than at 35% SWC.

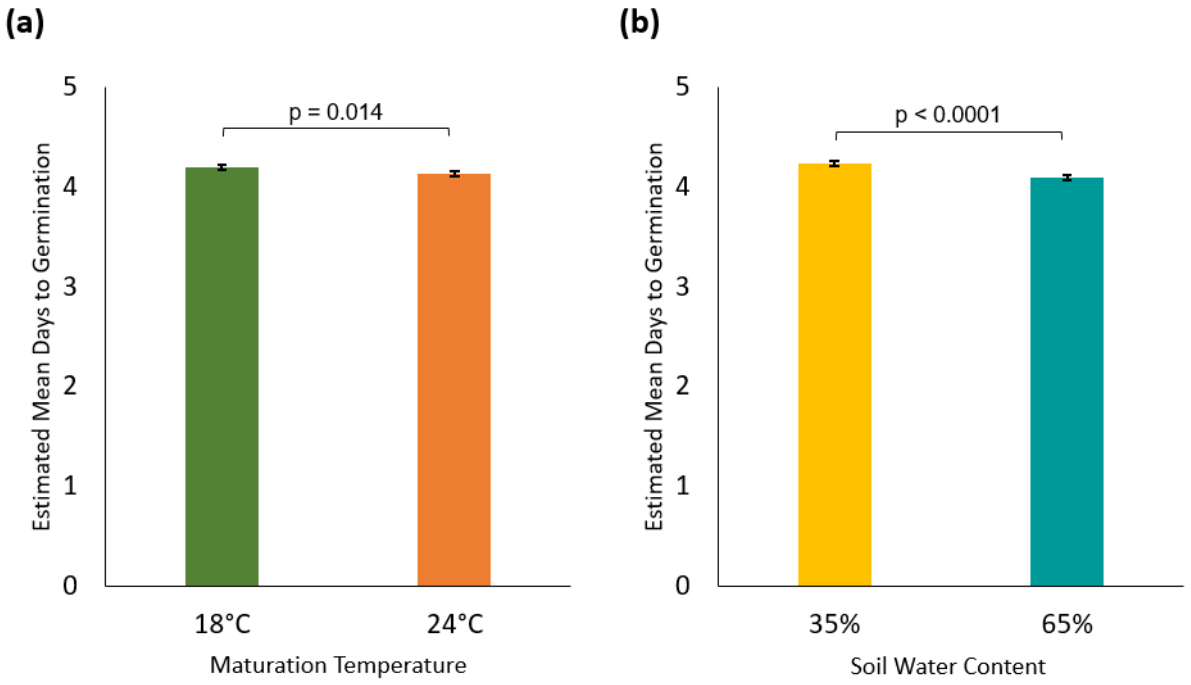

Figure 3.1: Linear model output of estimated mean days to germination, by (a) maturation temperature, (b) soil water content. Error bars show SE and annotations show p -values calculated using Welch's t-test.

Table 3.1: Details of linear model output of estimated mean days to germination. All values shown to 3 significant figures. SE = standard error. df = degrees of freedom. CF = 95% confidence limit.

	Estimated Mean Days to Germination	SE	df	Lower CL	Upper CL
	Maturation Temperature				
18°C	4.20	0.0244	115	4.16	4.25
24°C	4.13	0.0266	146	4.08	4.18
	Watering Treatment				
35%	4.24	0.0255	131	4.19	4.29
65%	4.09	0.0251	126	4.04	4.14

3.3.3. Creating a linear model to estimate environmental effects on seedling vigour

One of the factors used to evaluate seed vigour is early seedling growth and seedling establishment, with more vigourous seedlings showing faster growth in early development and increased rates of establishment. Seedlings with faster early growth rates, particularly in the 12-day timespan studied in this experiment, would reach a plant area of 1000mm^2 in a shorter timeframe. In this analysis, the time in days taken for a seedling to reach an area of 1000mm^2 was used to measure a threshold of seedling size as a proxy to estimate seedling vigour; as the experiments were not carried out at a low temperature, seedling emergence rather than seedling vigour was tested.

3.3.3.1. Effects of experimental factors on early seedling growth

To determine if blocking factors needed to be included in the linear model, ANOVA was used as the basis of exploratory analysis of experimental factors.

Seedlings were split across two experimental runs, due to size limitations on the belt. The effect of experimental run was examined using ANOVA to see if T1000 differed significantly between the two runs.

This ANOVA model estimated time taken for a seedling to reach 1000mm^2 (T1000) in experimental run (R) i, with the residual term e_i:

$$T1000_i = \mu + R_i + e_i$$

The results of this ANOVA, shown in Table 3.2, show that seedling emergence was significantly different between the two experimental runs. As a result, experimental run was included as a factor in the overall linear model to estimate T1000.

Table 3.2: ANOVA statistical results table for the effect of experimental run on T1000.

	Df	**Sum Squares**	**Mean Squares**	**F value**	**Pr(>F)**
Experimental Run	1	712	712.4	209.6	< 2E-16
Residuals	3028	10293	3.4		

Seedlings were imaged across the course of a full day, and so images were taken at different times. As plants were imaged in the order they were located on the belt, it is possible that any impact of location within the glasshouse would be reflected in time of day. To check if measurements were affected by the timing of imaging, the effect of hour of day was examined using ANOVA:

$$T1000_i = \mu + H_i + e_i$$

This ANOVA model estimated time taken for a seedling to reach 1000mm^2 (T1000) at hour of day (H) i, with the residual term e_i. The results of this ANOVA, shown in Table 3.3, show that time of day had a significant effect on T1000 measurements and so this was included in the model.

Table 3.3: ANOVA statistical results table for effect of hour on T1000.

	Df	Sum Squares	Mean Squares	F value	Pr(>F)
Hour	10	739	73.91	21.74	<2e-16
Residuals	3019	10266	3.4		

3.3.3.2 Optimizing linear model for estimating T1000

A general linear model was used to account for significant variables in the experiment and produce estimated values for time taken to reach 1000mm^2 (T1000) for each line. The formula used for the model was:

$$T_{ijklm} = \mu + MT_j \times W_k \times v_i + R_l + H_m + e_{ijklm}$$

This model estimates the days to 1000mm^2 of variety (v) i in maturation temperature (MT) j with watering treatment (W) k during experimental run (R) l and at hour of day (H) m, where e_{ijklm} is the residual term. In this model, experimental run and hour of day were included as random effects, while variety, maturation temperature and watering treatment were fixed effects. To optimise the model, interaction effects between variety, maturation temperature and watering were included.

From the results of the linear model (Table 3.4), it is apparent that variety, watering treatment and seed maturation temperature are all significant factors when considering time taken to reach 1000mm^2, with watering treatments having a greater effect than maturation temperature. No significant interaction was found between the soil water content and seed maturation temperature, suggesting that any effects these treatments had were additive. Significant interaction effects were found between variety and maturation temperature, and variety and soil water content. Responses to these treatments therefore differed between varieties.

Table 3.4: Statistical results table for preliminary model estimating T1000. DenDF = denominator degrees of freedom. Results produced using Type III Analysis of Variance Table with Satterthwaite's method.

	Sum	Mean	DF	DenDF	F	Pr(>F)

	Sq	Sq			value	
Variety	2139.96	25.78	83	2693.1	18.9257	< 2.2e-16
Maturation Temperature (MT)	16.45	16.45	1	2692.7	12.0752	0.00052
Soil Water Content (SWC)	2893.34	2893.34	1	2693.6	2123.85	< 2.2e-16
Variety x MT	418.13	5.5	76	2693.1	4.0385	< 2.2e-16
Variety x SWC	194.28	2.34	83	2693	1.7182	7.37E-05
MT x SWC	1	1	1	2692.5	0.7362	0.39097
Variety x MT x SWC	116.39	1.53	76	2693.1	1.1241	0.21887

3.3.3.3 Defining the optimized mixed linear model used to estimate T1000

After evaluating the significant parameters affecting T1000, a general linear model was created to account for significant variables in the experiment and produce estimated values for time taken to reach 1000mm^2 (T1000) for each line. The formula used for the model was:

$$T_{ijklm} = \mu + MT_j + W_k + v_i + (v_i \times MT_j) + (v_i \times W_k) + R_l + H_m + e_{ijklm}$$

This model estimates the days to 1000mm^2 of variety (v) i in maturation temperature (MT) j with watering treatment (W) k during experimental run (R) l and at hour of day (H) m, where e_{ijklm} is the residual term. In this model, experimental run and hour of day were included as random effects, while variety, maturation temperature and watering treatment were fixed effects. Interaction terms between variety and maturation temperature as well as between variety and watering treatment were included.

Table 3.5: Statistical results table for optimized model estimating T1000. DenDF = denominator degrees of freedom. Results produced using Type III Analysis of Variance Table with Satterthwaite's method.

	Sum Sq	Mean Sq	DF	DenDF	F value	Pr(>F)
Variety	2162.43	26.05	83	2770.2	19.0675	< 2.2e-16
Maturation Temperature (MT)	18.24	18.24	1	2770	13.3463	0.00026

Soil Water Content (SWC)	3066.36	3066.36	1	2770.6	2244.16	< 2.2e-16
Variety x MT	433.08	5.7	76	2770.2	4.1705	< 2.2e-16
Variety x SWC	184.05	2.22	83	2770.2	1.6229	0.00038

The results of the model show that all variety, maturation temperature and soil water content all significantly affected T1000, with soil water content having the largest effect. The interactions between variety and maturation temperature and variety and soil water content are also significant, indicating that responses to the treatments used in the experiment differ between varieties.

3.3.4 Impact of maturation temperature on early seedling growth

To examine the effect of seed maturation temperature on seedling emergence and early growth, the estimated means for days to reach 1000mm^2 produced from the linear model detailed above was compared between seeds produced by each variety at 18°C and at 24°C. The results of this model are shown in Figure 3.2.

The effect of maturation temperature on seedling emergence across the diversity set was statistically significant, as shown by the significant effect of maturation temperature overall and of the significant interaction between maturation temperature and variety (Table 3.5, above). In seven varieties, T1000 could not be estimated at 24°C. Of the remaining 77 varieties, there was an equal split in which seed maturation temperature produced seedlings reaching 1000mm^2 more quickly; 38 varieties reached 1000mm^2 more quickly after maturation at 18°C.

The largest difference was seen in the variety Shannon X Winner, which reached 1000mm^2 an estimated 3.07 days more slowly when seeds were matured at 24°C. The majority of varieties showed only a small difference, with 68 of 84 varieties (approximately 80%) tested predicted to have a difference of less than 1 day between the two maturation temperatures. Although the effect of maturation temperature is

statistically significant, the effect size is extremely small. The difference in plant behaviour would make no real difference in a crop growth situation.

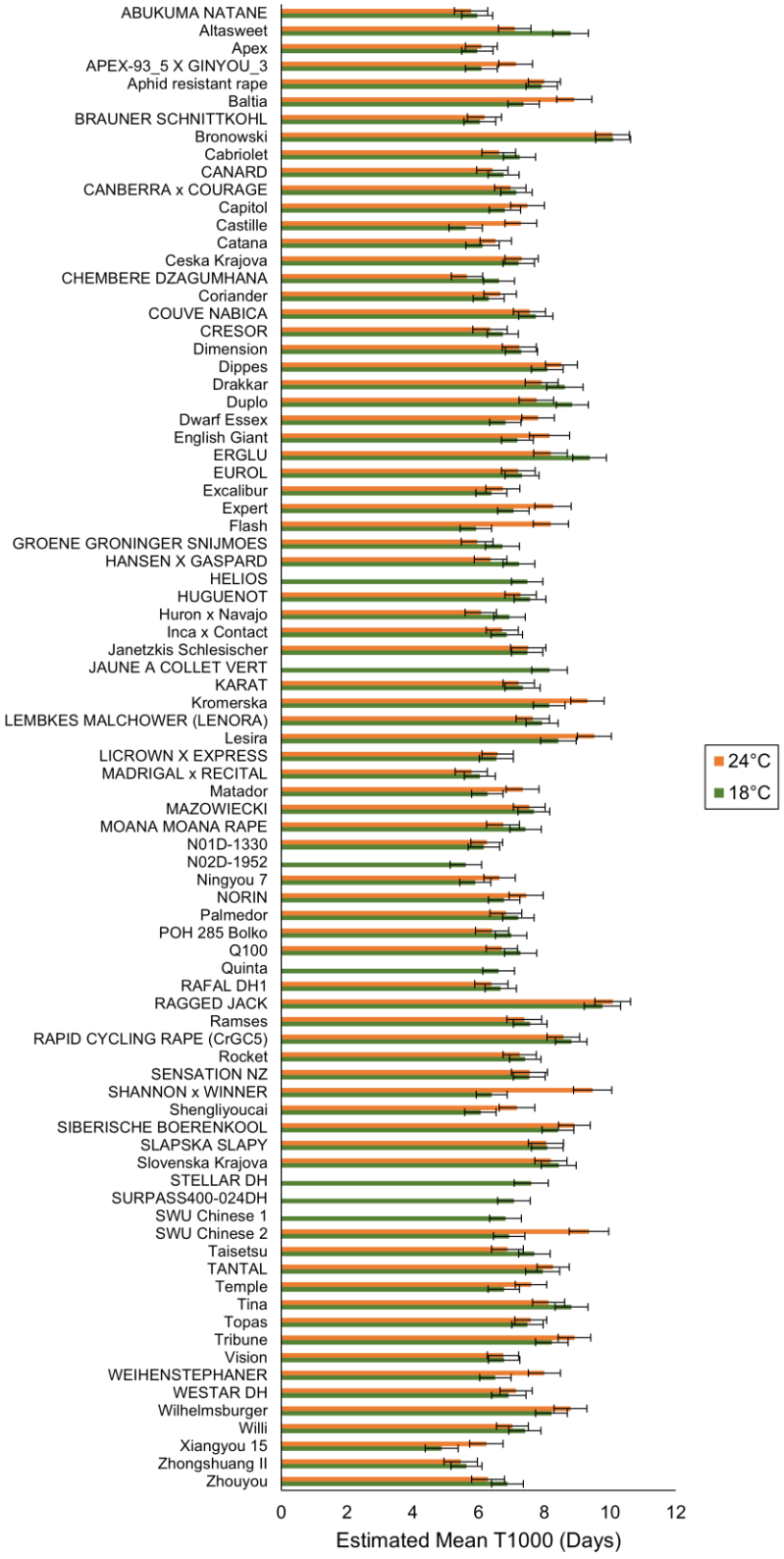
ABUKUMA NATANE
Altasweet
Apex
APEX-93_5 X GINYOU_3
Aphid resistant rape
Baltia
BRAUNER SCHNITTKOHL
Bronowski
Cabriolet
CANARD
CANBERRA x COURAGE
Capitol
Castille
Catana
Ceska Krajova
CHEMBERE DZAGUMHANA
Coriander
COUVE NABICA
CRESOR
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE GRONINGER SNIJMOES
HANSEN X GASPARD
HELIOS
HUGUENOT
Huron x Navajo
Inca x Contact
Janetzkis Schlesischer
JAUNE A COLLET VERT
KARAT
Kromerska
LEMBKES MALCHOWER (LENORA)
Lesira
LICROWN X EXPRESS
MADRIGAL x RECITAL
Matador
MAZOWIECKI
MOANA MOANA RAPE
N01D-1330
N02D-1952
Ningyou 7
NORIN
Palmedor
POH 285 Bolko
Q100
Quinta
RAFAL DH1
RAGGED JACK
Ramses
RAPID CYCLING RAPE (CrGC5)
Rocket
SENSATION NZ
SHANNON x WINNER
Shengliyoucai
SIBERISCHE BOERENKOOL
SLAPSKA SLAPY
Slovenska Krajova
STELLAR DH
SURPASS400-024DH
SWU Chinese 1
SWU Chinese 2
Taisetsu
TANTAL
Temple
Tina
Topas
Tribune
Vision
WEIHENSTEPHANER
WESTAR DH
Wilhelmsburger
Willi
Xiangyou 15
Zhongshuang II
Zhouyou
24°C
18°C
0
2
4
6
8
10
12
Estimated Mean T1000 (Days)

Figure 3.2: Estimated mean days to 1000mm^2 (T1000) by variety, as produced by mixed linear model. Error bars show standard error. Legend colour indicates seed maturation temperature.

3.3.5 Impact of soil water content on early seedling growth

To examine the effect of soil water content (SWC) on T1000, the estimated means for days to reach 1000mm^2 produced from the linear model detailed above was compared between seedlings from each variety grown at 35% SWC and 65% SWC. The results of this model are shown in Figure 3. Seven varieties did not have sufficient data to estimate T1000 and these were not included in the graphical output.

All varieties reach 1000mm^2 faster at 65% SWC compared to 35% SWC. The variety with the smallest difference was Moana Moana Rape, reaching 1000mm^2 0.98 days faster at 65% SWC than at 35%. The largest difference was seen in the variety Westar DH, reaching 1000mm^2 an estimated 3.09 days faster at 65% SWC compared to 35%. Differences in the extent of the response between varieties indicate that some varieties may be more drought resistant than others. However, across the diversity set reduced soil water content was always associated with slower seedling growth.

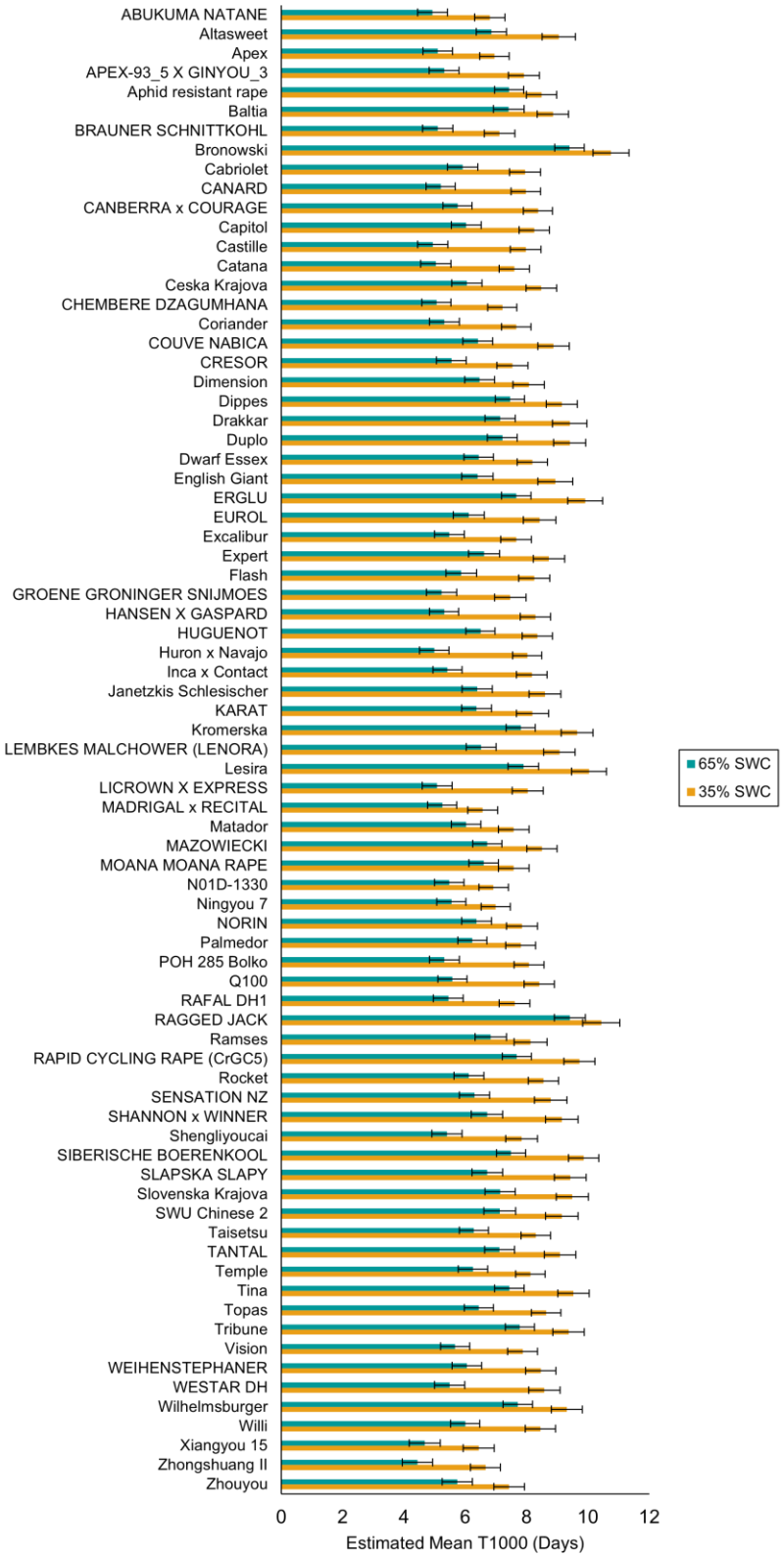
ABUKUMA NATANE
Altasweet
Apex
APEX-93_5 X GINYOU_3
Aphid resistant rape
Baltia
BRAUNER SCHNITTKOHL
Bronowski
Cabriolet
CANARD
CANBERRA x COURAGE
Capitol
Castille
Catana
Ceska Krajova
CHEMBERE DZAGUMHANA
Coriander
COUVE NABICA
CRESOR
Dimension
Dippes
Drakkar
Duplo
Dwarf Essex
English Giant
ERGLU
EUROL
Excalibur
Expert
Flash
GROENE GRONINGER SNIJMOES
HANSEN X GASPARD
HUGUENOT
Huron x Navajo
Inca x Contact
Janetzkis Schlesischer
KARAT
Kromerska
LEMBKES MALCHOWER (LENORA)
Lesira
LICROWN X EXPRESS
MADRIGAL x RECITAL
Matador
MAZOWIECKI
MOANA MOANA RAPE
N01D-1330
Ningyou 7
NORIN
Palmedor
POH 285 Bolko
Q100
RAFAL DH1
RAGGED JACK
Ramses
RAPID CYCLING RAPE (CrGC5)
Rocket
SENSATION NZ
SHANNON x WINNER
Shengliyoucai
SIBERISCHE BOERENKOOL
SLAPSKA SLAPY
Slovenska Krajova
SWU Chinese 2
Taisetsu
TANTAL
Temple
Tina
Topas
Tribune
Vision
WEIHENSTEPHANER
WESTAR DH
Wilhelmsburger
Willi
Xiangyou 15
Zhongshuang II
Zhouyou
65% SWC
35% SWC
0
2
4
6
8
10
12
Estimated Mean T1000 (Days)

Figure 3.3: Estimated mean days to 1000mm² (T1000) by variety, as produced by general linear model. Error bars show standard error. Legend colour indicates soil water content.

3.3.6 Impact of crop type on early seedling growth

As there were significant differences in seedling emergence between *Brassica napus* varieties, this could be related to crop type. To test this seedling emergence was compared between different crop types. ANOVA suggests significant differences in seed vigour between crop types (Table 3.6). The ANOVA used to discover if T1000 differed between crop type was:

$$T_{ijk} = CT_i \times MT_j \times W_k + e_{ijk}$$

This ANOVA model estimated days taken to reach 1000mm² (T) for crop type (CT) i at maturation temperature (MT) j and watering treatment (W) k, with a residual term e_{ijk}. The results of this ANOVA are described in Table 3.6.

Table 3.6: ANOVA table of statistical results for crop type effect on days to reach 1000mm2. df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	Pr(>F)
Crop Type	5	343	69	28.326	< 2E-16
Watering	1	3280	3280	1353.088	< 2E-16
Temperature	1	26	26	10.808	0.001
Crop Type : Watering	5	15	3	1.212	0.3006
Crop Type : Temperature	5	47	9	3.883	0.0016
Watering : Temperature	1	1	1	0.302	0.5829
Crop Type : Watering :	5	8	2	0.622	0.6833

Temperature					

As the ANOVA suggested differences in T1000 between crop types, the mean time taken to reach 1000mm² was calculated using the results of the ANOVA and Tukey post-hoc testing (p = 0.05). This confirmed differences in T1000 between crop types (Figure 3.4). Swedes take the longest time to reach 1000mm², while Semi-winter OSR reaches 1000mm² in the shortest time (Figure 3.4). This suggests that Semi-winter OSR varieties have the fastest early growth rate of seedlings, while swede varieties grow the most slowly.

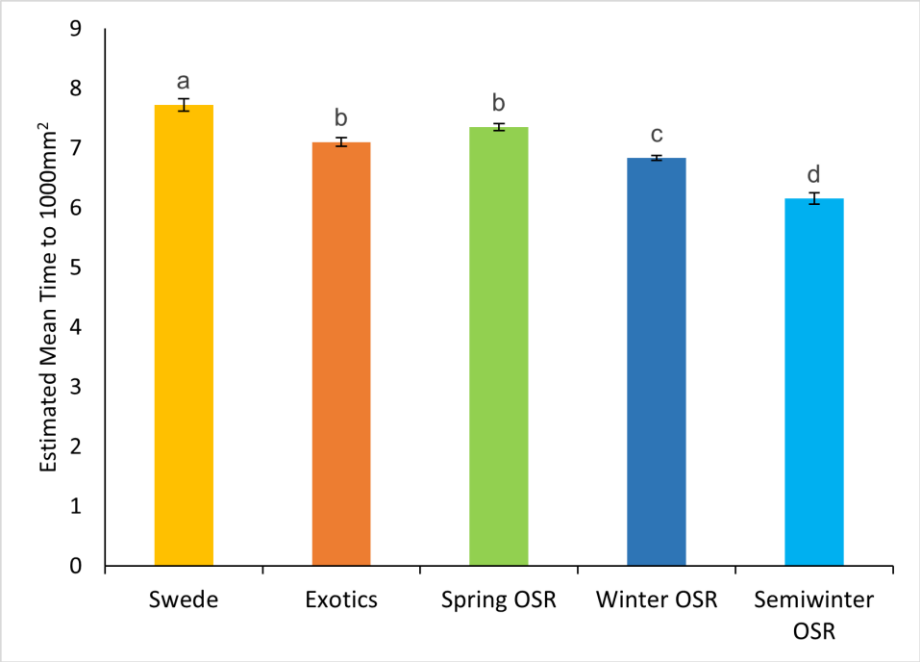

Figure 3.4: Estimated mean days to 1000mm² by crop type estimated using ANOVA. Error bars show standard error. Significance groups calculated using Tukey's post-hoc test, p = 0.05.

The results of the ANOVA also suggested interactions between crop type and maturation temperature. To further explore the effects of maturation temperature on T1000, a linear model was created. ANOVA results from Table 3.6 were used to identify interaction terms to include in the model. Experimental factors of belt run and hour

were previously identified as affecting T1000 so were included. The model used to estimate T1000 based on crop type was:

$$T_{ijklm} = \mu + MT_j + W_k + CT_i + (v_i \times MT_j) + R_l + H_m + e_{ijklm}$$

This model estimates the days to 1000mm^2 of crop type (CT) i in maturation temperature (MT) j with watering treatment (W) k during experimental run (R) l and at hour of day (H) m, where e_{ijklm} is the residual term. In this model, experimental run and hour of day were included as random effects, while variety, maturation temperature and watering treatment were fixed effects. Interaction between variety and maturation temperature was included.

The results of this model (Figure 3.5) indicate that crop types have different responses to seed maturation temperature.

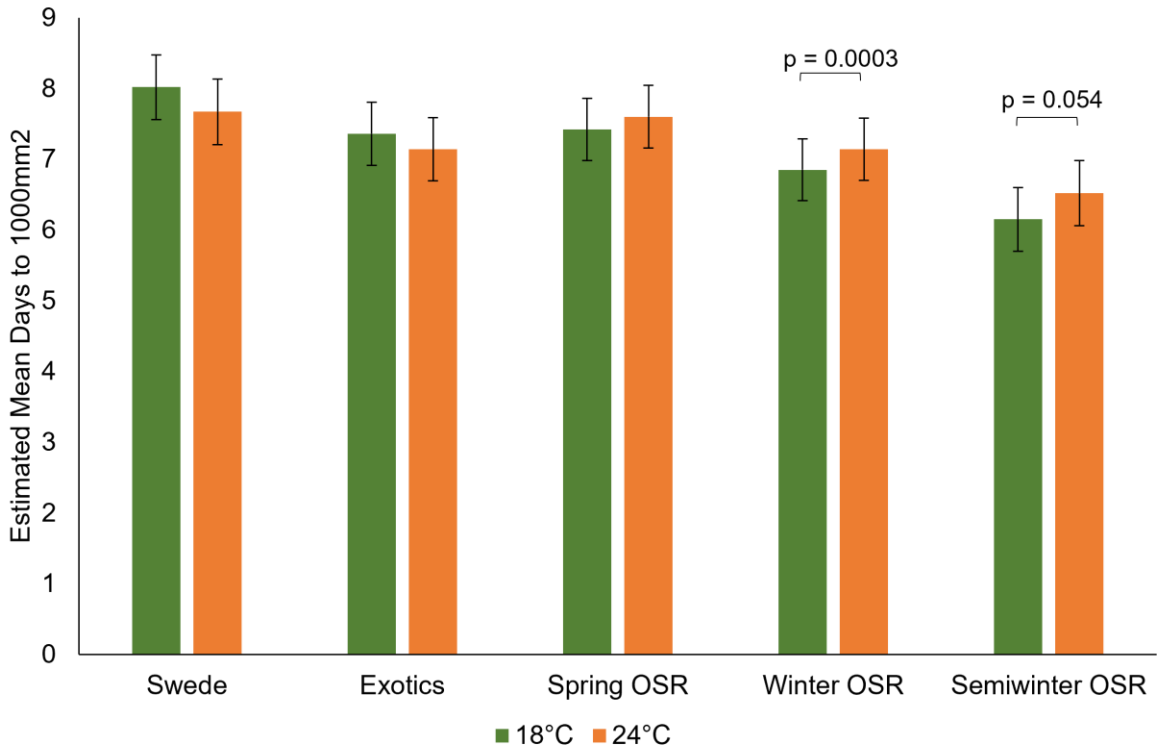

Figure 3.5: Estimated mean days to reach 1000mm2 (T1000) by crop type at two seed maturation temperatures. Error bars show standard error. Legend colours indicate seed maturation temperature. Paired bars with no p-value marked were not significantly different. P-values calculated using Welch's t-test.

The only crop type with a statistically significant effect of seed maturation temperature on T1000 was Winter OSR, with the difference in Semi-winter OSR approaching the threshold for statistical significance (Figure 3.5). However, even in Winter OSR the difference in time taken to reach 1000mm^2 was 0.28 days, equating to a real time difference of 6 – 7 hours. As seen when comparing individual varieties, although the effect of maturation temperature on T1000 was significant, the effect size was small with only extremely minor real-world impacts.

No interaction was identified between crop type and watering treatment, suggesting that all crop types respond similarly to drought as tested by the two watering treatments (Figure 3.6).

Across all crop types, seedlings took an estimated 2 days longer to reach 1000m^2 at 35% SWC compared to 65% SWC. This further confirms the strong negative effect of drought

on early seedling growth seen across all varieties. Although there were differences between the responses of individual varieties to reduced soil water content, these differences were not statistically significant when considered at the level of crop type.

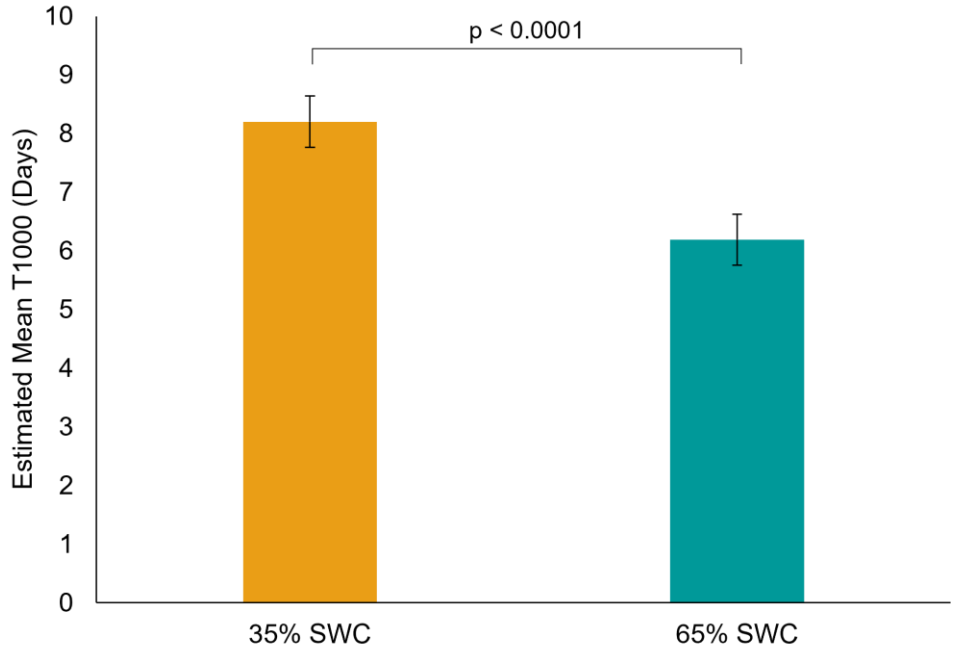

Figure 3.6: Estimated mean time to reach 1000mm² (T1000) by watering treatment across all crop types, as estimated using mixed linear model. Error bars show standard error. P-values calculated using Welch's t-test.

3.3.7 Relationship between seedling emergence and TGW

The line with the fastest T1000 in this experiment, Xiangyou 15, was also found to be the line with the highest overall thousand-grain weight in the large glasshouse phenotyping experiment. Swedes were the crop type found to have the slowest T1000. As swedes have not been bred as an oilseed crop, they have low TGW. This suggested there was a possible relationship between TGW and T1000. Exploratory correlation analysis was carried out to test this hypothesis.

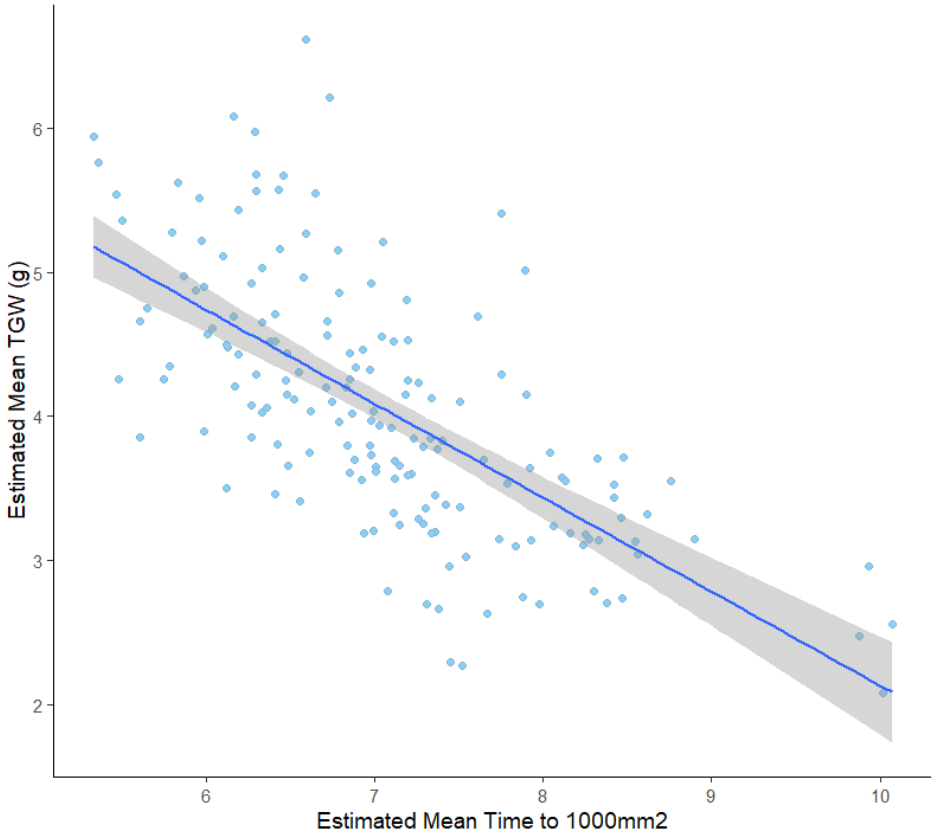

Figure 3.7: Scatterplot of TGW and T1000 by variety. Line of best fit shown in dark blue fitted using linear model method, formula T1000 = 9.87 – 0.695(TGW). Grey area around line of best fit indicates 95% confidence level.

Table 3.7: Results of correlation analysis between mean T1000 and TGW for each line to 3 significant figures.

Correlation test	**Value**	**p-value**
Spearman's rho	-0.723	$< 2 \times 10^{-16}$
Pearson correlation	-0.673	$< 2 \times 10^{-16}$
R^2	0.453	

The relationship between TGW and time taken to reach 1000mm^2 can be seen in Figure 3.7, where lines with a higher mean TGW have a shorter time taken to reach 1000mm^2. Table 3.7 also shows the results of the statistical correlation analysis. The results of Spearman and Pearson correlation tests indicate a strong correlation between the two factors, while the R-squared value of 0.453 indicates a moderate correlation.

From these results I can conclude that TGW is strongly negatively correlated with time taken to reach 1000mm^2, suggesting that larger seeds take less time to reach 1000mm^2.

These results suggest that TGW is an important contributor to early seedling growth. In future it would be interesting to expand this model to include TGW. This would be an interesting further analysis to fully understand the contribution of TGW to early seedling vigour in light of other factors such seed maturation temperature and soil water content.

3.4 Conclusions & Discussion

The results described in this chapter aimed to describe the effects of plant environment on germination and seedling growth. The environmental effects studied were the water content of the soil to explore the effects of drought on germinating seeds, as well as the seed maturation temperature experienced by the seeds prior to harvesting. Seed maturation temperature was studied due to its effects on seeds previously detailed in Chapter 2 of this thesis, as well as its known role in affecting seed germination (Fenner, 1991; Penfield and Springthorpe, 2015). Soil water content (SWC) was also studied as the ability of a *Brassica* crop germinate and grow well in dry soils is of particular importance to farmers and breeders. It is known that drought conditions can affect germination and crop growth and so varieties of *Brassica napus* which can withstand low water conditions would be valuable (Nelson et al., 2022). Preliminary experiments carried out at IBERS showed that 65% SWC was sufficient for *Brassica napus* seeds to germinate reliably, while 35% SWC was a condition of relative drought.

The initial experiments on germination show that the rate of germination across the dataset was not strongly affected by either seed maturation temperature or by soil water content. Seeds germinated slightly faster after maturation at 24°C compared to 18°C, which is similar to the effect of increasing maturation temperature seen in Arabidopsis (Springthorpe and Penfield, 2015). Germination rate - defined in this experiment as the time taken for the hypocotyl hook to unfurl - was also slightly faster in seeds germinated

at 65% SWC compared to 35% SWC, which corresponds to previous studies indicating that drought has a negative impact on germination (Zhang et al, 2014). Although there was a statistical significance, the effect size for both sets of treatments was extremely small and so improving the germination rate is not a likely target for improving tolerance to high seed development temperatures or for improving drought tolerance.

Early seedling growth was also studied, using the time taken to reach 1000mm^2 as a measure of early plant growth, and as such a proxy for seedling emergence. Similar to the effect of maturation temperature on germination rate, the impact of maturation temperature on seedling vigour was statistically significant but extremely small. Although there was variation between lines, this variation was only a minor effect. This shows that in the *Brassica napus* DFFS, the maturation temperature experienced by the seed does not have a major impact on early seedling growth. These results contrast the effect of maturation temperature seen in *Brassica oleracea,* where higher seed maturation temperature was found to increase the speed of seedling establishment, suggested to be due to differences in ABA regulation in the endosperm (Chen et al, 2021). However, it should be noted that the experiment for seedling emergence was stopped after 12 days. Further differences in seedling growth between maturation temperatures may have become apparent with a longer timeframe for measurements. It may also be the case that *Brassica napus* seedling vigour is controlled by different mechanisms to those identified in *Brassica oleracea*.

Soil water content had a strong impact on early seedling growth, with seedlings growing at 35% SWC taking more time to reach 1000mm^2 compared to seedlings at 65% SWC, suggesting a slower rate of growth during early seedling growth. This was true for all varieties; although variation in the scale of response to drought was identified between varieties, seedlings grown in drier conditions took 1 – 3 days longer to reach 1000mm^2. These results suggest that drought conditions always have a negative impact on early growth in *Brassica napus*. No variety could be identified as resistant to drought from the results of this experiment. This supports previous assessments of drought effects on seedling establishment (Batool et al, 2022). It is possible that further genome association analysis could identify genes of interest in the differences in these responses, although this did not form part of this project.

The effect of crop type was also found to have a significant impact on seedling emergence, which varied depending on maturation temperature. Only Winter OSR was significantly affected by maturation temperature. When considering responses of different crop types to drought there was no significant variation, with all crop types negatively affected by lower soil water content. No crop type was found to have a significantly improved tolerance to drought. In the oilseed crop *Camelina sativa*, spring-type crops were found to have improved germination performance compared to winter-type and crop type was also identified as a factor contributing to drought tolerance (Čanak et al, 2020). This could suggest that some feature of winter oilseed crops in contrast to spring types may be related to later seed performance.

These results highlight the important of using irrigation or timing sowing appropriately for *Brassica napus*, as it is clear that dry soil provides adverse conditions for young seedling vigour, which is important for overall crop establishment and the ability to withstand further stresses from biotic and abiotic sources.

The results of these experiments also indicated that larger seeds produced seedlings which reached 1000mm^2 more quickly. The crop type with the highest seedling vigour was also Semi-winter OSR, which is also the crop type which generally produces the largest seeds. Additionally, seeds which were produced at 18°C were larger compared to those produced at 24°C; a seed maturation temperature of 18°C was also associated with improved seedling vigour, suggesting this effect may be related to a difference in seed size. Seed weight has been previously noted to impact seedling growth characteristics in *Brassica oleracea* (Bettey et al, 2000). Larger seeds are likely to contain larger embryos, as by the final stage of seed development the embryo makes up the majority of the internal seed structure. Additionally, seeds produced under heat stress may be smaller and have defective embryo development (Mácová et al, 2022); this may further relate reduced seed growth to seedling vigour. It is possible that improved seedling vigour is related to favourable development of the embryo prior to germination and emergence.

Chapter 4 - Using associative transcriptomics to identify the genetic basis of temperature responses in *Brassica napus*

4.1 Introduction

Previous experiments in this thesis have shown that yield parameters (TGW and SPP) are affected by temperature during vernalisation and during seed maturation. However, these effects varied by both variety and crop type, suggesting a range of diversity in *Brassica napus*. The question that arises from these results is then: what causes the difference between these varieties?

Different crop types and varieties of *Brassica napus* have been bred for adaptation to different environments, and selection for different phenotypes has results in genetic differentiation of the genes controlling these phenotypes. Using GWAS can then reveal what the underlying genetic differences may be between lines with contrasting responses to temperature, and so improve our knowledge of the genetic basis of temperature responses in *Brassica napus*. Additionally, identifying genes which confer resistance to high seed maturation temperature or low vernalisation temperature, and are associated with reliable yield, could be valuable information for future breeding strategies to produce climate-change resilient crops.

Past studies have been carried out in *Brassica napus* to identify genes of interest for control of seed size and seed number per pod, as well as for a range of other yield-related and agronomic traits. Many studies have been based on identifying QTLs for yield-related traits (Shi et al, 2009; Basunanda et al, 2010; Fan et al, 2010; Ding et al, 2012; Yang et al, 2012; Zhang et al, 2012). While this can provide a general location for a candidate gene, it is unlikely to be specific enough to identify a single gene. Additionally, the genetic variation in a QTL mapping experiment is generally based only on the genetic differences between the parents, which may not represent the breadth of variety seen across a population.

Studies have also been carried out using GWAS to identify candidate genes controlling yield-related traits in *Brassica napus* (Li et al, 2014; Körber et al, 2016; Dong et al, 2018; Khan et al, 2019; Pal et al, 2021). A common feature of these studies is that they use phenotyping data obtained from field studies carried out across a number of growing seasons. This provides data which is relevant to how *Brassica napus* grows in the field but has the additional element that growing conditions are uncontrolled. Analyses may have to account for environmental factors (Pal et al, 2021), such as differences in rainfall, soil types and disease in the plants. Other recent studies have explored the effect of temperature on reproductive development in *Brassica napus* without using genetic approaches to identify specific candidate genes responsible for these effects (Magno Massuia de Almeida et al, 2021; Mácová et al, 2022).

Previous studies have varied in how many SNPs are used for identifying differences in *Brassica napus* with SNP density ranging from 6,000 to upwards of 600,000 SNPs across the genome (Körber et al, 2016; Dong et al, 2018). The number of varieties included in these analyses also ranges from 157 to several hundreds of varieties. An increased number of varieties provides a wider range of information on how genetic background is linked to phenotype.

The analysis detailed in this chapter uses a GWAS approach with a dataset of 350k SNPs produced from RNAseq analysis (Harper et al, 2012). This SNP array provides more coverage than many studies carried out elsewhere and has also been used to successfully identify sources of genetic variation in previous studies, including variation linked to seed size (Miller et al, 2019). The data for this analysis was gathered from the previous phenotyping experiment, which used CERs and temperature-controlled glasshouses to ensure consistent growth temperatures across 12 different temperature treatments.

The use of consistent growth temperatures also allowed for clear comparisons between temperatures, allowing the effects of individual treatments for both vernalisation and seed maturation to be explored. Additionally, all plants in the phenotyping were grown under controlled conditions, meaning that all plants received the same watering regime and were grown in the same soil. Growth in controlled conditions rather the field enabled the impacts of disease to be limited using preventative and treatment measures.

GWAS uses the combination of phenotypic measurement data, transcriptome sequencing and kinship analysis to identify genes which are likely to be associated with changes in phenotype within a population (Korte and Farlow, 2013). Linkage disequilibrium mapping uses the non-random association of genes at different mapping to locate markers associated with a particular phenotype, such as yield resistance to temperature stress (Atwell et al, 2011; Bac-Molenaar et al, 2015; Wei et al, 2021). GWAS requires diverse germplasm and a large number of varieties, so that sufficient genetic variation is present to yield useful results. Transcriptome data from RNAseq is here used to provide sequence data; this allows the genetic variation associated with phenotypic differences to be identified.

The overall hypothesis being tested in this chapter is that there are genes responsible for the temperature responses discovered so far in *Brassica napus*, and that these genes can be identified using GWAS. Once these genes have been identified, this chapter aims to characterize the impact of SNPs found in the GWAS analysis and produce candidate genes for further study.

4.2 Results

The results from the phenotyping experiment, described in Chapter 2, were used as the basis for the GWAS analysis described in this chapter. As plants were grown under 12 different temperature treatments, 12 separate GWAS analyses were run for both SPP and TGW. This enabled the identification of associations which may have been significant in some treatments. Once all of these individual analyses were run, all 12 treatments were combined to carry out a meta-analysis of TGW and SPP data. Datasets to compare the effects of temperature by using ratios were also produced and run as individual GWAS analyses. The full list of GWAS analyses run are listed in Table 4.1.

Table 4.1: Datasets used to run GWAS analyses. The first numbers indicate vernalisation temperature; the seconds numbers indicate weeks of vernalisation; the third number indicates maturation temperature. Where "vs" is used, this indicates a ratio of the trait value (either TGW or SPP) between treatments.

Original Treatments	***Comparative Treatments***

5°C 6 wks 18°C	5°C vs 15°C 6 wks 18°C
5°C 6 wks 24°C	5°C vs 15°C 6 wks 24°C
5°C 12 wks 18°C	5°C vs 15°C 12 wks 18°C
5°C 12 wks 24°C	5°C vs 15°C 12 wks 24°C
10°C 6 wks 18°C	5°C 6 wks 18°C vs 24°C
10°C 6 wks 24°C	5°C 12 wks 18°C vs 24°C
10°C 12 wks 18°C	10°C 6 wks 18°C vs 24°C
10°C 12 wks 24°C	10°C 12 wks 18°C vs 24°C
15°C 6 wks 18°C	15°C 6 wks 18°C vs 24°C
15°C 6 wks 24°C	15°C 12 wks 18°C vs 24°C
15°C 12 wks 18°C	***Meta-analysis***
15°C 12 wks 24°C	All original treatments combined

Due to the large amount of analysis used to identify genetic associations with TGW and SPP, this chapter will detail for one treatment how the analysis was carried out. After this, summary tables will be used, and the data presented to identify significant SNPs will be identified from across the breadth of GWAS analyses.

4.2.1 Example GWAS

A GWAS analysis was run on TGW data for the temperature treatment 5°C vernalisation for 6 weeks and maturation temperature 18°C, using TASSEL software (Bradbury et al, 2007). The dataset used for this experiment contained 355,536 SNPs from leaf transcriptome data across the DFFS. Minor allele frequency was set to 5% to remove any rare variant SNPs which may cause false positives. This left 177,725 SNPs across the dataset.

GWAS was performed using a general linear model:

$$\text{Trait} = \mu + Q1 + Q2 + Q3 + \text{marker}$$

In this model, Q1, Q2 and Q3 are measures of kinship used to account for relatedness. This kinship matrix was produced using STRUCTURE (matrix created by Guangyuan

Lu, unpublished data; Pritchard et al, 2000). The number of groups in the population, *k*, was tested from 1 to 9 with the highest ΔK value found at k = 4.

A mixed linear model was also run to compare how effectively the models could fit the data.

$$\text{Trait} = \mu + Q1 + Q2 + Q3 + \text{marker} + k$$

Where k is the kinship adjustment run as a random effect and Q1 + Q2 + Q3 are part of the fixed effect Q matrix to adjust for relatedness.

Goodness of fit describes how well the model produced fits the observations expected from the population. In order to check the goodness of fit of the two models run in TASSEL, quantile-quantile plots (Q-Q plots) showing observed quantile distributions from the model against expected quantile distributions can be examined. From the Q-Q plots, it can be seen that using a general linear model to run GWAS results in some under-correction of the model, increasing the risk of false positives (Figure 4.1). In contrast, using a mixed linear model to run GWAS (Figure 4.2) results in over-correction of the associations. This increases the likelihood of false negative results, meaning that actual associations can be missed. To ensure that existing associations were able to be identified, the GLM approach was used for all analyses.

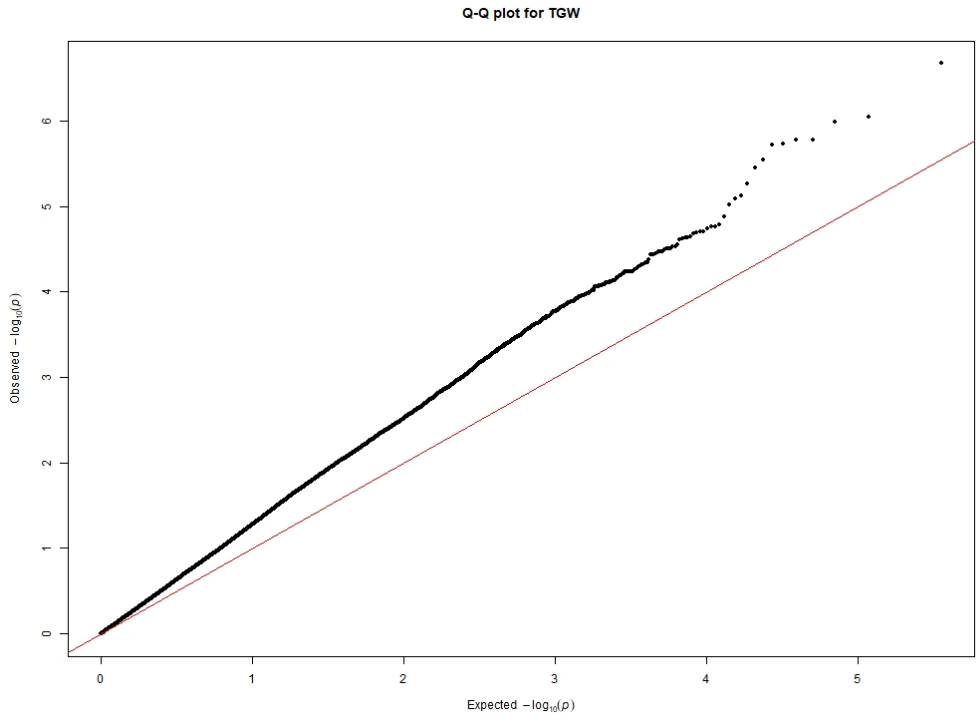

Figure 4.1: Q-Q plot for TGW at 5°C 6 weeks vernalisation, 18°C maturation temperature. Statistics produced by GLM.

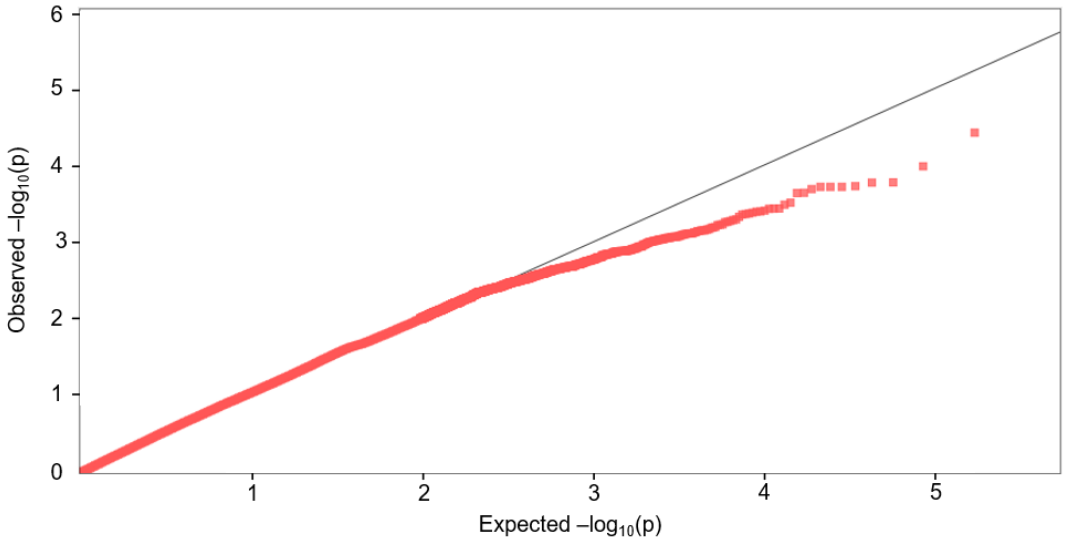

Figure 4.2: Q-Q plot for TGW at 5°C 6 weeks vernalisation, 18°C maturation temperature. Statistics produced by MLM.

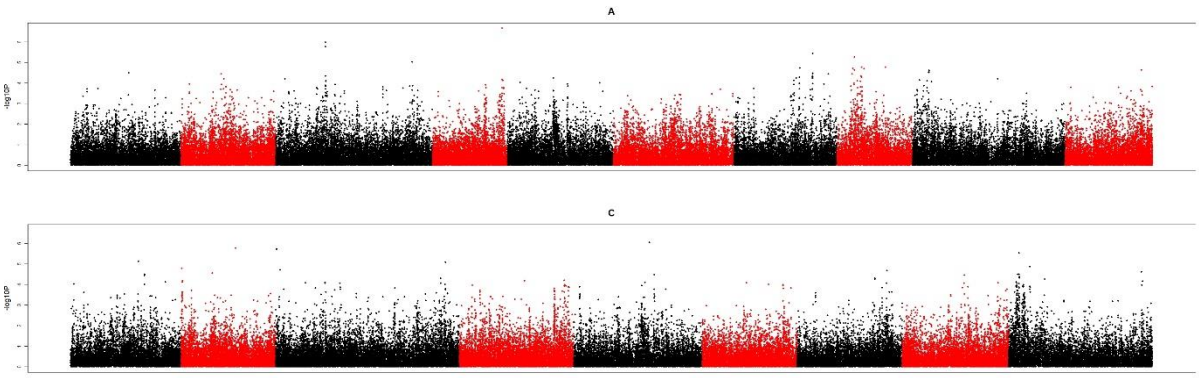

Figure 4.3: Manhattan plot for A and C genomes of Brassica napus produced from GLM of TGW for plants grown at 5°C vernalisation for 6 weeks and matured at 18°C.

The results of the GLM run on TGW data for this treatment were visualized using a Manhattan plot (Figure 4.3). The p-value was adjusted using the Bonferroni method:

$$\alpha_B = \alpha/m$$

Where α is the original p-value cut-off of 0.05, α_B is the adjusted p-value and m is the number of markers (177,725). The calculated Bonferroni adjusted p-value is 2.813 x 10^{-7}, which results in a significant -log(p) value of 6.55. Using this cut-off, only a single SNP marker has a high enough value to be significant (Table 4.2), even when using a slightly under-correcting GLM.

Table 4.2: Significant SNP markers identified using Bonferroni-adjusted p-value (2.813E-07), from GLM of TGW at 5°C 6wks 18°C.

Marker	Chr	Position	p
Cab045199.1:468:G	A04	21589385	2.14E-07

As the Bonferroni adjusted p-values were too stringent to produce candidate genes, associations were instead identified manually (Figure 4.4) to locate areas of interest within the genome. Figure 4.4a shows the distribution of TGW of seeds produced from plants vernalised at 5°C for 6 weeks and matured at 18°C. The group with the greatest frequency is 3.5 – 4.0 g, with TGW spanning a range of 2.0 – 8.5 g showing variation across the group. Using this method, three regions of interest were identified on

chromosomes A03, A04 and A07 (Table 4.2). The Manhattan plots used to identify these regions are shown in Figure 4.4b – d, with the regions of interest highlighted in blue. These regions were selected due to the appearance of increased probability of closely clustered genetic associations with TGW compared against the background level of noise in the plot.

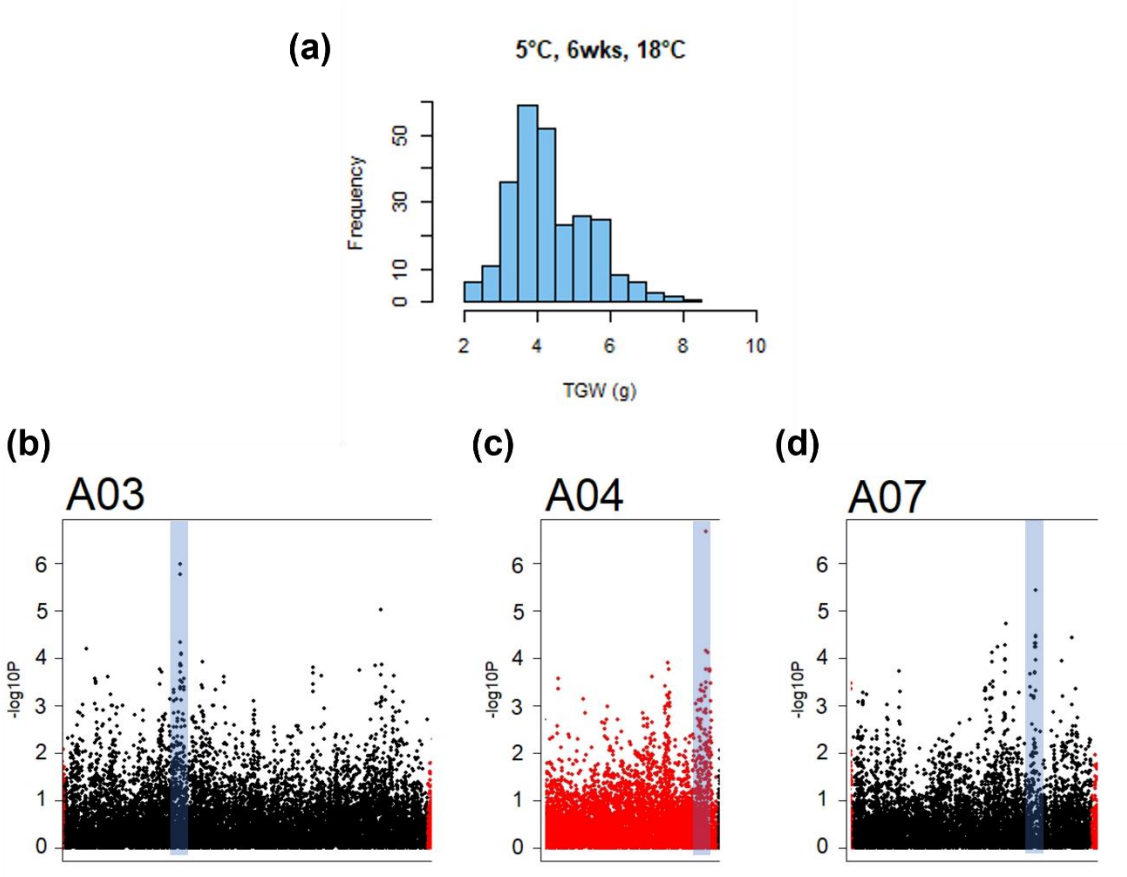

Figure 4.4: (a) Distribution of TGW for seeds produced following 5°C 6wks vernalisation treatment and 18°C maturation temperature treatment. (b)-(d) shows Manhattan plots of selected associations, located on chromosome: (b) A03; (c) A04; (d) A07.

The SNP with the most significant p-value in the region identified on chromosome A03 (p = 1.05E-06) was found in the gene BnaA03g15800D. The SNP with the most significant p-value in the region identified on chromosome A04 (p = 2.14E-07) was found in the gene BnaA04g25390D. The SNP with the most significant p-value in the

region identified on chromosome A07 (p = 3.60E-06) was found in the gene BnaA07g28280D. This is an orthologue of Arabidopsis gene AT1G66580.

Table 4.3: SNP markers manually identified as possible associations of interest, from GLM of TGW at 5°C 6wks 18°C.

Marker	Chr	Position	p	Minor Allele Frequency
Cab002689.1:2435:C	A03	8737064	1.05E-06	26
Cab002689.1:2446:A	A03	8737064	1.68E-06	33
Cab002689.1:3435:A	A03	8737064	4.58E-05	26
Cab002693.1:93:C	A03	8754976	8.37E-05	23
Cab002695.1:129:C	A03	8765839	7.87E-05	22
Cab002695.1:198:C	A03	8765839	8.29E-05	18
Cab045199.1:468:G	A04	21589385	2.14E-07	7
Cab045199.1:142:C	A04	21589385	6.82E-05	8
Cab021443.1:186:A	A07	24170412	3.60E-06	5
Cab021443.1:273:T	A07	24170412	5.81E-05	9
Cab021443.1:276:T	A07	24170412	5.81E-05	9

GWAS was run in this manner for all temperature treatments, for both TGW and SPP. As well as running GWAS on TGW across the treatments, GWAS was also carried out using comparative datasets. For these, the difference in SPP and TGW was calculated between vernalisation temperatures of 5°C and 15°C, and between 18°C and 24°C maturation temperatures. This led to a total of 22 GWAS analyses for each trait.

The data produced by the GWAS allowed for identification of which *Brassica napus* gene the associated SNP marker was present in. However, more precise information which would have enabled analysis of how each SNP affected the function of the gene or protein was not available without further sequencing. As a result of this the following

analyses can assess the relationship between the SNP present and the phenotype but cannot examine the impact of the SNP on the activity or function of the gene.

4.2.2 Identification of associations of interest for TGW

As shown in the example in section 4.2.1, associations discovered in GWAS results were manually assessed to identify SNPs of interest. As using the Bonferroni-adjusted p-value for most treatments produced no significant results, the associations were also assessed using *ad hoc* p-values across all GWAS analyses. Here the significance threshold was adjusted manually to identify genes which could plausibly be associated with differences in TGW but did not pass the Bonferroni-adjusted significance threshold. A p-value of 0.0001 equates to a -log(p) of 4, which in these Manhattan plots filters out the majority of noise. The associations identified for TGW are summarized in Table 4.3.

Table 4.4: Summary of GWAS association locations for TGW, with region located indicated by Mb on the chromosome as described in the pangenome (He et al, 2015). Totals columns show the number of treatments which produced SNPs in each region below the significance threshold indicated in the column header.

Chr	Location (Mb)	Total < 0.001	Total < 0.0001	Total < 2.81E-07
A01	25.9 - 26.1	3	1	0
A03	8.7 - 8.8	6	3	0
A03	14.8 - 14.9	3	1	0
A03	16.2 - 16.4	4	2	0
A04	21.5 - 21.7	10	5	1
A06	26.5 - 26.8	4	2	0
A07	24.1 - 24.2	9	5	0
A08	17.8 - 18	4	1	0
C06	36.8 - 37.7	13	5	0
C08	25.3 - 25.9	5	1	0

For TGW, eight associations were identified on the A genome and two on the C genome. Full details can be found in Appendix B, Table B8.10. Only one of these associations on chromosome A04 passed the Bonferroni threshold for significance. Additional lower

thresholds were used to identify which associations might be found across a range of treatments, comprised of data from the 12 temperature treatments used in the glasshouse phenotyping experiment as well as an additional 10 datasets created by comparing the trait values between contrasting temperatures.

4.2.2.1 Meta-analysis of GWAS results for TGW

Genomic associations with TGW and SPP were also explored using a METAL meta-analysis (Willer et al, 2010) combining data from all 12 of the temperature treatments used in the glasshouse phenotyping experiment. Meta-analysis allows the association p-values produced from separate GWAS analyses to be combined, allowing for a more robust evaluation of the effect of associations found across multiple studies. The distribution of TGW across all of the 12 temperature treatments combined in this meta-analysis can be seen in Figure 4.5, indicating the spread of TGW values assessed in this analysis.

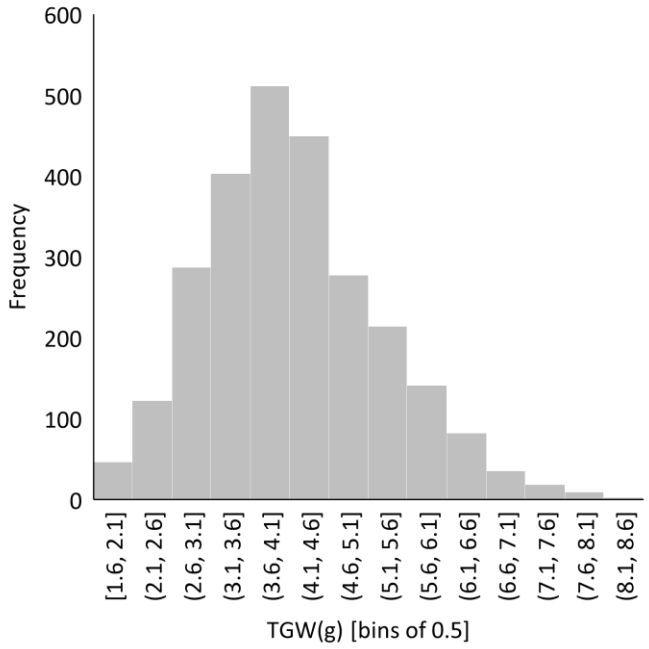

Figure 4.5: Distribution of TGW across all 12 treatments used in meta-analysis.

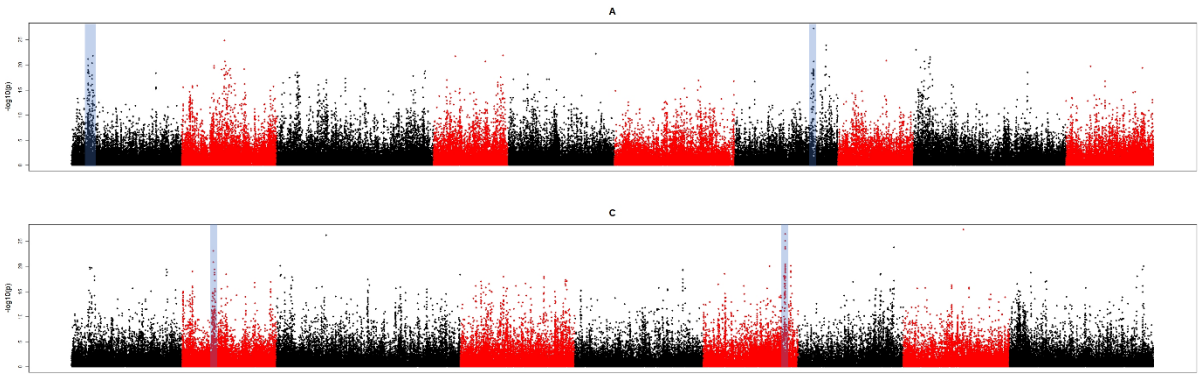

Figure 4.6: Manhattan plot output from METAL meta-analysis across all 12 treatments for TGW.

Four main associations were identified on chromosomes A01, A07, C02 and C06 (Figure 4.6) for TGW using the meta-analysis method. Associations were identified using the most significant p-values with consistent associations at a single location.

The details of the associations identified from the meta-analysis for TGW are described in Table 4.5.

Table 4.5: SNP associations found using METAL meta-analysis across 12 treatments for traits TGW. Peak location range is in Mb from beginning of the indicated chromosome. p-values are the most significant in the given range.

Chromosome	Peak Location Range	p-value
A01	2.0 - 2.7	1.69E-22
A07	24.1 - 24.2	5.49E-28
C02	6.9 - 7.0	7.84E-24
C06	37.6 - 37.7	3.54E-27

4.2.3 Identification of loci of interest for seed weight

All associations identified from individual GWAS treatments and meta-analysis were further explored to identify possible candidate genes for control of seed size. Where these associations have been identified, it is possible that they are affecting seed weight

in general, or specifically the effect of temperature on seed weight. To identify which gene was most likely to be responsible for the association between the SNP and TGW, nearby SNP markers were correlated with which genes were nearby. The SNP markers are produced from the SNP array (Harper et al, 2012). These were then linked to their orthologues in *Arabidopsis thaliana* using the *Brassica napus* pangenome (He et al, 2015). Previously, a distance of 150kb has been used to account for linkage disequilibrium in *Brassica napus* (Pan et al, 2021), so genes 150kb upstream or downstream of the most significant SNP were considered.

4.2.3.1 Association with TGW on Chromosome A01 identified in the gene BnaA01g25880D

An association was identified with TGW at 25.9Mb on Chromosome A01. Preliminary analysis identified a SNP in this region (Figure 4.7), where the SNP was a change between a C and T. This SNP was identified in the gene BnaA01g25950D, which codes for a protein of unknown function and has been identified as an orthologue of AT3G19680. Genes within 150kb of this SNP are listed in Table B1 (Appendix B). From the genes in this region of interest, the most likely gene to be related to changes in seed size was identified as an orthologue of AT3G19700 (*IKU2*), notated in *Brassica napus* as BnaA01g25940D. BnaA01g25950D, the gene in which this SNP was identified, is 1.1kb away from *BnaA01.IKU2*.

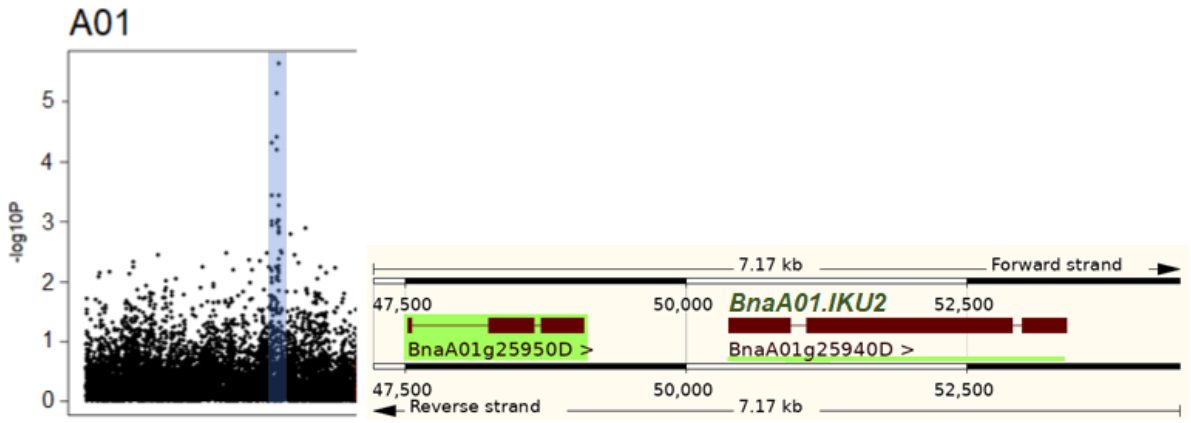

Figure 4.7: (a) Preliminary Manhattan plot of chromosome A01 (15°C 6wks 24°C) (b) SNP marker and candidate gene locations from chromosome A01. Gene containing the SNP is highlighted in green; BnaA01.IKU2 is underlined in green.

Table 4.6: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaA01g25950D affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom

	Df	**Sum Squares**	**Mean Squares**	**F value**	**P value**
Vernalisation Temperature					
SNP Call	3	274.7	91.56	79.622	< 2e-16
Vernalisation Temperature (VT)	2	13.5	6.76	5.881	0.00283
SNP Call x VT	6	6.9	1.15	1.003	0.42142
Residuals	2551	2933.4	1.15		
Maturation Temperature					
SNP Call	3	274.7	91.56	83.457	< 2e-16
Maturation Temperature (MT)	1	119.2	119.21	108.665	< 2e-16
SNP Call x MT	3	31.6	10.55	9.613	2.61E-06
Residuals	2555	2803	1.1		

Two-way ANOVA results indicate that for vernalisation temperature, the SNP present at this location did not affect the response of TGW to temperature, as there was no significant interaction between vernalisation temperature and SNP (Table 4.6); this can also be seen in Figure 4.8a where no significant difference was identified.

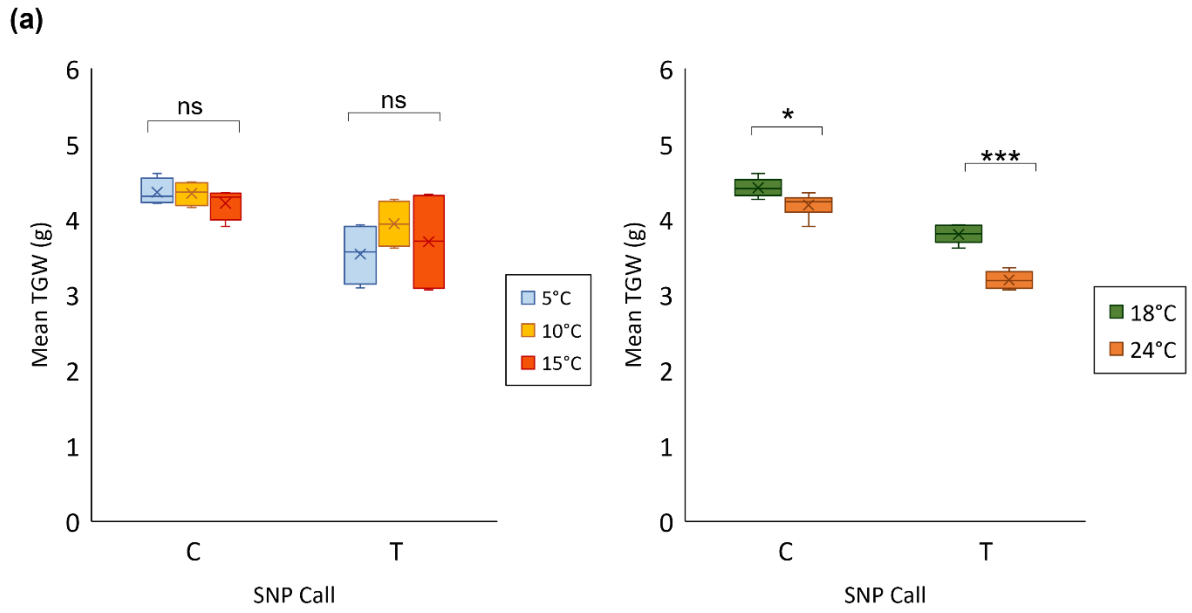

*Figure 4.8: Boxplot showing differences in TGW at two alleles at this location by vernalisation temperature. (b) Boxplot showing differences in TGW at two alleles at BnaA01g25950D by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 57 for C and n = 30 for T.*

For maturation temperature, ANOVA identified a significant interaction between SNP call and maturation temperature (Table 4.6), suggesting that this SNP had a significant effect on changes in TGW in responses to temperature. Varieties which contained a C SNP call within BnaA01g25950D had larger seeds and were less strongly affected by temperature when compared to varieties containing a T SNP call (Figure 4.8b). This suggests that varieties containing the T allele produce smaller seeds and are more sensitive to temperature.

When further analysis of this region of chromosome A01 was completed, the most significant SNP associated with TGW was found within BnaA01g25880D (Figure 4.9).

This has been identified as an orthologue of AT3G19820. The gene in which the most significant SNP was identified was 48kb from BnaA01g25940D.

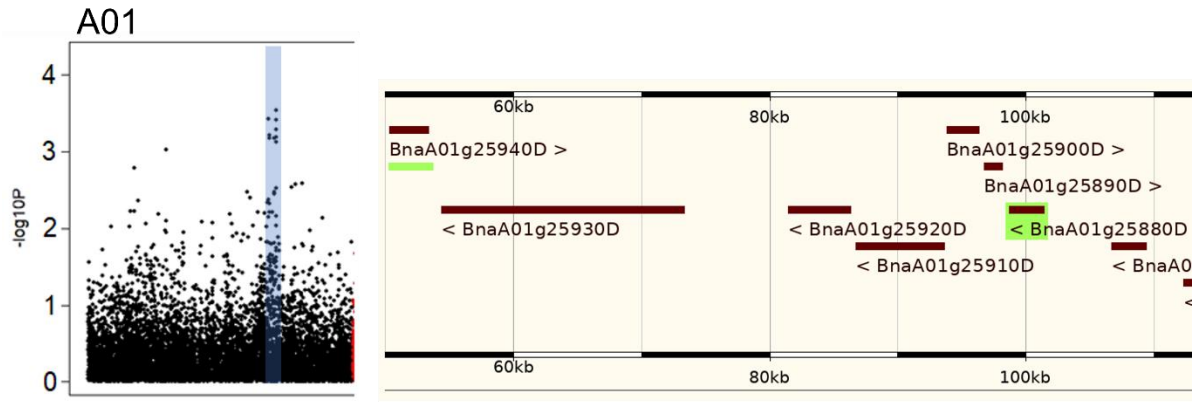

Figure 4.9: (a) Manhattan plot of chromosome A01 plants grown under 15°C 6wks vernalisation and matured at 24°C) (b) SNP marker and candidate gene locations from chromosome A01. Gene containing the SNP is highlighted in green; BnaA01.IKU2 is underlined in green.

Table 4.7: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaA01g25880D affecting TGW. Df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	P value
Vernalisation Temperature					
SNP Call	4	240.4	60.1	53.096	< 2e-16
Vernalisation Temperature (VT)	2	14.4	7.21	6.366	0.00175
SNP Call x VT	7	15.3	2.18	1.93	0.06105
Residuals	2533	2867	1.13		
Maturation Temperature					
SNP Call	4	240.4	60.1	55.73	< 2e-16
Maturation Temperature (MT)	1	119.4	119.45	110.764	< 2e-16
SNP Call x MT	4	41.4	10.36	9.606	1.04E-07
Residuals	2537	2735.9	1.08		

The SNP call used to identify this association was in the gene BnaA01g25880D. The SNP present at this location was a change between an A and a G. Using data from the phenotyping experiment, mean TGW for each treatment was compared between lines containing both SNP calls. A two-way ANOVA was also used to assess the effect of genotype and environment on TGW.

ANOVA results indicate that for vernalisation temperature, the SNP present at this location was not associated with the response of TGW to vernalisation temperature, as there was no significant interaction between vernalisation temperature and SNP (Table 4.7). As seen in Figure 4.10a, TGW was not significantly different across vernalisation temperatures.

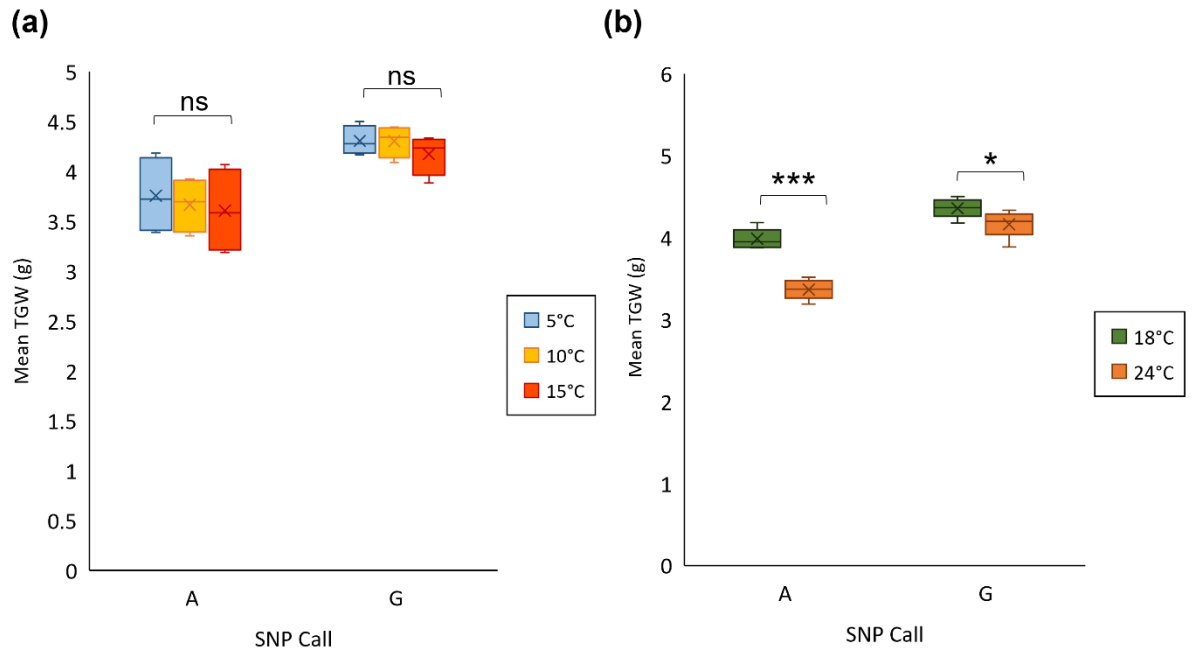

Figure 4.10: (a) Boxplot showing differences in TGW in varieties containing two alleles of BnaA01g25880D *by vernalisation temperature. (b) Boxplot showing differences in TGW in varieties containing two alleles of* BnaA01g25880D *by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 27 for A and n = 55 for G.*

As seen in Figure 4.10b, comparing TGW between lines with different SNP calls indicates that the difference between maturation temperatures for the A allele is 0.62g, while for the G allele the difference between maturation temperatures is 0.19g. Two-way ANOVA results (Table 4.7) also indicate a significant interaction between this SNP call and maturation temperature. This suggests that varieties with the A allele are more strongly affected by increased maturation temperature. Additionally, TGW is higher in varieties containing the G allele compared to A. Alongside the difference in TGW between alleles, *IKU2* was considered a possible candidate gene due to its known role in seed size control and expression in the endosperm (Luo et al, 2005).

In summary, two different SNPs were identified in the region containing *BnaA01.IKU2*. Varieties which contained these SNPs were compared to see if these SNPs were in linkage disequilibrium. As shown in Table 4.8, varieties which contained a G SNP call in BnaA01g25880D most commonly contained a C SNP call in BnaA01g25950D while varieties which contained an A SNP call at BnaA01g25880D also most commonly contained a T SNP call in BnaA01g25950D. These two combinations were present in 72 out of the 94 varieties for which SNP information was available. The separate analysis of these SNPs suggested that both the G and C SNP calls were associated with varieties with larger seeds, where TGW was less affected by temperature (Figures 4.8 and 4.10). Although not perfectly correlated, this could indicate that both of these SNP calls are linked to a region which impacts TGW, with both the G and C SNP calls linked to an allele associated with larger seed size which is less impacted by seed maturation temperature. As the *BnaA01.IKU2* gene is found between these two genes this supported further analysis into the role of *IKU2* in controlling seed size in response to temperature.

Table 4.8: Number of varieties which contain particular SNP calls in two genes on A01 associated with TGW. R, Y and N denote ambiguous SNP calls.

BnaA01g25880D SNP Call	**BnaA01g25950D SNP Call**	**No. of Varieties**

G	C	52
A	T	20
A	C	4
G	T	3
R/N	C	1
R/N	T	7
G	Y/N	1
A	Y/N	3
R	Y	3

4.2.3.2 Association with TGW on Chromosome A03 identified in the gene BnaA03g26160D

For the association identified at 14.8Mb in Chromosome A03, the most significant SNP was found within BnaA03g26160D (Figure 4.11). This has been identified as an orthologue of Arabidopsis gene AT4G02480. This gene is 34kb from BnaA03g26120D (*AXR6 / CULLIN1*).

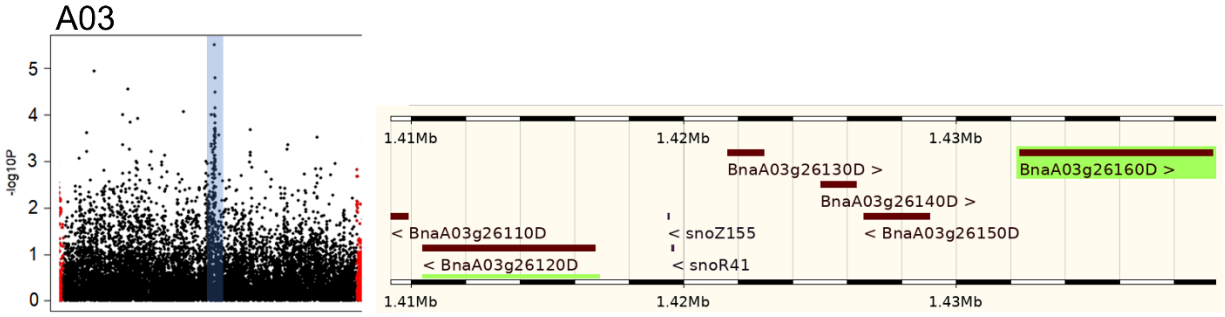

Figure 4.11: AXR6 (a) Manhattan plot of chromosome A03 (18v24 10°C 6wks). (b) SNP marker and candidate gene locations from chromosome A03. Gene containing the SNP is highlighted in green; BnaA03.AXR6 is underlined in green.

The SNP call used to identify this association was in the gene BnaA03g261260D. The SNP present at this location was a change between a C and a T. Using data from the original phenotyping experiment, mean TGW for each treatment was compared between

varieties containing both SNPs. A two-way ANOVA was also used to assess the effect of gene and environment on TGW.

There were no differences in TGW across vernalisation temperature for either allele; however, varieties with the T allele had smaller seeds across all three vernalisation temperatures (Figure 4.12a). The effect of vernalisation temperature on TGW was found to be significant (Table 4.9). No significant interaction was found between vernalisation temperature and the SNP present (Table 4.9), indicating that the response of TGW to vernalisation temperature was not affected by the SNP.

Table 4.9: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaA03g261260D affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	**Sum Squares**	**Mean Squares**	**F value**	**P value**
Vernalisation Temperature					
SNP Call	3	344.4	114.8	102.038	< 2e-16
Vernalisation Temperature (VT)	2	10.6	5.3	4.714	0.00905
SNP Call x VT	6	3.6	0.6	0.533	0.78365
Residuals	2551	2869.9	1.13		
Maturation Temperature					
SNP Call	3	344.4	114.8	106.453	< 2e-16
Maturation Temperature (MT)	1	109.8	109.76	101.781	< 2e-16
SNP Call x MT	3	19.2	6.39	5.924	0.000504
Residuals	2555	2755.2	1.08		

As seen in Figure 4.12b, comparing TGW between varieties with different SNP calls indicates that the difference between maturation temperatures for the C allele is 0.50g, while for the T allele there is no significant difference between the means for each treatment. This suggests that TGW was affected by temperature in varieties containing the C allele, while TGW in plants containing the T allele was not affected by seed maturation temperature. The effect of this SNP on TGW response to maturation temperature was found to be significant by two-way ANOVA (Table 4.7). TGW was lower in all treatments in plants containing the T allele compared to the C allele, suggesting that the C allele may be associated with increased seed size.

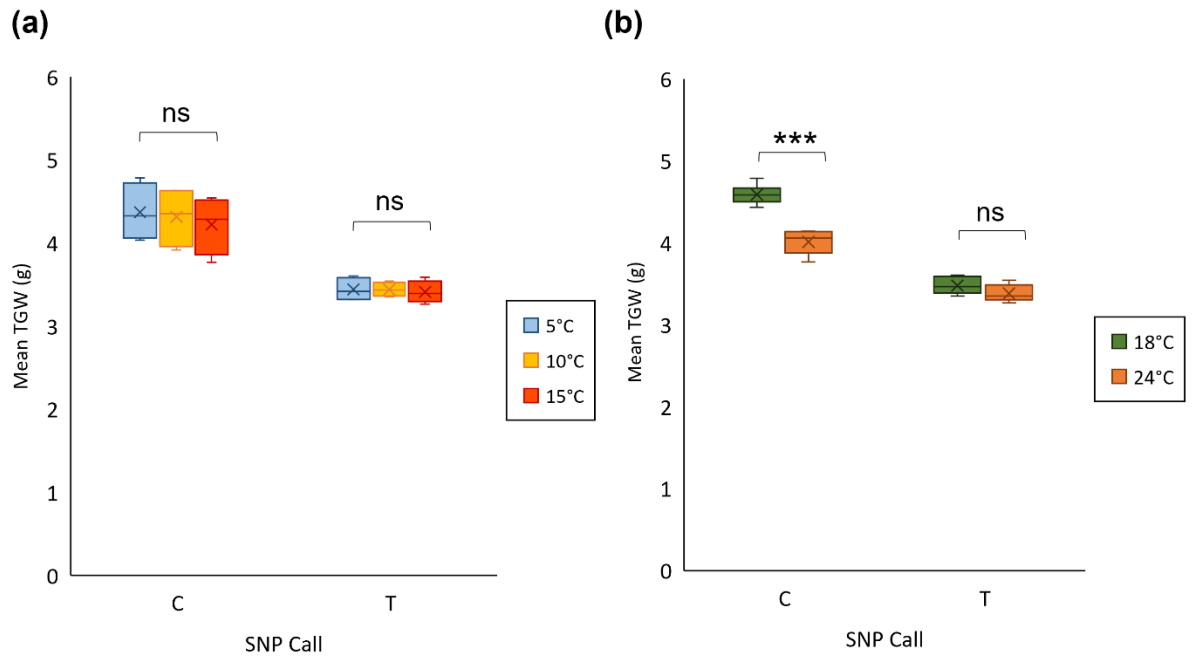

*Figure 4.12: (a) Boxplot showing differences in TGW for varieties with two alleles of BnaA03g261260D by vernalisation temperature. (b) Boxplot showing differences in for varieties with two alleles of BnaA03g261260D by maturation temperature. Asterisks used to indicate significant, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 49 for C and n = 27 for T.*

AXR6 (*AUXIN RESISTANT 6*, also known as *CULLIN1*) is known to be expressed in the embryo in *Arabidopsis* (Shen et al, 2002), supporting that it may have a role in temperature responses during seed development. *AXR6* has been identified as having an

involvement in auxin responses throughout development, including in the embryo (Hobbie et al, 2000) and some mutations in *AXR6* have been identified as temperature sensitive (Quint et al, 2005).

4.2.3.3 Association with TGW on Chromosome A03 identified in the gene BnaA03g28950D

For the association identified at 16.2Mb in Chromosome A03 (Figures 4.13a,b), the specific SNP was found within BnaA03g28950D. This has been identified as an orthologue of AT3G05100.

The *Brassica napus* genome in this region shows close synteny with a region of Arabidopsis chromosome 3 which also contains the gibberellin receptor *GID1A*. Homologs of genes found either side of *GID1A* (AT35G5120) in *Arabidopsis* (Figure 4.13c), namely AT35G5100 and AT35G5140 (*RBK2),* are found in this area of the *Brassica napus* genome (Figure 4.13d). Although *GID1A* is not known to be present in this location in the *Brassica napus* Darmor reference genome, the C05 chromosome does contain an annotated copy of *GID1A* (BnaC05g46680D). As such, it is possible that some *Brassica napus* varieties also contain a *GID1A* gene on chromosome A05 in the region implicated by the GWAS. This suggests that, although unconfirmed, a homolog of *GID1A* may be present at this location in the genome in *Brassica napus* and may be behind this association.

Table 4.10: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaA03g28950D affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	**Sum Squares**	**Mean Squares**	**F value**	**P value**
Vernalisation Temperature					
SNP Call	3	84.1	28.018	22.913	1.24E-

					14
Vernalisation Temperature (VT)	2	13.8	6.878	5.624	0.00365
SNP Call x VT	6	11.3	1.884	1.541	0.16083
Residuals	2551	3119.4	1.223		
Maturation Temperature					
SNP Call	3	84.1	28.02	23.9	2.98E-15
Maturation Temperature (MT)	1	120.2	120.22	102.552	< 2e-16
SNP Call x MT	3	29	9.66	8.238	1.87E-05
Residuals	2555	2995.3	1.17		

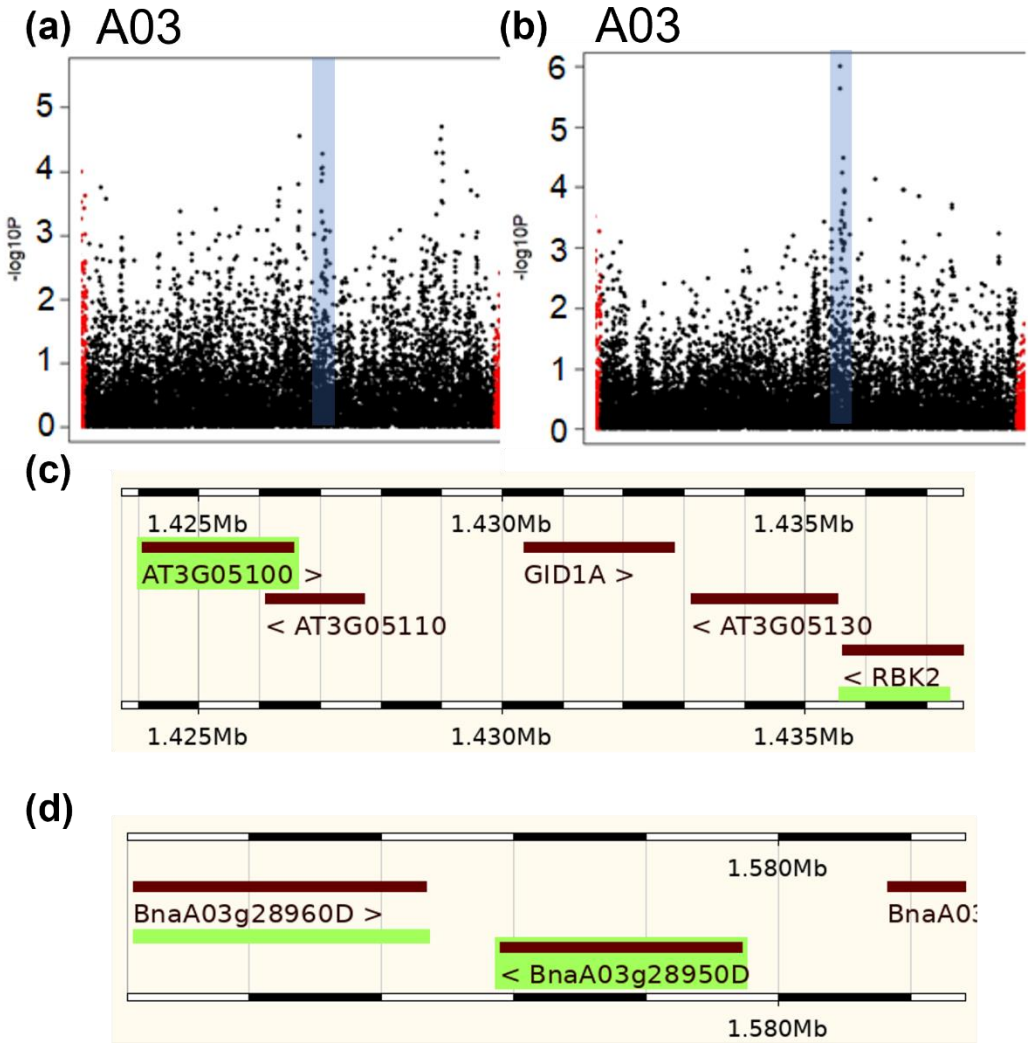

Figure 4.13: Manhattan plots showing association on chromosome A03, produced using GLM from (a) 5°C 12wks vernalisation and 18°C maturation treatment (b) preliminary analysis of the difference in TGW between 18°C and 24°C maturation temperatures after 5°C 12wks vernalisation treatment. Indication of gene locations of GID1A and SNP markers in (c) Arabidopsis (d) Brassica napus, from Ensembl plants. BnaA03g28960D (SNP containing gene and Arabidopsis orthologue are highlighted in green) is a homolog of AT3G05100. BnaA03g28950D is a homolog of RBK2 (gene and orthologue underlined in green).

The SNP call used to identify this association was in the gene BnaA03g28950D. The SNP present at this location was a change between an A and a G. Using data from the phenotyping experiment, mean TGW for each treatment was compared between lines containing the identified SNP against the SNP present in the majority of varieties. A two-way ANOVA was also used to assess the effect of gene and environment interaction associated with the SNP in BnaA03g28950D.

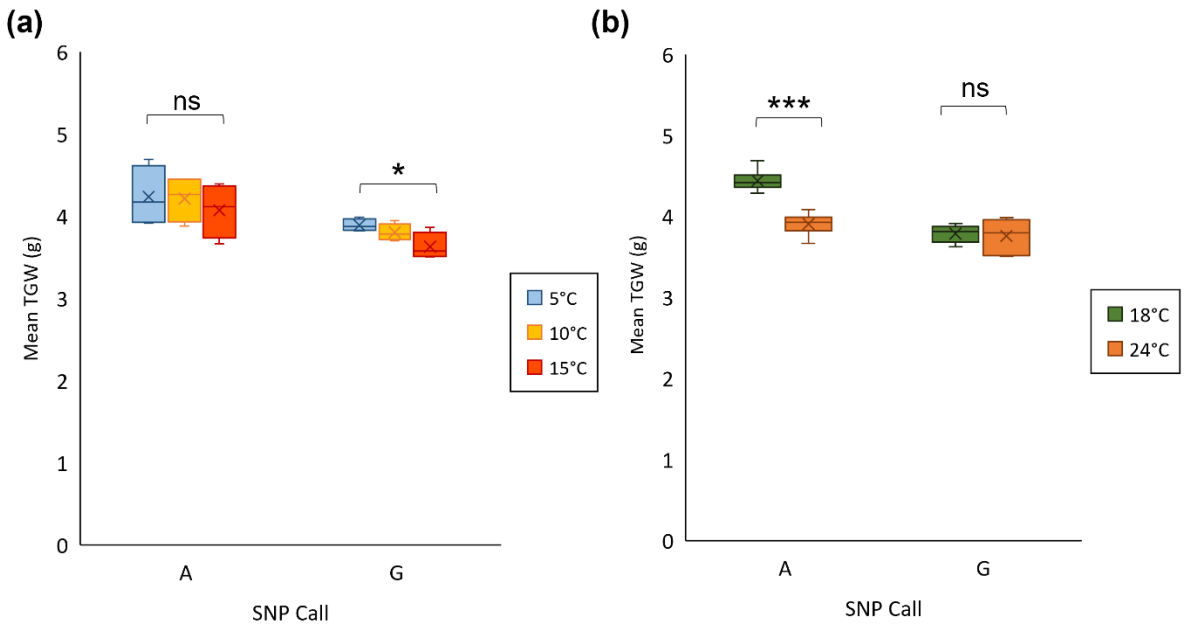

*Figure 4.14: (a) Boxplot showing differences in TGW for varieties containing two alleles of BnaA03g28950D by vernalisation temperature. (b) Boxplot showing differences in TGW for varieties containing two alleles of BnaA03g28950D by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 53 for A and n = 25 for G.*

Varieties which contained the A SNP showed no difference across the vernalisation temperatures; in varieties containing the G allele, TGW was 0.26g lower at 15°C compared to 5°C (Figure 4.14a). However, two-way ANOVA showed no significant interaction between SNP call and vernalisation temperature (Table 4.10), indicating that the difference in response is not due to the presence of this SNP. As seen in Figure 4.14b, comparing TGW between lines with different SNP calls indicates that the difference between maturation temperatures for the A allele is 0.53g, while for the G allele there is no significant difference between the temperatures. Two-way ANOVA showed a significant interaction between SNP and maturation temperature (Table 4.10), suggesting that the SNP present in this location had an impact on the TGW response to maturation temperature. This suggests that varieties containing the A allele are sensitive to seed maturation temperature, while varieties with the G allele are not. Additionally, at 18°C TGW is higher in varieties containing the A allele compared to G. GID1A encodes a gibberellin receptor; gibberellin is important for development of the integuments in the

seed (Kim et al, 2005). The degree of integument growth determines the size of the final cavity which the embryo can grow to fill and is therefore an important determinant of seed size (Garcia et al, 2005). *GID1A* is expressed in the *Arabidopsis* seed (Griffiths et al, 2006; Voegele et al, 2011; Gallego-Giraldo, 2014), suggesting this process could be involved in seed size and temperature responses. However, the presence of an orthologue of *GID1A* at this location in *Brassica napus* could not be confirmed.

4.2.3.4. Association with TGW on Chromosome A08 identified in the gene BnaA08g17820D

For the association identified at 17.8Mb in Chromosome A08, the specific SNP was found within BnaA08g17820D (Figure 4.15). This has been identified as an orthologue of AT1G29670. This SNP was located in a *Brassica napus* orthologue of the Arabidopsis GDSL-motif containing lipase gene *GDSL1*.

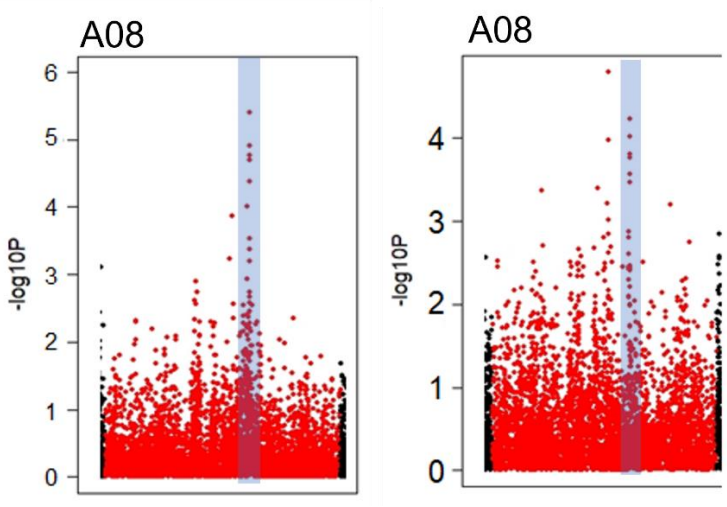

Figure 4.15: Manhattan plot of chromosome A08 showing associations for TGW from plants vernalised at 15°C for 12 weeks, comparing between 18°C and 24°C maturation temperature. Left panel shows preliminary analysis while right panel shows final analysis.

Using data from the original phenotyping experiment, mean TGW for each treatment was compared between varieties containing the differences at this SNP location. Two-way ANOVA was also used to identify interactions between the SNP call and the environment.

The SNP present at this location was a change between an A and a G. There were no differences in TGW between varieties containing either of these alleles at any of the vernalisation temperatures studied (Figure 4.16a). Two-way ANOVA showed no significant interaction between vernalisation temperature and SNP (Table 4.11), indicating that this SNP did not affect the response of TGW to vernalisation temperature.

Table 4.11: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaA08g17820D affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	**Sum Squares**	**Mean Squares**	**F value**	**P value**
Vernalisation Temperature					
SNP Call	2	15	7.395	5.929	0.002697
Vernalisation Temperature (VT)	2	18	8.84	7.088	0.000852
SNP Call x VT	4	11	2.675	2.145	0.07288
Residuals	2554	3185	1.247		
Maturation Temperature					
SNP Call	2	14.8	7.4	6.182	0.0021
Maturation Temperature (MT)	1	131.9	131.86	110.224	< 2e-16
SNP Call x MT	2	23.1	11.53	9.634	6.78E-05
Residuals	2557	3058.8	1.2		

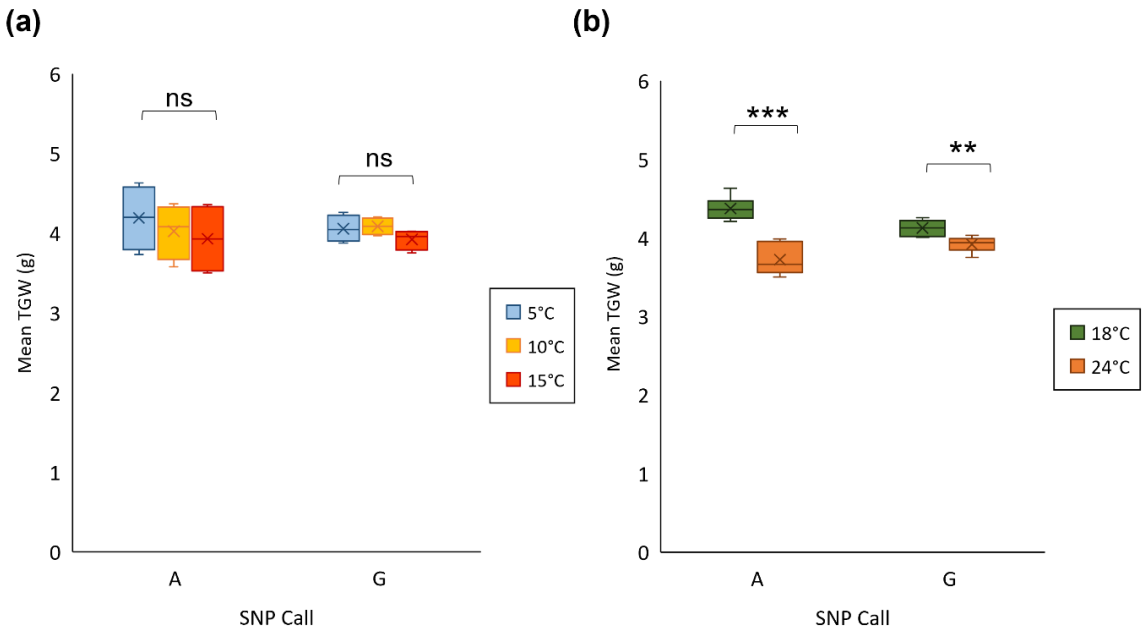

*Figure 4.16: (a) Boxplot showing differences in TGW for varieties containing two alleles of BnaA08g17820D by vernalisation temperature. (b) Boxplot showing differences in TGW for varieties containing two alleles of BnaA08g17820D by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 33 for A and n = 59 for G.*

As seen in Figure 4.16b, comparing TGW between lines with different SNP calls indicates that the difference between maturation temperatures for the A allele is 0.65g, while for the G allele there is no significant difference between means at each temperature treatment. The interaction between maturation temperature and SNP call was found by two-way ANOVA to be significant (Table 4.11). This suggests that TGW in plants containing the A allele of *GDSL1* is temperature sensitive, while TGW in plants containing the G allele is not affected by seed maturation temperature. As the SNP call from the GWAS was directly located on the gene *GDSL1*, it was considered a likely candidate for the gene responsible for the association due to the role of GDSL family genes in lipid metabolism in seeds (Dong et al, 2016). *GDSL1* has also been found to be expressed in the developing seed (Ding et al, 2019a).

4.2.3.5 Association with TGW on Chromosome A06 identified in the gene BnaA06g31400D

For the association identified at 26.5Mb in Chromosome A06, the specific SNP was found within BnaA06g31400D (Figure 4.17). This has been identified as an orthologue of AT2G02560. This SNP was located within a *Brassica napus* orthologue of the Arabidopsis gene *CAND1*.

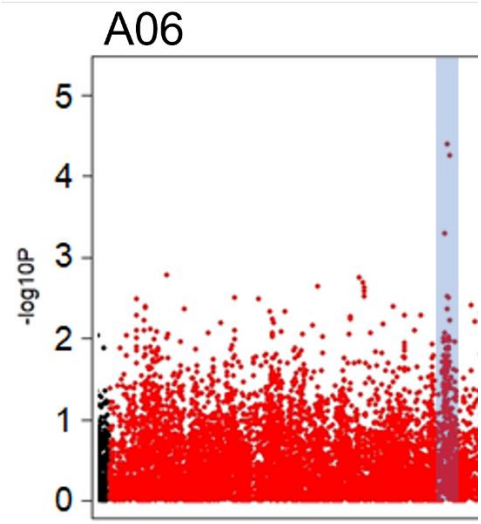

Figure 4.17: Manhattan plot showing associations with TGW on Chromosome A06 from the temperature treatments 15°C 12wk vernalisation and comparison between maturation temperature 18°C against 24°C.

The SNP present at this location within the *Brassica napus* CAND1 orthologue was a change between a T and a G. Using data from the original phenotyping experiment, mean TGW for each treatment was compared between varieties containing these two SNP calls. Two-way ANOVA was also used to assess the effect of the interaction between genotype and environment.

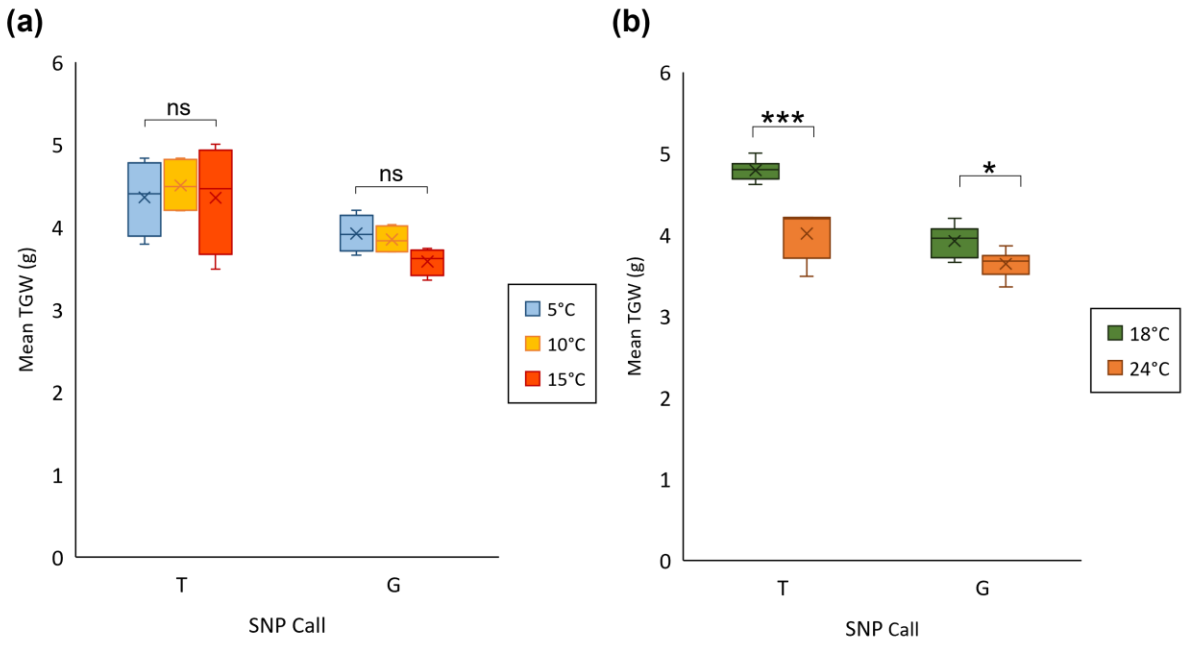

*Figure 4.18: (a) Boxplot showing differences in TGW for varieties containing two alleles of BnaA06g31400D by vernalisation temperature. (b) Boxplot showing differences in for varieties containing two alleles of BnaA06g31400D by maturation temperature. asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 9 for T and n = 32 for G.*

As seen in Figure 4.18a, TGW was not significantly different across vernalisation temperatures for either allele. Table 4.12 also shows no significant interaction between vernalisation temperature and SNP call. The SNP present in this gene does not affect how TGW changes in response to vernalisation temperature.

As seen in Figure 4.18b, comparing TGW between lines with different SNP calls indicates that the difference between maturation temperatures for the T allele is 0.77g, while for the G allele the difference between maturation temperatures is 0.28g. Two-way ANOVA results show that there is a significant interaction between maturation temperature and SNP call. This suggests that in varieties with the T allele, TGW is more strongly affected by increasing seed maturation temperature than varieties with the G allele. Maturation temperature and SNP call both show significant effects on TGW individually (Table 4.12).

Table 4.12: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaA06g31400D affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	P value
Vernalisation Temperature					
SNP Call	3	38	12.804	10.319	9.47E-07
Vernalisation Temperature (VT)	2	17	8.26	6.657	0.00131
SNP Call x VT	6	8	1.405	1.132	0.34067
Residuals	2551	3165	1.241		
Maturation Temperature					
SNP Call	3	38.4	12.8	10.745	5.13E-07
Maturation Temperature (MT)	1	125	125.04	104.93	< 2e-16
SNP Call x MT	3	20.5	6.85	5.748	0.000646
Residuals	2555	3044.5	1.19		

CAND1 is expressed throughout the plant, including in seeds (Cheng et al, 2004) and so is a possible candidate for direct functions impacting seed size. Additionally, CAND1 interacts with AXR6, described above (Zheng et al, 2002).

4.2.3.6 Association with TGW on Chromosome C02 identified in an orthologue of AT5G60710

For the association identified at 6.9Mb in Chromosome C02, the specific SNP was found within the gene model Bo2g024660. This SNP was found within an orthologue of AT5G60710 (*REVOLUTA*) (Figure 4.19), although the specific gene name in *Brassica napus* could not be identified. The list of nearby genes (found in Table B6, in Appendix B) contained another gene of possible interest, an orthologue of the LEA family gene AT5G22870. However, little information was available about the gene, which was 115kb

away from the SNP. As the SNP was identified within a *Brassica napus* orthologue of *REVOLUTA* (*REV*), this gene was explored further.

Using data from the original phenotyping experiment detailed in Chapter 2, mean TGW for each treatment was compared between lines containing the two major SNP variants. Two-way ANOVA was also used to identify interactions between genotype and environment.

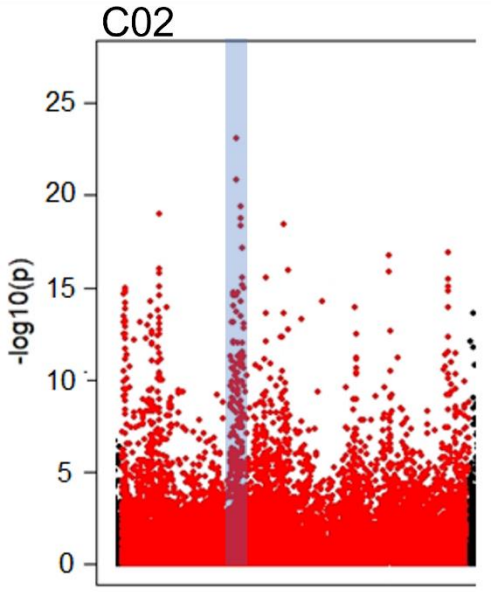

Figure 4.19: Manhattan plot showing associations with TGW on Chromosome C02, identified using meta-analysis.

The SNP present at this location was a change between an A and a T. As seen in Figure 4.20a, TGW was not significantly different across vernalisation temperatures for varieties containing the T allele. Increasing the vernalisation temperature from 5°C to 15°C caused a 0.88g decrease in TGW for varieties containing the A allele. While both SNP call and vernalisation temperature were found to affect TGW individually, no significant interaction was found between genotype and vernalisation environment, suggesting that this SNP does not affect how TGW changes in response to maturation temperature.

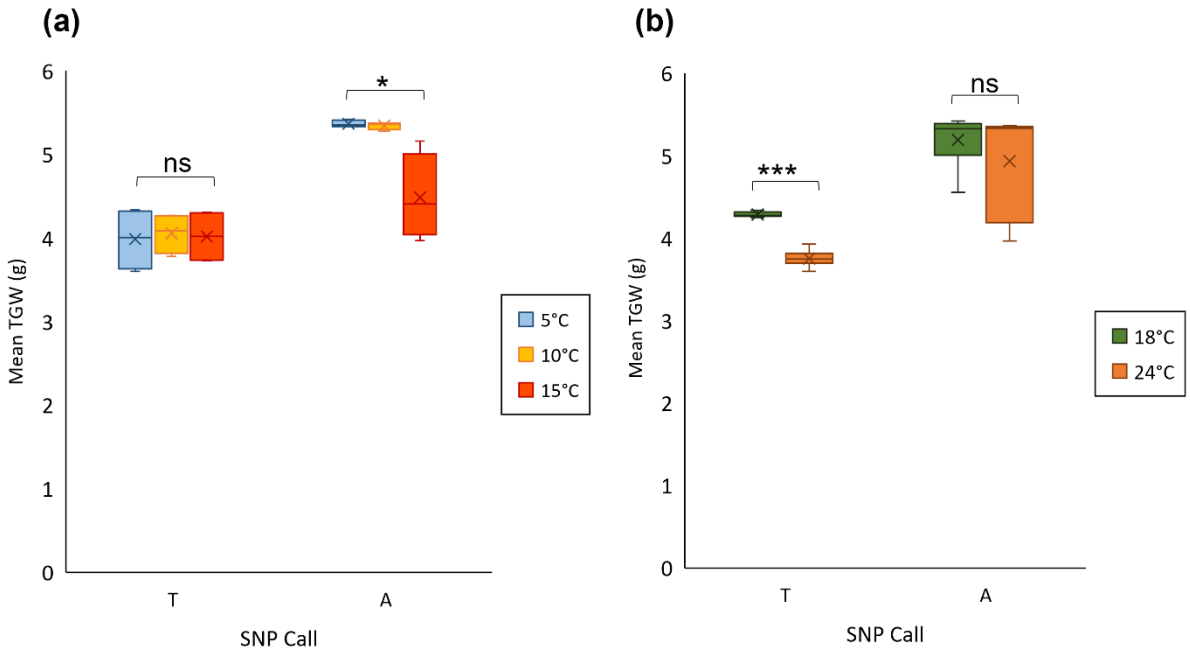

*Figure 4.20: (a) Boxplot showing differences in TGW for varieties with two alleles at Bo2g024660 by vernalisation temperature. (b) Boxplot showing differences in TGW between varieties with different SNP calls in Bo2g024660 by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 38 for T and n = 7 for A.*

As seen in Figure 4.20b, comparing TGW between lines with different SNP calls indicates that the difference between maturation temperatures for the T allele is 0.53g, while for the A allele there is no significant difference between the maturation temperatures. Two-way ANOVA showed a significant interaction between SNP call and maturation temperature (Table 4.13), indicating that the SNP affected the change in TGW in response to maturation temperature. This suggests that varieties with the T allele are sensitive to seed maturation temperature, while varieties with the G allele are not.

REV is expressed in seeds throughout seed development as well as in female reproductive tissues (Sieber et al, 2004; Le et al, 2010), which may suggest a possible

link of maternal effects on seed size. REV is involved in development of the maternal integuments (Kelley et al, 2009; Hashimoto et al, 2018) and mutants in *REV* show increased seed size in *Arabidopsis* (Talbert et al, 1995).

Table 4.13: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in Bo2g024660 on Chromosome C06 affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	P value
Vernalisation Temperature					
SNP Call	3	304.7	101.55	89.475	< 2e-16
Vernalisation Temperature (VT)	2	17.5	8.75	7.707	0.00046
SNP Call x VT	6	11	1.84	1.62	0.1375
Residuals	2551	2895.3	1.13		
Maturation Temperature					
SNP Call	3	304.7	101.55	94.137	< 2e-16
Maturation Temperature (MT)	1	135.7	135.7	125.792	< 2e-16
SNP Call x MT	3	31.9	10.63	9.849	1.86E-06
Residuals	2555	2756.3	1.08		

4.2.3.7 Association with TGW on Chromosome C06 identified in the gene BnaC06g31160D

For the association identified at 37.6Mb in Chromosome C06, the specific SNP was found within BnaC06g31160D (Figure 4.21). This has been identified as an orthologue of AT1G75820. This SNP was located within a *Brassica napus* orthologue of the Arabidopsis gene *CLV1*. *CLV1* is expressed in the seed and embryo in *Arabidopsis* (Le et al, 2010) and functions in control of cell proliferation (Fletcher et al, 1999).

Nearby genes included two members of the GDSL-like lipase/acylhydrolase superfamily (Table B7, in Appendix B). However, the SNP present within the *Brassica napus CLV1* orthologue was most significant, and due to the role of *CLV1* in cell proliferation this gene was followed up for further analysis.

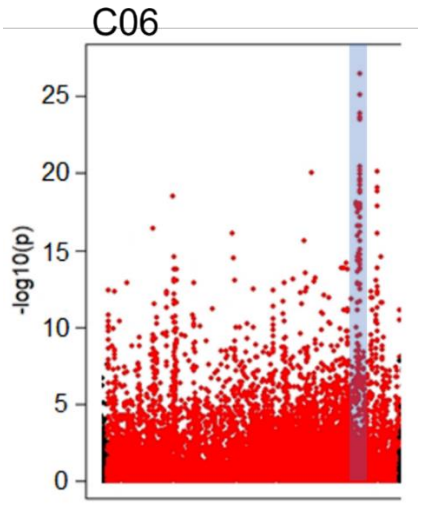

Figure 4.21: Manhattan plot indicating SNP associations associated with TGW in chromosome C06, identified using meta-analysis.

The SNP present at this location was a change between an A and a C. Using data from the original phenotyping experiment, mean TGW for each treatment was compared between varieties containing these two SNPs, and two-way ANOVA was used to explore the interaction between genotype and environment.

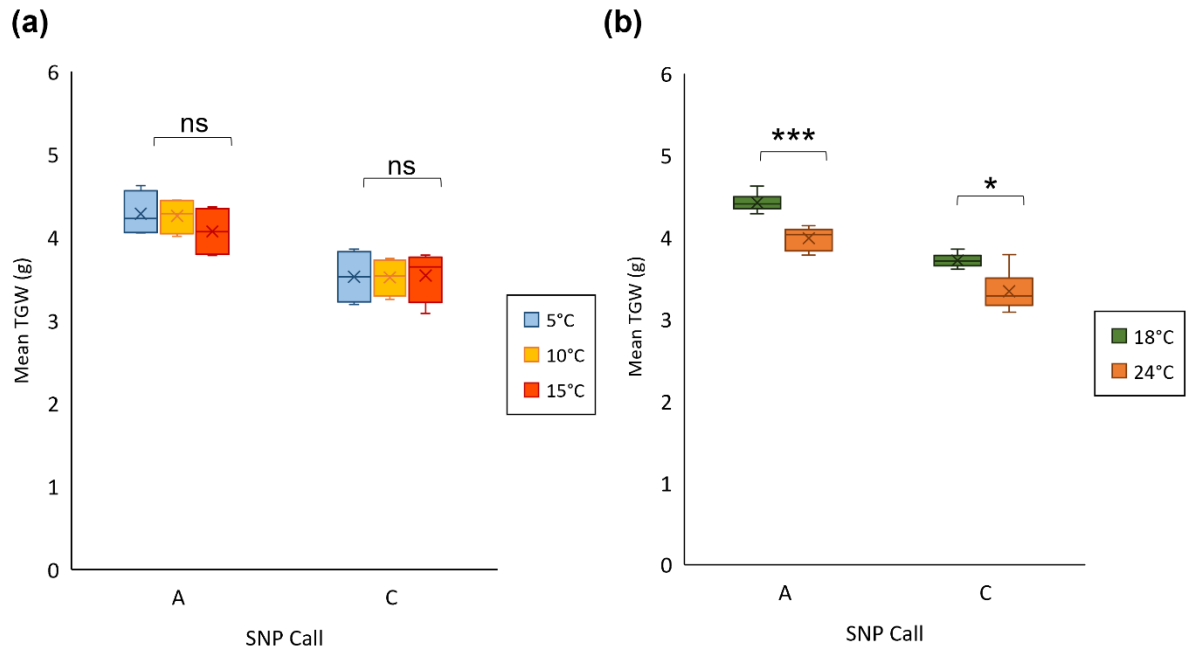

*Figure 4.22: (a) Boxplot showing differences in TGW for varieties with two different SNP calls in BnaC0631160D by vernalisation temperature. (b) Boxplot showing differences in for varieties with two different SNP calls in BnaC0631160D by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 63 for A and n = 21 for C.*

Table 4.14: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaC06g31160D affecting TGW. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	P value
Vernalisation Temperature					
SNP Call	3	229.8	76.61	65.935	< 2e-16
Vernalisation Temperature (VT)	2	26.5	13.27	11.419	1.16E-05
SNP Call x VT	6	8.3	1.39	1.192	0.307
Residuals	2551	2963.9	1.16		
Maturation Temperature					
SNP Call	3	229.8	76.61	68.346	<2e-16
Maturation Temperature (MT)	1	132.7	132.74	118.425	<2e-16
SNP Call x MT	3	2.2	0.73	0.648	0.584
Residuals	2555	2863.8	1.12		

TGW was not significantly different across the range of vernalisation temperatures studied for varieties containing either of these SNPs (Figure 4.2a). As two-way ANOVA testing did not find a significant association between vernalisation temperature and SNP call (Table 4.14), this SNP is unlikely to have any effect on TGW responses to vernalisation temperature. As seen in Figure 4.22b, comparing TGW between lines with different SNP calls indicates that the difference between maturation temperatures for the A allele is 0.43g, while for the C allele this difference is 0.37g. Two-way ANOVA did not show a significant interaction between SNP call and maturation temperature. This suggests that varieties with both alleles are affected by seed maturation temperature, without an effect of SNP on how TGW is affected. The results of the two-way ANOVA indicated a significant effect of SNP on TGW individually. At both temperatures, varieties containing the A allele have larger seeds, suggesting that the A allele may be associated with increased seed size. This could possibly be related to the cell proliferation function of *CLV1*.

4.2.3.8 Summary of regions of interest for genetic associations with TGW

The SNPs described in the above section were those which were associated with TGW and where genes could be identified which appeared likely to have roles in determining seed size. For all other associations produced from the GWAS results, no clear candidate genes were identified within 150kb of the most significant SNP.

It is possible that these regions contained genes with previously unknown roles in seed size determination, or that genes controlling seed size were not annotated in these regions. However, associations with clear candidate genes were prioritized for further study (Table 4.15).

Table 4.15: Summary of TGW SNP associations and candidate genes identified, if any.

Chromosome	Peak Location	Candidate Gene	Arabidopsis orthologue	Reference
A01	25.9 - 26.1	BnaA01g25940D	AT3G19700 (HAIKU2)	Luo et al, 2005
A01	2.6 - 2.7	N/A	N/A	
A03	8.7 - 8.8	N/A	N/A	
A03	14.8 - 14.9	BnaA03g26120D	AT4G02570 (AXR6/CULLIN1)	Shen et al, 2002
A03	16.2 - 16.4	Hypothetical *BnaA03.GID1A*	AT3G5120 (GID1A)	Griffiths et al, 2006
A04	21.5 - 21.7	N/A	N/A	
A06	26.5 - 26.8	BnaA06g31400D	AT2G02560 (CAND1)	Zheng et al, 2002
A07	24.1 - 24.2	N/A	N/A	
A08	17.8 - 18	BnaA08g17820D	AT1G29670 (GDSL1)	Ding et al, 2019
C02	6.9 - 7.0	Orthologue of Bo2g024660	AT5G60710 (REVOLUTA)	Talbert et al, 1995
C06	37.6 - 37.7	BnaC0631160D	AT1G75850 (CLAVATA1)	Fletcher et al, 1999; Leyser and Furner, 1992
C06	36.8 - 37.7	N/A	N/A	
C08	25.3 - 25.9	N/A	N/A	

4.2.4 Identification of associations of interest for seed number per pod

As shown in the example in section 4.2.1, associations identified from GWAS results were narrowed down. As using the Bonferroni-adjusted p-value for most treatments produced no significant results, the associations were also assessed using *ad hoc* thresholds for p-values across all GWAS analyses. In this case, the significance threshold

was adjusted manually to identify genes which could plausibly be associated with differences in SPP but did not meet the Bonferroni-adjusted significance threshold. A p-value of 0.0001 equates to a -log(p) of 4, which in these Manhattan plots filters out the majority of noise. The associations identified for SPP are summarized in Table 4.16. Full details can be found in Appendix B, Table B8.11.

Table 4.16: Summary of GWAS association locations for SPP, with region located indicated by Mb on the chromosome as described in the pangenome (He et al, 2015). Totals columns show the number of treatments which produced SNPs in each region below the significance threshold indicated in the column header.

Chr	Location (Mb)	Total < 0.001	Total < 0.0001	Total < 2.81E-07
A02	34.6 - 34.7	5	2	0
A05	1.5 - 1.6	3	1	0
A08	15.4 - 16	11	6	0
C01	38.6 - 39.1	11	6	0
C04	35.1 - 35.2	2	0	0
C07	43.6 - 43.7	2	0	0
C06	16.7 - 16.9	4	0	0
C09	2.7 - 2.9	10	5	0

Associations for SPP were identified in the same manner and summarised in Table 16. For SPP, three associations were identified on the A genome and five on the C genome. None of these manually identified associations passed the Bonferroni threshold for significance. Additional lower cut-off points using *ad hoc* p-values were used to identify which associations were likely to be found across a range of treatments, comprised of data from the 12 temperature treatments used in the glasshouse phenotyping experiment as well as an additional 10 datasets created by comparing the trait values between contrasting temperatures.

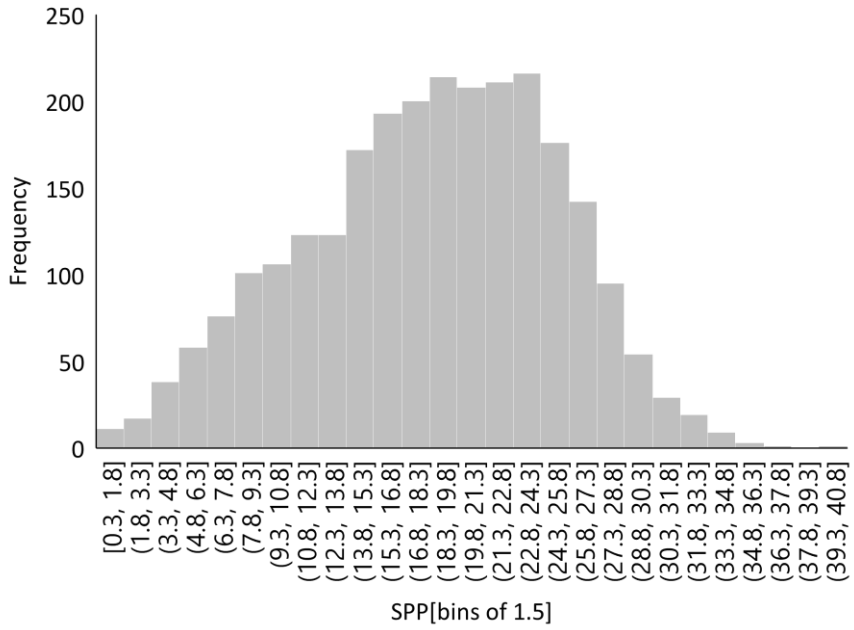

Figure 4.23: Distribution of SPP across all 12 treatments used in meta-analysis.

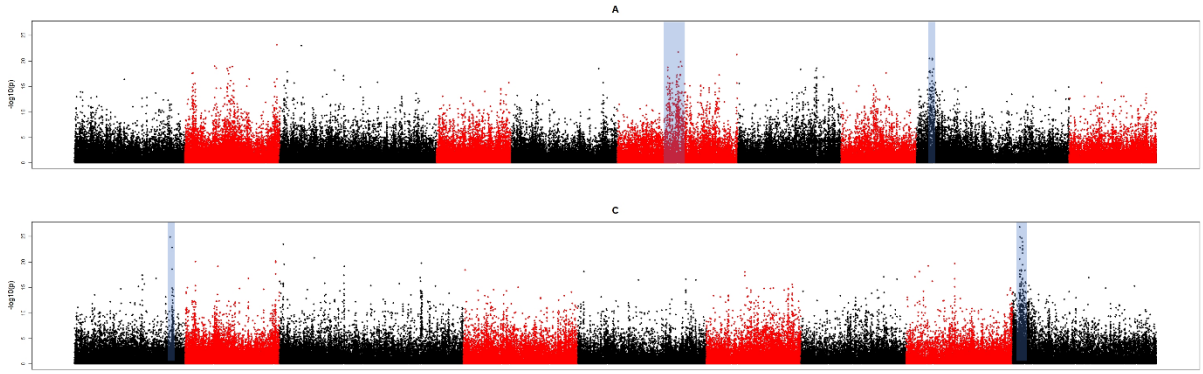

Figure 4.24: Manhattan plot showing associations with SPP from METAL.

The distribution of SPP across all 12 temperature treatments combined in this meta-analysis can be seen in Figure 4.23. Four main associations were identified on chromosomes A06, A09, C01 and C09 for SPP using the meta-analysis method (Figure 4.24).

The details of the associations identified from the meta-analysis for SPP are described in Table 4.17. These associations were then considered further alongside those identified

using GWAS analyses for individual temperature treatments. The aim was to identify possible candidate genes which could be causing them.

Table 4.17: SNP associations found using METAL meta-analysis across 12 treatments for SPP. Peak location range is in Mb from beginning of the indicated chromosome. p-values are the most significant in the given range.

Chromosome	Peak Location Range (Mb)	p-value
A06	15.7 - 17.9	2.18E-22
A09	2.0 - 2.5	3.48E-21
C01	37.8 - 38.8	1.33E-25
C09	1.9 - 2.1	1.56E-27

4.2.4.1 Association with SPP on Chromosome C09 in BnaC09g02780D

For the association identified at 2.8Mb in Chromosome C09, the specific SNP was found within BnaC09g02780D, a gene which has been identified as an orthologue of AT5G58750 (Figure 4.25). This gene is 8kb from BnaC09g02740D, an orthologue of *AN3/GIF1* (full details of nearby genes within 150kb of the SNP in Table B8, Appendix B). Mutations in *AN3* have been previously shown to have an impact on seed size as well as seed number and the gene is expressed in seeds and reproductive tissues in *Arabidopsis* (Meng et al, 2016).

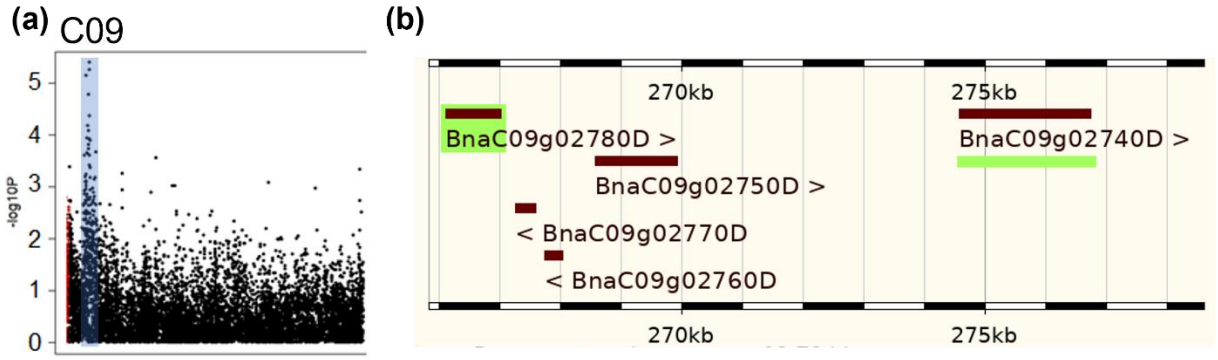

Figure 4.25: (a) Manhattan plot of chromosome C09 (b) SNP marker and candidate gene locations from chromosome C09 from GWAS analysis of vernalisation at 10°C for 6 weeks, comparing maturation temperature

between 18°C and 24°C. Gene in which SNP was identified is highlighted in green; the location of BnaC09.AN3 is underlined in green.

The SNP call used to identify this location was in the gene BnaC09g02780D, with the difference associated with TGW being a change between an A and a G. Using data from the original phenotyping experiment detailed in Chapter 2, mean SPP for each treatment was compared between lines containing these two SNPs. Two-way ANOVA was used to assess the interaction between genotype and environment.

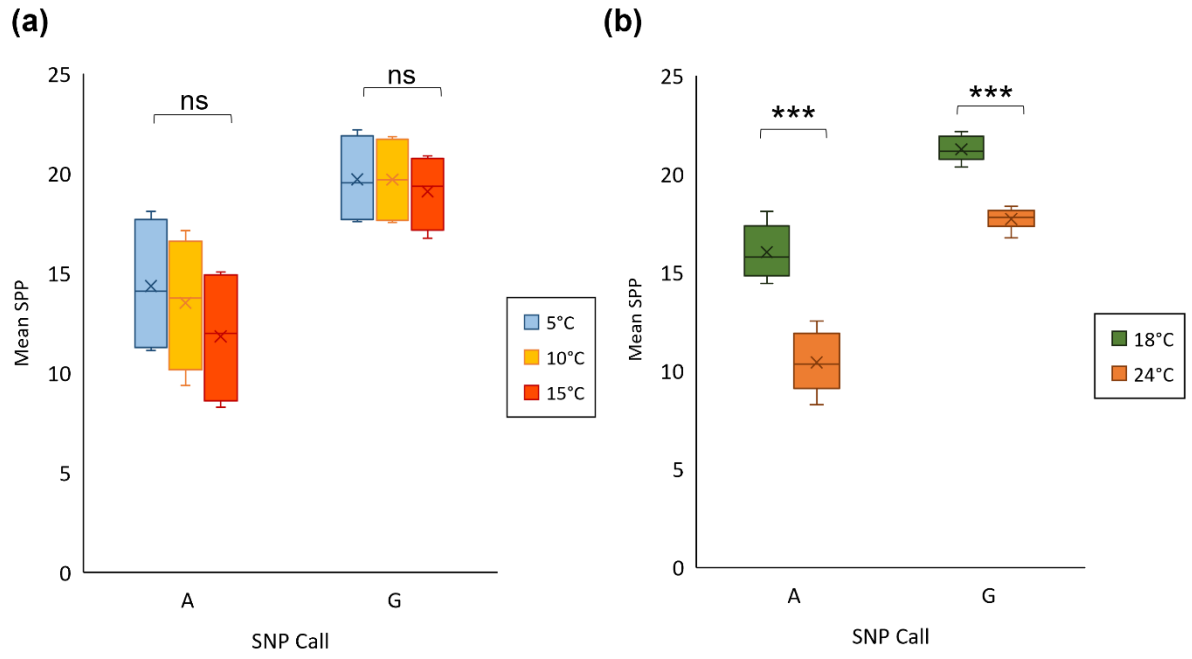

*Figure 4.26: (a) Boxplot showing differences in TGW for varieties containing two alleles of BnaC09g02780D by vernalisation temperature. (b) Boxplot showing differences in TGW in varieties containing two of BnaC09g02780D by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n = 12 for A ; n = 49 for G.*

SPP was not significantly different across the vernalisation temperatures studied in varieties containing either allele (Figure 4.26a). Results from two-way ANOVA (Table 4.18) indicate no significant interaction between SNP call and vernalisation temperature, suggesting the SNP present at this location does not affect changes in SPP resulting from vernalisation temperature responses. As seen in Figure 4.26b, comparing

SPP between lines with different SNP calls indicates that the difference in mean SPP between maturation temperatures for the A allele is 5.6, while for the G allele this difference was 3.5. Two-way ANOVA identified a significant interaction between SNP call and the maturation temperature that plants were exposed to post-vernalisation (Table 4.18). This suggests that plants containing the G allele were less sensitive to temperature. Additionally, plants with the G allele have a higher number of SPP produced at all temperatures compared to the A allele, suggesting there may be a link between the G allele and increased seed production.

Table 4.18: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaC09g02810D affecting SPP. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	P value
Vernalisation Temperature					
SNP Call	3	10432	3477	82.623	<2e-16
Vernalisation Temperature (VT)	2	229	114	2.717	0.0663
SNP Call x VT	6	354	59	1.401	0.2103
Residuals	2551	107364	42		
Maturation Temperature					
SNP Call	3	10432	3477	88.499	<2e-16
Maturation Temperature (MT)	1	7213	7213	183.567	<2e-16
SNP Call x MT	3	341	114	2.892	0.0341
Residuals	2555	100392	39		

4.2.4.2 Association with SPP on Chromosome C01 identified in the gene BnaC01g36120D

For the association identified at 38.7Mb in Chromosome C01, the specific SNP, a change between an A and a G, was found within BnaC01g36120D (Figure 4.27a), identified using the gene marker Bo1g134770. This has been identified as an orthologue of AT3G13920. This SNP is 41kb from the gene marker Bo1g134790 (Figure 4.27b), which corresponds to an orthologue of *MYB26* (BnaC01g37360D).

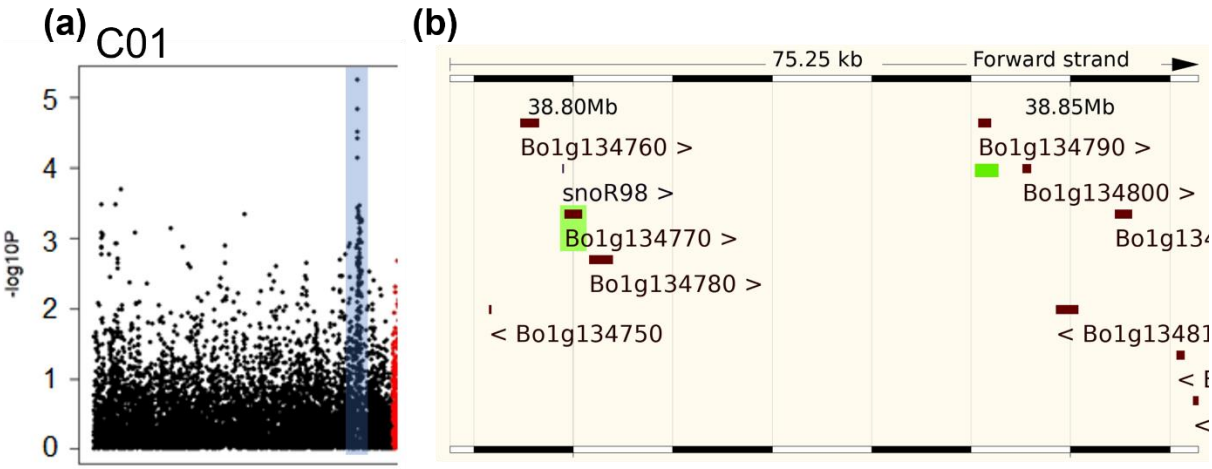

Figure 4.27: (a) Manhattan plot of chromosome C01 showing associations with SPP (b) SNP marker locations from chromosome C01. The gene in which the SNP was identified in highlighted in green, while the gene marker corresponding to BnaC01.MYB26 is underlined in green.

The gene *MYB26* was identified as a candidate gene on chromosome C01 affecting seed number per pod. In *Arabidopsis thaliana,* the *MYB26* gene is also known as *MALE STERILE 35 (MS35)* and plays a vital role in anther development (Yang et al, 2007). Mutants in this gene are male-sterile as the anthers do not dehisce and so do not produce seeds (Dawson et al, 1999; Steiner-Lange et al, 2003). In *Brassica rapa*, reduced expression levels of *MYB26* have been shown to cause male sterility (Liu et al, 2016).

A two-way ANOVA was used to examine the effects of this SNP on seed number per pod (Table 4.19). The SNP call itself has a significant effect on SPP. However, vernalisation temperature was not found by two-way ANOVA to have a significant effect on SPP ($p = 0.292$). While no significant difference was found between vernalisation temperatures, varieties containing the A allele showed a larger number of seed per pod across all vernalisation temperatures (Figure 4.28a). There was no significant interaction between

SNP call and vernalisation temperature (Table 4.19), suggesting that vernalisation temperature does not affect SPP in any varieties regardless of what SNP is present.

These results show that maturation temperature did have a significant effect on SPP, and a significant interaction was identified between SNP call and maturation temperature ($p = 0.000737$). Varieties containing the A allele tended to have higher numbers of seeds per pod, with seed production not affected by maturation temperature; this is in contrast to varieties containing the G allele, which produced fewer seeds per pod at 24°C compared to 18°C (Figure 4.28b). This suggests that the SNP present in this location has an effect on how SPP changes in responses to maturation temperature.

Table 4.19: Two-way ANOVA results for Gene x Environment interaction. Genotype is indicated by the SNP present in BnaC01g36120D affecting SPP. The environment is either vernalisation temperature or maturation temperature. Df = degrees of freedom

	Df	**Sum Squares**	**Mean Squares**	**F value**	**P value**
Vernalisation Temperature					
SNP Call	2	8293	4146	96.503	<2e-16
Vernalisation Temperature (VT)	2	106	53	1.232	0.292
SNP Call x VT	4	247	62	1.436	0.219
Residuals	2554	109733	43		
Maturation Temperature					
SNP Call	2	8293	4146	103.32	< 2e-16
Maturation Temperature (MT)	1	7077	7077	176.35	< 2e-16
SNP Call x MT	2	395	197	4.92	0.00737
Residuals	2557	102614	40		

Due to the role of *MYB26* described in previous studies, it is most likely that the effect of *MYB26* on seed number is due to effects on male fertility. Temperature is known to affect male fertility and pollen viability (Young et al, 2004; Barnabás et al, 2008). which

corresponds with the results found shown in Table 4.19. As this project was focused on maternal effects and seed development, *MYB26* was not studied further here. However, differences in *MYB26* between varieties of *Brassica napus* poses an interesting basis for further study into improving seed production.

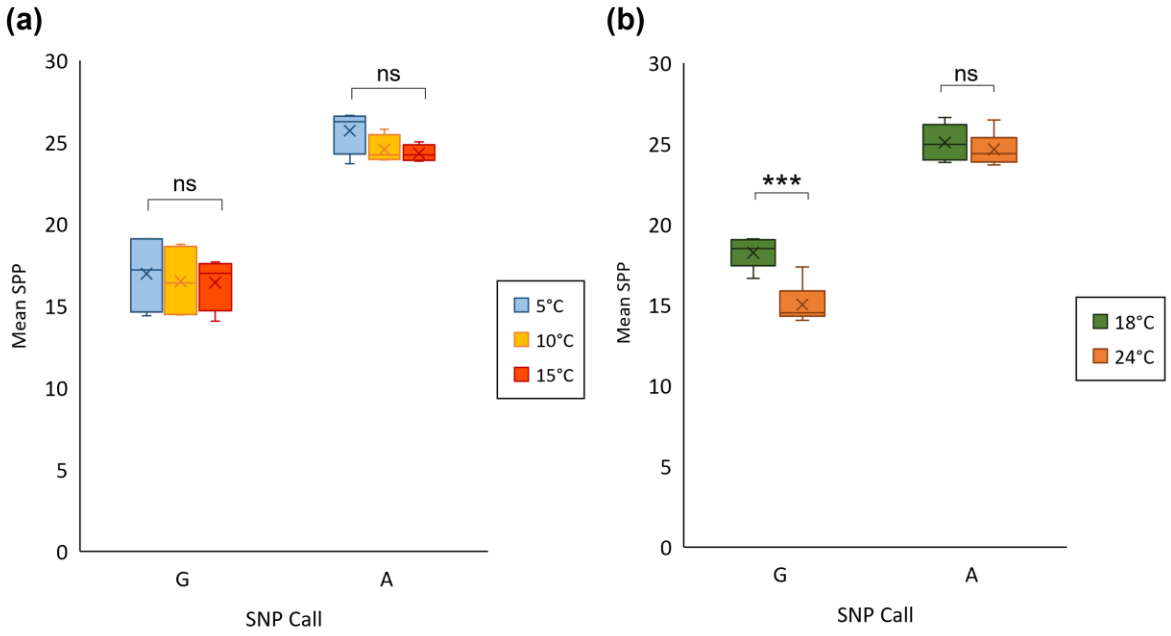

*Figure 4.28: (a) Boxplot showing differences in TGW for varieties with two alleles of BnaC01g36120D by vernalisation temperature. (b) Boxplot showing differences in TGW for varieties with two alleles of BnaC01g36120D by maturation temperature. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. For varieties per allele, n =5 for A and n = 41 for G.*

4.2.4.3 Summary of SPP SNPs identified

For all other associations produced from the GWAS results, no clear genes with functions identified as likely to affect seed number per pod were identified within 150kb of the most significant SNP. It is possible that these regions contained genes with unknown roles in seed production and fertility, or that genes controlling seed number

were not annotated in these regions. Associations with clear candidate genes were prioritized for further study (Table 4.20).

Table 4.20: Summary of SPP SNP associations and candidate genes identified.

Chromosome	Peak Location	Possible Candidate Gene	Arabidopsis Orthologue	Reference
A02	34.6 - 34.7	N/A	N/A	
A05	1.5 - 1.6	N/A	N/A	
A06	15.7 - 17.9	N/A	N/A	
A08	15.4 - 16	N/A	N/A	
A09	2.0 - 2.5	N/A	N/A	
C01	38.7 – 38.9	BnaC01g37360D	AT3G13890 (MYB26)	Dawson et al, 1999
C01	37.8 - 38.8	N/A	N/A	
C04	35.1 - 35.2	N/A	N/A	
C06	16.7 - 16.9	N/A	N/A	
C07	43.6 - 43.7	N/A	N/A	
C09	2.8 - 2.9	BnaC09g02740D	AT5G28640 (AN3/GIF1)	Kim and Kende, 2004; Meng et al, 2017
C09	1.9 - 2.1	N/A	N/A	

4.2.5 Narrowing down the gene list for further study

The genes prioritized for further study were *IKU2*, *AN3* and *GDSL1*, due to their known roles in seed size.

IKU2 and *AN3* are both known seed size genes which were identified through GWAS analysis as being associated with variable temperature responses dependent on SNP calls. However, there were limited previous studies to describe the effects of temperature on how these genes were regulated, or if they behaved differently under heat stress. Additionally, IKU2 and AN3 both function in a similar manner in the endosperm with relation to the MINI3 signalling cascade determining development of the endosperm (Garcia et al, 2003; Horiguchi et al, 2005; Meng et al, 2016). As a result of their well-described roles in seed size, *IKU2* and *AN3* were selected for further analysis.

GDSL1 is a member of the GDSL family of genes, generally involved in lipid metabolism. Additionally, GDSL1 has been shown to affect seed germination and oil content in *Arabidopsis* and *Brassica napus* (Ding et al, 2019). Furthermore, the GWAS SNP with the strongest correlation to seed size was in the *GDSL1* gene; this makes *GDSL1* a promising target for further experimental study.

The other candidate genes identified were not selected for further study. This was due to a number of practical concerns which made these genes lower priority than the genes selected. This included that in *Arabidopsis*, *axr6* homozygous mutants suffered from a seedling lethal phenotype; this would have made follow-up experiments in an *Arabidopsis* model difficult. The SNPs associated with *CAND1* and *REV* showed significant interactions, but had a low minor alle frequency – in both, less than 10 varieties contained the minor allele of the SNP used to identify the association from GWAS. As such, these genes were considered lower priority.

The chromosome segment containing a putative *GID1A* gene was identified as important for integument development, which is a factor contributing to overall seed size. However, it was uncertain if a copy of *GID1A* was present at this location of the genome in *Brassica napus*. For the association close to *CLV1*, no significant interaction was identified between gene and environment, meaning there was no strong evidence that the SNP associated with this candidate gene was affecting how TGW changes in response to temperature. As this study focuses on responses to temperature, *CLV1* was not selected for further study. The minor allele frequency for the SNP identified near to

MYB26 was low. Additionally, if this SNP was associated with TGW due to differences *MYB26*, the effect would be due to male fertility effects and so not closely aligned with the goals of this project. Due to these factors, MYB26 was not selected for further study.

4.3 Conclusions & Discussion

The results from comparing SNPs defined as associated with IKU2 suggests that the C allele is favourable for seed size, as plants carrying this allele have larger seeds with TGW remaining high at increased temperatures. *IKU2* is a gene involved in the regulation of endosperm cellularization, which affects the final size of a seed. As an association with seed size was noted close to the location of *IKU2* in the genome, these results suggest that a particular allele of IKU2 could be associated with the production of larger seeds in a temperature-resistant manner. Studies have also indicated that the impact of IKU2 on the endosperm is caused by the genotype of the endosperm tissue itself (Garcia et al, 2005), and so is a zygotic effect rather than maternal. This matches with the fact that the association attributed here to *IKU2* was found in comparisons of seed maturation temperature, rather than vernalisation temperature. Additionally, no studies were found to have investigated whether IKU2 regulation or activity was affected by seed maturation temperature.

Both alleles associated with *AN3* from the GWAS analysis displayed a response to temperature during seed set. However, plants with a G at this location produced a larger number of seeds per pod than plants with an A allele grown at the same temperature. AN3 represses MINI3 to control seed size through regulation of seed expansion; *AN3 Arabidopsis* mutants also showed a reduced number of seeds per silique and reduced fertility (Meng et al, 2016). This may suggest that compared the G- containing allele, an A at this location has a more similar gene function to *Arabidopsis* mutants which produced a lower number of seeds per silique. *AN3* is expressed in *Arabidopsis* in the seed as well as in the ovule and carpel (Meng et al, 2016), which could be related to a role in determining how many viable seeds are produced in a pod. Mutants in *AN3* have

abnormal development of the ovule and gynoecium (Lee et al, 2014). *AN3* expression in seeds alongside its known role in seed size control suggest that there could be differences in alleles of *AN3* between *Brassica napus* lines which may also be related to seed size as well as seed number. Furthermore, differences in seed size based on this SNP call across the DFFS can be explored further, which may help determine the role of *AN3* in determining seed size to temperature responses.

GDSL1 is part of a family of GDSL-type genes which code for GDSL-type esterase/lipase proteins (GELPs), which are widely conserved throughout many plant species including *Arabidopsis*, rice, soybean, wheat, and tomato (Cenci et al, 2022; Chepyshko et al, 2012; Ling et al, 2006; Su et al, 2020; Yang et al, 2022; Sun et al, 2022). These genes were identified in *Brassica rapa* (Dong et al, 2016), revealing 121 genes in this family containing conserved sequences and Ser-Gly-Asn-His motifs; 71% of these genes contained five exons. The GDSL family of genes could encode enzymes involved in lipid metabolism, because changes in *GDSL1* activity/function have been shown to impact seed germination and oil content in both *Arabidopsis* and *Brassica napus* (Ding et al, 2019a). In both species,overexpression of native *GDSL1* led to increased lipid catabolism and decreases in seed oil content, as well as a greater rate of germination; the opposite response was also observed in plants where GDSL1 expression was silenced. As GDSL1 can affect seed properties such as oil content, it is possible that it is responsible for changes in seed development behind the association identified by GWAS. The overall TGW of the seed could be affected in this instance by changes in oil content and lipid breakdown in the developing seed. The differences in response between plants with different alleles suggests that although varieties with the A allele of GDSL1 are temperature sensitive, varieties containing the G allele are not affected by temperature changes. This could mean that one of these alleles is affected in GDSL1 function or expression in a way that allows the G allele to maintain seed oil content at high temperatures while A allele cannot.

The remaining identified candidate genes were not selected for further analysis; however, this does not mean they may not have roles in the seed size response to temperature.

AXR6 and *CAND1* were both identified as candidate genes from GWAS analysis when comparing TGW between seed maturation temperatures of 18°C and 24°C. These genes are also known to interact, with *CAND1* interacting with the unneddylated form of *AXR6* and regulating the formation of a ubiquitin E3 ligase complex containing *AXR6* (Zheng et al, 2002). Neddylation and deneddylation are post-transcriptional modifications where the NEDD8 protein is conjugated or unconjugated from cullin protein subunits, such as those present in *AXR6* (Mergner and Schwechheimer, 2014). Previous studies have also shown that deneddylation of *AXR6* increases under heat stress (Singh et al, 2019). It is possible that the interaction between *AXR6* and *CAND1* is related to control of seed size in a temperature-sensitive manner, and although it was not studied here it could be an interesting study in the future.

Chapter 5 - Exploring the genetic basis of temperature responses in *Brassica napus* seeds

5.1 Introduction

The candidate genes selected for further study in this chapter have been known to affect seed size, but it is not known how these seed size pathways may be impacted by changes in seed maturation temperature. As previous work in this project has indicated that increased seed maturation temperature in *Brassica napus* leads to smaller seeds, examining how the function of these genes may be affected under high temperature could lead to an improved understanding of these seed size pathways and how they could be used to improve crop yield.

As previously discussed, *IKU2* is a protein kinase with a known function in seed size control. *IKU2* promotes endosperm proliferation, with *iku2* mutants showing a small seed phenotype due to precocious endosperm cellularization (Luo et al, 2005). Overexpression of the *Brassica napus IKU2* gene in Arabidopsis leads to increased seed mass (Xiao et al, 2016). However, no studies have examined if this seed phenotype is affected by the temperature at which the seeds develop.

AN3 has been shown to affect the number of seeds produced per silique in *Arabidopsis*, as well as its known role in regulation of seed size by controlling seed expansion via the *MINI3* pathway (Meng et al, 2016). This project has so far identified a putative correlation between genotype at *AN3* on chromosome C09 and seed number per pod in *Brassica napus,* and as *AN3* is known to impact seed size then this effect may also be present. *AN3* has been shown to affect plant responses to abiotic stress, with *an3* knockouts showing improved drought tolerance (Meng and Yao, 2015). However, the effect of temperature on AN3 activity and control of seed size has not yet been studied.

Seed oil content is known to be affected by temperature, with increased seed maturation temperature resulting in reduced seed oil accumulation (Huang et al, 2019). GDSL1 is a member of the GDSL family of lipases and overexpression of GDSL1 in *Brassica napus* is associated with reduced seed oil content as well as reduced seed size (Ding et al,

2019). Previous studies have not specifically explored how GDSL1 affects seed properties under different seed maturation temperatures.

Although the effect of seed maturation temperature on seeds has been explored here in *Brassica napus*, studies have also explored the effect of temperature on *Arabidopsis* seeds. High temperatures can cause a reduction in seed set in *Arabidopsis* (Kipp-Sinanis, 2008; Zinn et al, 2010).

5.2 Hypothesis & Aims

After previously identifying and selecting candidate genes, this chapter details experimental analysis of how variation in these genes may be affecting seed size and its relationship to temperature.

Having previously identified TGW changes associated with difference SNP calls ascribed to *IKU2*, the first hypothesis explored in this chapter is that varieties with different SNP calls have differences in *IKU2* sequence which have an impact on how *IKU2* controls seed size.

Different SNP calls associated with *AN3* were linked to changes in seed number per pod. The second hypothesis tested here is that these differences are similarly linked to changes in TGW.

If different SNP calls associated with *AN3* are affecting TGW, the third hypothesis tested is that varieties with different SNP calls have differences in *AN3* sequence which have an impact on how *AN3* controls seed size.

Different alleles of *GDSL1* were linked to changes in TGW. The fourth hypothesis tested is that varieties with different *GDSL1* alleles have differences in *GDSL1* sequence which have an impact on how *GDSL1* controls seed size.

As *GDSL1* is linked to seed oil content, the fifth hypothesis tested is that there is a relationship between TGW and oil content which is affected by *GDSL1* allele.

A further aim of this chapter is to ascertain if the impacts of temperature on *Brassica napus* seed yield parameters can be replicated in *Arabidopsis*, testing the hypothesis

that mutants in *Arabidopsis thaliana* could be used so understand more about how these genes interact with temperature.

5.3 Results

5.3.1 Exploring *IKU2* as a candidate gene following association with BnaA01g25950D and seed size

5.3.1.1 Identification of IKU2 haplotypes

The SNP originally used to identify *IKU2* as a candidate gene was not found within a *Brassica napus IKU2* gene but was found in a neighbouring gene BnaA01g25950D, 1.1kb from *IKU2*. Following further analysis discussed in Chapter 4.2.3.1, a second SNP was discovered 48kb away in the gene BnaA01g25880D. This section tests the hypothesis that the SNP call identified in BnaA01g25950D correlated with an unknown source of variation in the *IKU2* gene.

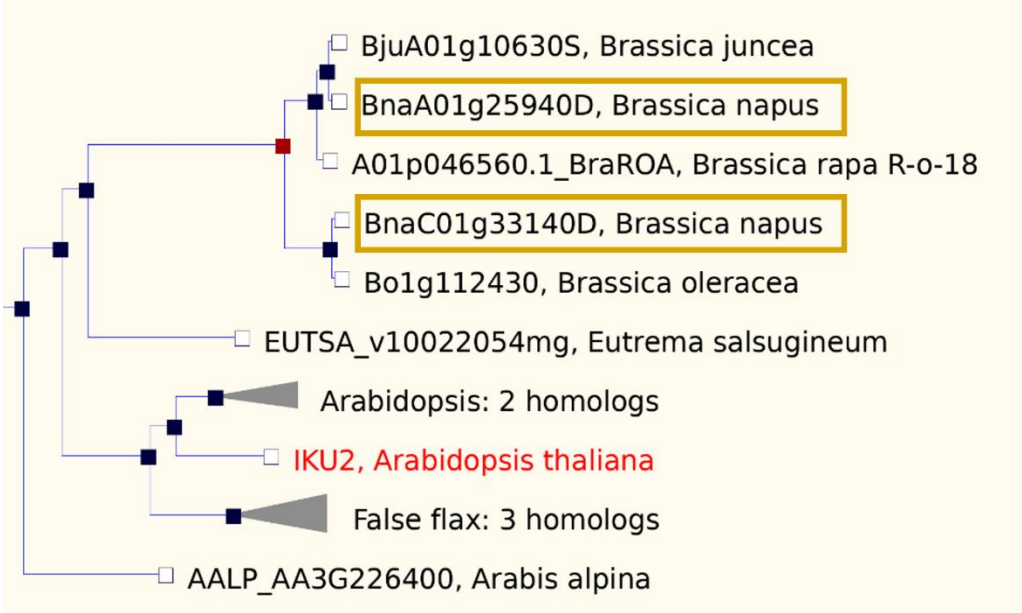

Figure 5.1: Plant Compara gene tree from Ensembl Plants for IKU2. Related genes in Brassica napus are identified by yellow boxes; the Arabidopsis IKU2 gene is highlighted in red. Blue squares indicate speciation nodes while red squares indicate duplication nodes.

No sequence data was available for *IKU2* across the diversity set. The only sequence data available was from the reference variety Darmor-bzh (Chalhoub et al, 2014), which was not included in the diversity set used in this study. The *Brassica napus* genome

contains only two genes which have been identified as orthologues of *IKU2* (Figure 5.1). These two genes are found on chromosomes A01 (BnaA01g25940D) and C01 (BnaC01g33410D). Using the Darmor reference sequence, high sequence similarity was seen between the multiple copies of *IKU2* in the *Brassica napus* genome, with a sequence similarity of 96.21%.

To further investigate the possible genetic causes behind the differences in SNP call seen in GWAS analysis, six varieties with contrasting SNP calls from the diversity set were selected for sequencing of *BnaA01.IKU2* (Table 1).

Table 5.1: Brassica napus varieties selected for sequence analysis.

Variety	**IKU2 SNP call**	**AN3 SNP call**
CANBERRA X COURAGE	C	G
Dimension	C	G
Inca x Contact	C	G
Altasweet	T	A
Janetskis Schlesischer	T	A
Quinta	T	A

The *IKU2* gene in *Brassica napus* is 3.1kb long. To improve efficiency of PCR amplification and sequencing, the genes was amplified in three sections. The primers used are described in Table 5.2. These primers were used to amplify IKU2 from the six varieties shown in Table 5.1 so that sequences could be compared between contrasting SNP calls.

Table 5.2: Primers used for amplifying and sequencing IKU2 in Brassica napus.

Sequence	T_m (°C)	Gene Section	PCR product length (bp)

GGAATCACATGCAACTCCGG	58.7	Start	1067
AGCTGGAGTCTAGGCAAACC	58.4		
CGCAATGACGCATCTTCTGA	57.3	Middle	1250
ACTCCAACCCTTTAGCAGCT	57.5		
TGCAGCATAACGAGTGAGGA	57.5	End	735
TCCAGCATACGAACCACAGA	57.7		

The locations of the primers described in Table 5.2 are shown on a representation of the *BnaA01.IKU2* gene in Figure 5.2. Primer sites overlapped to ensure that areas toward the edges of the amplified region would be sequenced accurately. These primers enabled amplification of the majority of the exons and introns within the gene; however, it is possible that variation was present at the 5' or 3' end of the gene or within promoter regions which was not captured in this experiment.

Figure 5.2: Gene model of IKU2 (5' to 3') showing the locations of primer pairs. Blue boxes indicate exons, with diagonal lines showing the UTR. Grey lines indicate introns. Primer locations are indicated by triangles above the gene model, with dashed lines indicating the amplified section.

The GWAS results details in Chapter 4 suggested that varieties with the C SNP call had larger seeds, and that seed size in these varieties was less sensitive to increased seed maturation temperature. Sequencing was used to identify any differences in *IKU2* associated with the SNP calls.

Figure 5.3: Gene model of IKU2 showing locations of non-synonymous mutations identified between haplotypes. Blue boxes indicate exons, with diagonal lines showing the UTR. Grey lines indicate introns. Yellow triangles indicate the locations of mutations.

Table 5.3: Non-synonymous mutations identified between IKU2 haplotypes.

Varieties	Amino acid sequence change
Dimension; Inca x Contact	N89D
Dimension; Inca x Contact	K117T
Dimension; Inca x Contact	N132I
Dimension; Inca x Contact	A197G
Dimension; Inca x Contact	F311I
Dimension; Inca x Contact	D358Y

Limited differences in sequence were found in *IKU2* between varieties with different SNP calls (Figure 5.3). Six non-synonymous mutations were identified in both Dimension and Inca x Contact which were two varieties containing the C SNP call (Table 5.3); these mutations were not present in CANBERRA X COURAGE, the remaining variety with a C SNP call. This suggests that these six mutations may be linked to the C SNP call, but that the correlation is not absolute. No mutations were identified in any of the three varieties with a T SNP call. This suggests that the T SNP call is identical to the Darmor reference sequence and the known functional *IKU2* gene. This may suggest that the C SNP call is associated with a gain of function in *IKU2*, although the mutations found here do not clearly show by what mechanism this could occur. Further experiments would be required to learn more about how this allele differs from that found in the Darmor reference sequence.

N89D, N132I and A197G were identified as changes in amino acid sequence resulting from mutations in the DNA sequence, but they are present in other copies of the *Brassica napus IKU2* gene so were considered to be most likely a result of non-specific primer binding to the A genome copy. Of the three remaining mutations, F311I does not cause any change in polarity of the amino acid as both phenylalanine and isoleucine are non-polar, so is a conserved amino acid change. At K117, the change from lysine to threonine involves a change from a basic, positively charged amino acid to a polar uncharged amino acid. Additionally, the loss of the long lysine tail may cause reductions in protein stability (Shah and Shaikh, 2016; Song et al, 2021). The D358Y mutation represents a change from a negatively charged amino acid to a polar amino acid. None

of these mutations overlap with the kinase domains or other active sites in the IKU2 protein, so it remains unclear if they affect the function of the IKU2 protein.

As only some of the varieties with the C SNP call showed these mutations, six additional varieties were sequenced to increase the sample size and gain a better understanding of how well these SNPs might correlate with the SNP call identified in the GWAS. Four additional varieties with a C SNP call sequenced were Apex, Topas, Ramses and Castille. Two additional varieties sequenced with a T SNP call were Weihenstephaner and Bronowski. Results of sequencing these twelve varieties are shown in Table 4.

Once again, varieties with a T SNP call were not found to have any non-synonymous mutations. Of the additional four varieties with a C SNP call, three also lacked non-synonymous mutations. However, all six non-synonymous mutations identified previously were present in Apex, which also contained the C SNP call. Increasing the range of varieties sequenced revealed only three of the seven lines tested with a C SNP call also contained the mutations described in Table 3, indicating that any correlation between the SNP call and presence of non-synonymous mutations was weak. Although there were clearly some differences in IKU2 sequences, this was not strong evidence to suggest that the SNPs identified here were consistently linked to the SNP call from the GWAS analysis.

Table 5.4: Mutations present in varieties tested for 6 non-synonymous mutations in BnaA01.IKU2. SNP Call indicates the SNP call present in BnaA01g259501D. Where no mutations were identified relative to the Darmor reference sequence, this is indicated by N/A.

Variety	**SNP Call**	**Mutations Present**
Canberra X Courage	C	N/A
Dimension	C	N89D, K117T, N132I, A197G, F311I, D358Y
Inca x Contact	C	N89D, K117T, N132I, A197G, F311I, D358Y
Apex	C	N89D, K117T, N132I, A197G, F311I, D358Y
Topas	C	N/A
Ramses	C	N/A

Castille	C	N/A
Altasweet	T	N/A
Janetskis Schlesischer	T	N/A
Quinta	T	N/A
Weihenstephaner	T	N/A
Bronowski	T	N/A

5.3.2 Exploring *AN3* as a candidate gene associated with seed number and seed size

5.3.2.1 Allele of *BnaC09.AN3* associated with TGW response to seed maturation temperature

In the GWAS analysis described in Chapter 4, the chromosome C09 copy of *AN3* was identified as a candidate gene due to its effects on SPP. However, as AN3 is also known to affect seed size in Arabidopsis (Meng et al, 2016), the diversity set data was examined to determine if a difference in TGW was also associated with the SNP call.

TGW was affected by maturation temperature in varieties with both alleles, with seeds produced at 24°C being smaller than those produced at 18°C (Figure 5.4). The difference in mean TGW between 18°C and 24°C for varieties with the A allele was 0.637g, while for the G allele this difference was 0.443g. Results of two-way ANOVA (Table 4) show a statistically significant interaction between SNP call and maturation temperature ($p = 0.0158$), indicating that the SNP call has an effect on how TGW changes in response to temperature. At both maturation temperatures, the "A" allele was also associated with larger seeds. This suggests that the difference between the alleles may be related to a differential temperature effect on seed size, with the A allele more strongly affected by temperature and linked to producing larger seeds.

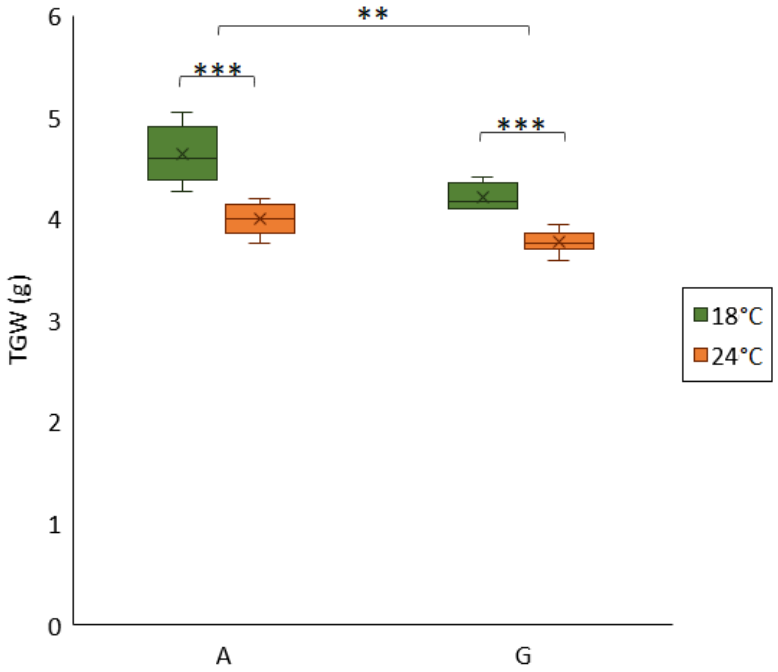

*Figure 5.4: Mean TGW in varieties containing two SNP calls of AN3. Asterisks indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***, produced by Student's t-test.*

Table 5.5: Two-way ANOVA results for effect of SNP in BnaC09g02780D on TGW by maturation temperature. Df = degrees of freedom.

	Df	Sum Squares	Mean Squares	F value	P value
SNP Call	3	45.2	15.06	12.632	3.39E-08
Maturation Temperature (MT)	1	125	125.03	104.878	< 2e-16
SNP Call x MT	3	12.4	4.12	3.456	0.0158
Residuals	2555	3046	1.19		

5.3.2.2 Identification of AN3 haplotypes

There are four orthologues of the Arabidopsis *AN3* gene in *Brassica napus*, found on chromosomes A02, A09, C02 and C09 (Figure 5.5). These genes in *Brassica napus* have approximately 80% similarity to the *Arabidopsis thaliana AN3* gene, as identified using BLAST (Table 5.6).

When *AN3* was identified as a candidate gene, the SNP used to identify the association in the GWAS results was 8kb away from the *Brassica napus AN3* gene on chromosome C09. This section tests the hypothesis that this SNP correlates with an unknown genetic variation in *BnaC09.AN3* which could be responsible for impacting seed traits. Sequence data for *BnaC09.AN3* was available only in the reference sequence Darmor-bzh (Chalhoub et al, 2014) and not for all the varieties in the diversity set. The Darmor reference sequence was used to design primers to try to amplify the C09 copy of *AN3* from varieties in the diversity set which contained either the A or G SNP call.

The *BnaC09.AN3* gene is 2.2kb long. The gene was amplified in three sections (Figure 5.6). The primers described in Table 5.7 were used to amplify *BnaC09.AN3* from the six varieties described above in Table 5.1 (Canberra x Courage; Dimension; Inca x Contact; Altasweet; Janetskis Schlesischer; Quinta) so that sequences could be compared between contrasting SNP calls.

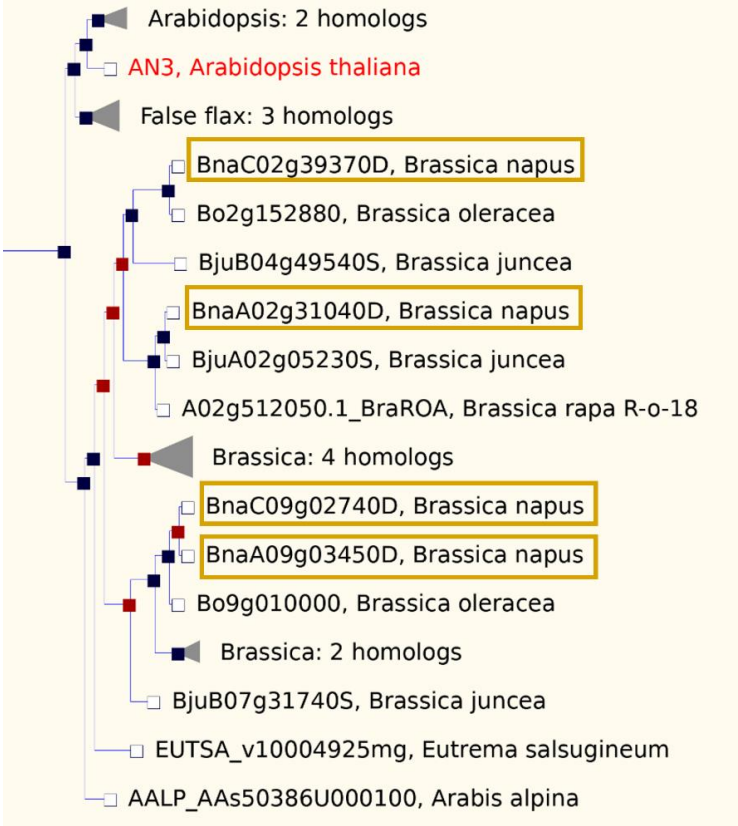

Figure 5.5: Plant Compara gene tree from Ensembl Plants for AN3. Related genes in Brassica napus are identified by yellow boxes; the Arabidopsis AN3 gene is highlighted in red. Blue squares indicate speciation nodes while red squares indicate duplication nodes. Grey triangles indicate collapsed subtrees which do not contain genes of interest.

Table 5.6: Similarity of Brassica napus AN3 orthologues to Arabidopsis AN3, identified using percentage identity to Arabidopsis AN3 genomic sequence with blastn.

Gene ID	**Chromosome**	**Exon Number**	**Percentage Identity**
BnaA02g31040D	A02	4	79.52%
BnaA09g03450D	A09	4	81.83%
BnaC02g39370D	C02	4	80.94%
BnaC09g02740D	C09	4	81.94%

Figure 5.6: Gene model of AN3 (3' to 5') showing the locations of primer pairs. Blue boxes indicate exons, with diagonal lines showing the UTR. Grey lines indicate introns. Primer locations are indicated by triangles above the gene model, with dashed lines indicating the amplified section.

Table 5.7: Primers used for amplifying and sequencing BnaC09.AN3 in Brassica napus.

Sequence	T_m (°C)	Gene Section	PCR product length (bp)
CTACCCCAGCAATGTCACCT	58.6	Start	530
TCAAGCTTGTTCGGTTTAGACT	58.7		
TCTTGTCCAGAGAGCTTCGT	57.5	Middle	848
ATTTACCTGGCTAAGCACGC	56.5		
GCGTGCTTAGCCAGGTAAAT	56.5	End	587
ACACTACCACGAAAAGTTCACT	58.6		

5.3.2.3 Identification of a deletion in *AN3*

When the six varieties chosen for analysis were amplified by PCR for sequencing, it was noted that varieties containing the G SNP produced two PCR products rather than a single product for the 5' end section of *AN3*, which was not seen in varieties containing the A SNP (Figure 5.7a). This fragment was estimated from gel electrophoresis results to have a size difference of 40-50bp compared to the expected band, indicating a possible deletion in one or more copies of *AN3* associated with the G SNP.

As this deletion correlated with the SNP call in the 6 lines tested, it was possible that this deletion was linked to this SNP and was the causative reason for the association seen in the GWAS results. To explore this, all 96 lines used in the phenotyping experiment were screened via PCR to determine if they produced one or two PCR fragments; this was then correlated with SNP call (Figure 5.7b). The majority of varieties with the A SNP produced a single band (75%) while the majority of varieties with the G SNP produced

two fragments (88%). Although not a perfect correlation, this indicated that the presence of this deletion is linked to the SNP call.

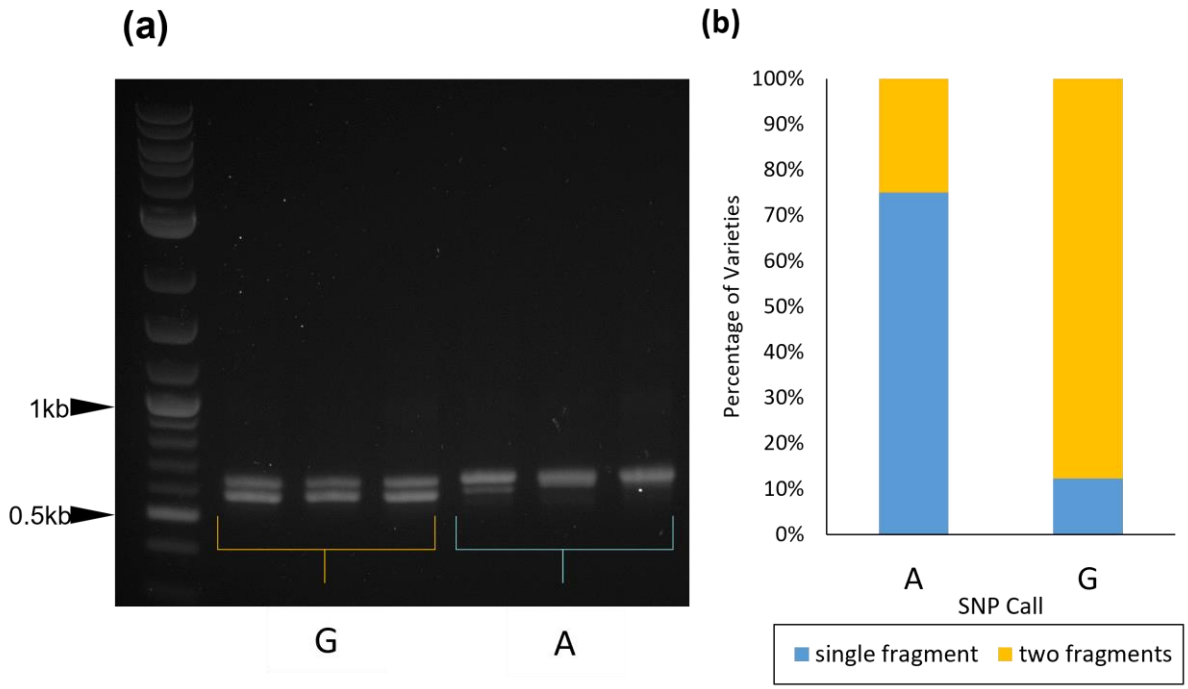

Figure 5.7: PCR results from BnaC09.AN3 amplification. (a) Electrophoresis gel results for 6 varieties with SNP call indicated. Left to right, varieties are: Canberra x Courage; Inca x Contact; Dimension; Altasweet; Janetskis Schlesischer; Quinta. (b) Percentage of varieties with each SNP call with single or double PCR band results from the DFFS.

5.3.2.4 Characterizing the deletion in *AN3*

Colony sequencing was used to isolate individual copies of *AN3* from two *Brassica napus* varieties with contrasting SNP calls and PCR results, namely Canberra x Courage and Quinta. *AN3* from the two different varieties was cloned into *E. coli* and sequenced from the resulting colonies.

There are four copies of *AN3* in the *Brassica napus* genome – BnaA02g31040D, BnaC02g39370D, BnaA09g03450D and BnaC09g02740D. The GWAS results in Chapter 4 identified the C09 copy as being related to changes in seeds (Chapter 4.2.4.1).

The *AN3* genes were then sequenced by sequencing colonies and matching the *AN3* sequence, to all copies of *AN3* from the Darmor reference genome (Figure 5.8). The deletion was confirmed to be present and identified as a 48-bp long deletion located in the 5' UTR of the *AN3* gene (Figure 5.8a). In the Darmor reference sequence, the A02 and C02 *AN3* copies lack this deletion, whereas the deletion can be seen in both the *A09* and *C09* copies (Figure 5.8b). The presence of the deletion in the C09 copy of *AN3* was also identified in the *Brassica napus* variety Canberra x Courage, which contains the SNP call G and showed two fragments in PCR. However, the presence of the deletion was not found in any of the *AN3* copies isolated by colony selection from variety Quinta, which contains the SNP call A and produced a single PCR fragment. This supports the hypothesis that there could be a link between the SNP call and the presence of the deletion; however, sequencing was only carried out in these two contrasting varieties and a more in-depth exploration would be required to fully draw this conclusion.

Figure 5.8: (a) Gene model of AN3 (3' to 5') showing the locations of primer pairs. Blue boxes indicate exons, with diagonal lines showing the UTR. Grey lines indicate introns. Location of deletion is indicated by pink arrow. (b)

Sequence alignment of Brassica napus AN3 genes isolated from colony sequencing. Mismatches are highlighted in red in the rows of the main diagram. Boxes on the left indicate what sequences are being compared; green labels indicate Darmor reference sequence and white labels indicate sequences obtained from colonies.

5.3.2.5 Deletion in *BnaC09.AN3* was not statistically associated with changes in TGW

As all the varieties present in the diversity set had now been genotyped using PCR to identify the presence or absence of the deletion using the number of PCR fragments as an indicator (Figure 5.7b), TGW was compared across the diversity set to determine if there was an effect of the deletion on TGW.

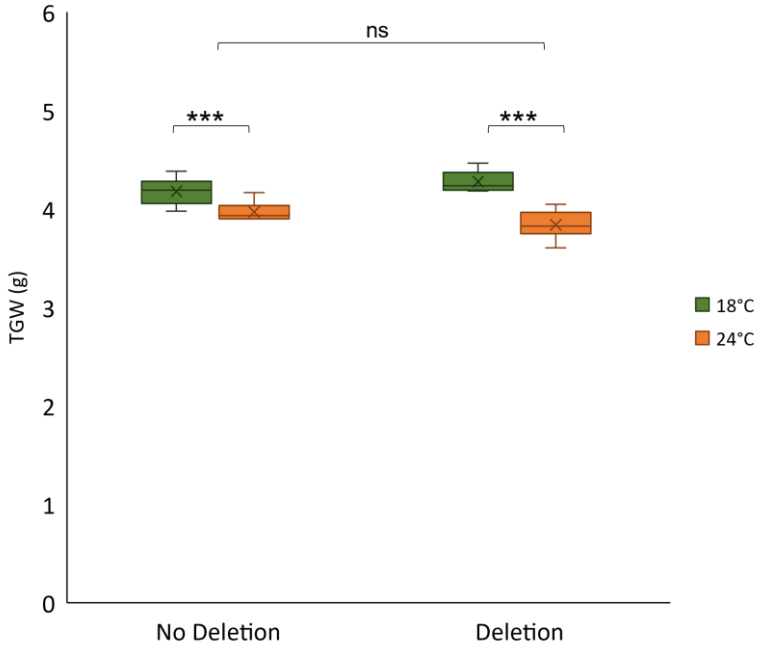

*Figure 5.9: TGW in varieties containing two SNP calls at BnaC09.AN3 by presence or absence of the AN3 C09 deletion. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. Between 18°C and 24°C, p values are: no deletion, p <0.001; deletion, p<0.001. Between deletion and no deletion, p values are: 18°C, p = 0.208; 24°C, p = 0.109 calculated by Students t-test.*

Table 5.8: ANOVA results analysing effect of deletion in BnaC09.AN3 on TGW across two maturation temperatures.

	Df	Sum Squares	Mean Squares	F value	P value
Deletion	1	1.9	1.92	1.592	0.207
Maturation Temperature (MT)	1	129.6	129.63	107.69	<2e-16
Deletion x MT	1	2.5	2.48	2.061	0.151
Residuals	2592	3120	1.2		

Despite the correlation of the deletion with the SNP calls alongside the previous correlation with seed size and SNP call, the presence or absence of the deletion in *AN3* was not found to have a significant effect on TGW (Figure 5.9). TGW was affected by maturation temperature in all varieties regardless of the presence of the deletion, which can be expected as it is known that TGW is reduced at 24°C compared to 18°C across the diversity set (as described in Chapter 2). Two-way ANOVA indicated that the presence of the deletion had no significant effect on TGW, and that there was no significant interaction between maturation temperature and the presence of the deletion (Table 5.8). These results indicate that the deletion identified using PCR does not affect TGW or how TGW changes in response to maturation temperature.

5.3.3 Exploring *GDSL1* as a candidate gene associated with seed size

5.3.3.1 Identification of *GDSL1* haplotypes

The *Brassica napus* genome likely contains at least three genes which are orthologues of the Arabidopsis gene *GDSL1* found on chromosomes A08 and C03, together with an additional copy which is not fully annotated but is likely to be found on chromosome A07 (Figure 5.10). The Brassica pangenome (He et al, 2015) suggests the existence of a fourth copy, which is not annotated on Ensembl plants, on chromosome C07; as seen in

Figure 5.10, it is possible this gene was inherited via the *Brassica oleracea* gene Bo7g045230.

From the GWAS results (details in Chapter 4.2.2.3), a SNP was identified on chromosome A08 within the *BnaA08.GDSL1* gene, BnaA08g17820D. *GDSL1* sequences for all 96 lines used in this experiment was available from exome capture data (Woodhouse et al, 2021). This allowed the SNP to be correlated with haplotypes and two main haplotypes were identified (Figure 5.11).

Four haplotypes were identified in total, although haplotype 4 was represented only by the variety Samourai and so was not analysed.

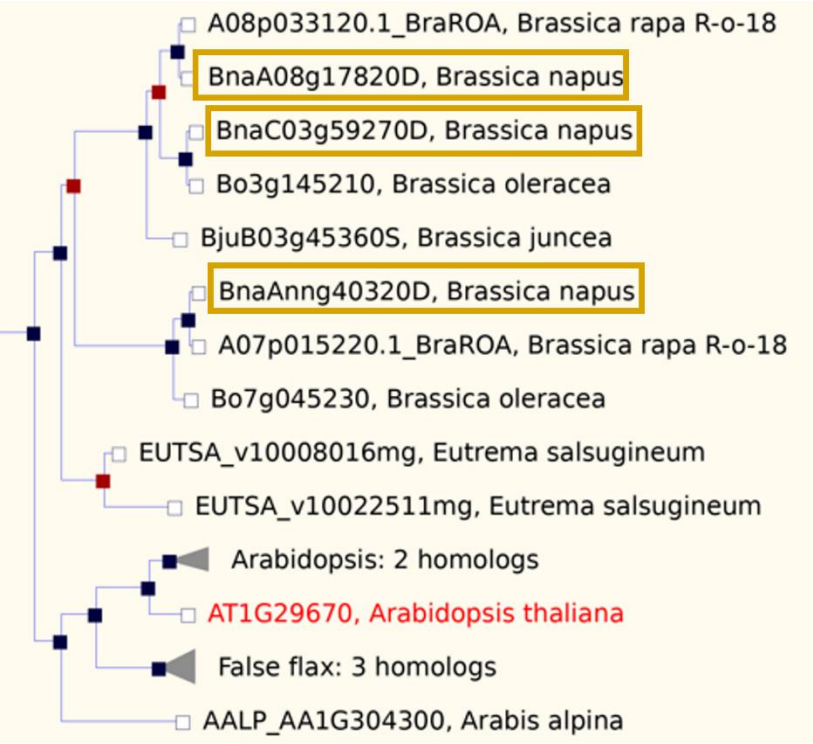

Figure 5.10: Plant Compara gene tree from Ensembl Plants for GDSL1. Related genes in Brassica napus are identified by yellow boxes; the Arabidopsis GDSL1 gene AT1G29670 is highlighted in red. Blue squares indicate speciation nodes while red squares indicate duplication nodes.

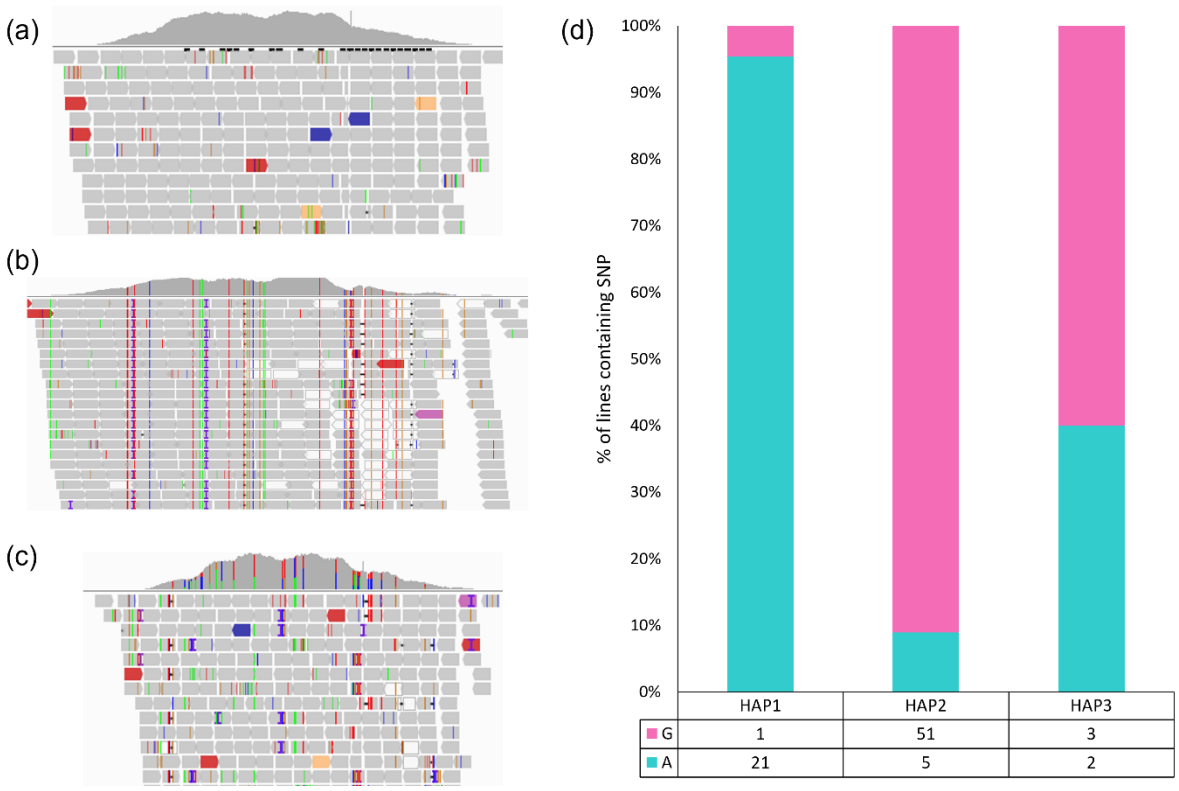

Figure 5.11 : GDSL1 haplotyping. (a) - (c) shows the IGV viewer output of short read sequences aligned to GDSL1 and consensus sequence for haplotypes 1, 2 and 3 respectively. (d) shows the percentage of varieties with each haplotype which contain either a G or A allele of GDSL1 from GWAS results. Numbers below the graphs show the real number of varieties in each group.

The two main haplotypes identified correlated with the SNP calls. The first haplotype, associated with the A SNP call, was identical to that in the Darmor reference sequence (Figure 5.11a). The second haplotype was associated with G SNP call and contained a number of SNPs and deletions compared to the Darmor reference sequence (Figure 5.11b). The third haplotype contained SNPs but lacked the deletion seen in haplotype 2 (Figure 5.11c); additionally, haplotype 3 was not found to be associated with a particular SNP call (Figure 5.11d). As haplotype 3 was found in only 5 varieties and was not associated with either SNP call, further analysis was focused on haplotypes 1 and 2. It is possible that haplotype 3 was associated with variation in copy number but this was not studied further.

The locations of the SNPs identified in the three haplotypes are visualized in Figure 5.12. While haplotype 1 matched the Darmor reference sequence (Figure 5.12a), both haplotypes 2 and 3 contained a number of SNPs at the same points (Figure 5.12b-c). Haplotype 2 contained a non-synonymous mutation in exon 3; in haplotype 3, the base call was ambiguous at this location so it is unclear if this mutation is shared between haplotypes 2 and 3. As the majority of SNPs present in haplotype 2 are also found in haplotype 3, the main difference between these haplotypes appears to be the presence of deletions.

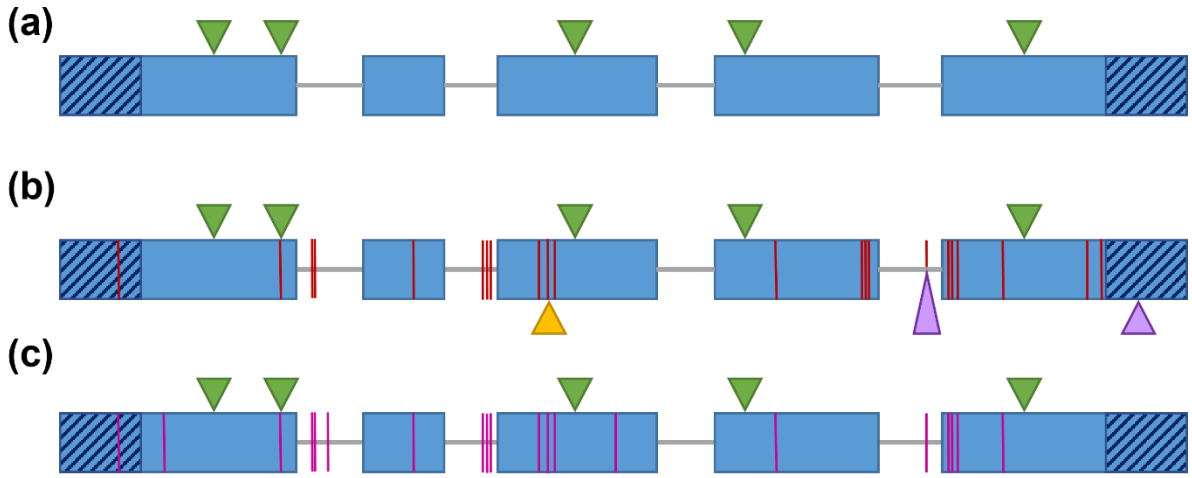

Figure 5.12: Gene model of GDSL1. Blue boxes indicate exons, with diagonal lines showing the UTR. Grey lines indicate introns. Green arrows above the boxes show the location of conserved blocks I-V as described in Ding et al, 2019. (a) haplotype 1; (b) haplotype 2. Yellow arrow below shows location of amino acid change. Purple arrows show location of deletions. Red vertical lines indicate SNPs. (c) haplotype 3. Purple vertical lines indicate ambiguous SNP calls.

Table 5.9: ANOVA results analysing effect of GDSL1 haplotype group on TGW by maturation temperature.

	Df	**Sum Squares**	**Mean Squares**	**F value**	**P value**
Haplotype	3	12.6	4.19	3.736	0.0108
Maturation Temperature (MT)	1	89.3	89.34	79.742	< 2e-16
Haplotype x MT	3	28.8	9.58	8.555	1.20E-05
Residuals	2301	2577.9	1.12		

In haplotype 1, mean TGW was significantly higher that HAP2 and was affected by temperature, with a difference in means between 18°C and 24°C of 0.765g whereas in haplotype 2 there was no significant effect of maturation temperature on TGW (Figure 5.13). Two-way ANOVA indicated a significant interaction between haplotype and maturation temperature (Table 5.9). This suggests that the presence of SNPs and deletions in haplotype 2 are correlating with GDSL1 function in a way that limits the effect of temperature on seed size, causing varieties with haplotype 2 to behave differently to haplotype 1.

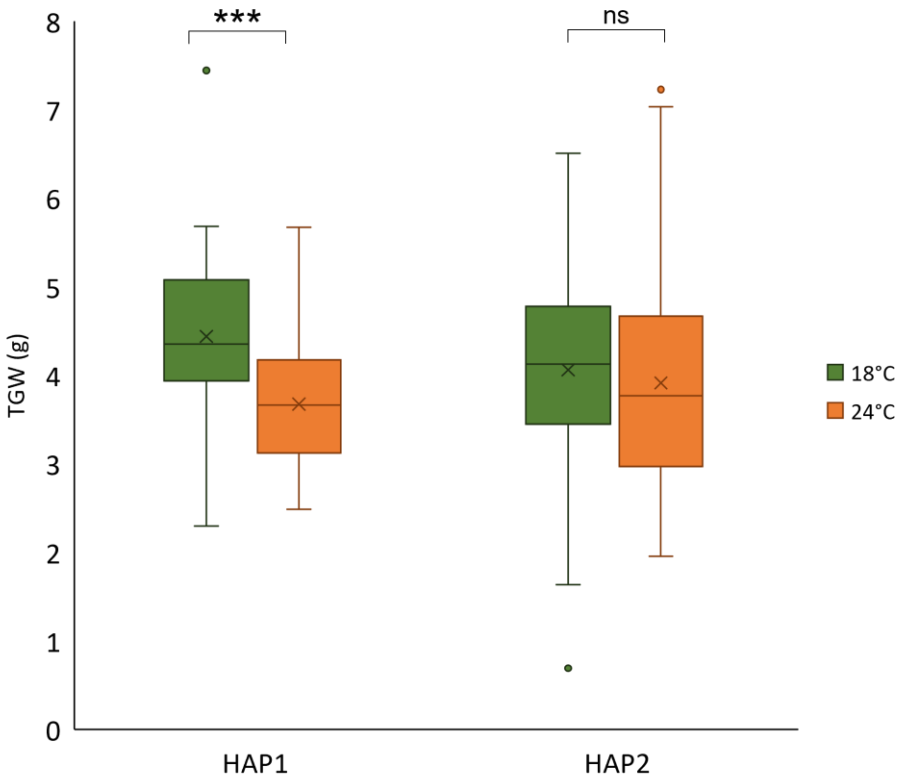

*Figure 5.13: A comparison of TGW between GDSL1 haplotypes at two maturation temperatures. Asterisks indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. Between 18°C and 24°C, p values are: HAP1, p = 6.08E-05; HAP2, p = 0.389 (p-values produced by Student's t-test). For HAP1, n = 23 varieties; for HAP2, n = 56 varieties.*

5.3.3.2 Differences in GDSL1 haplotypes

The structure of *GDSL1* is detailed in Figure 5.14, with five exons and four introns. *GDSL1* contains five conserved sequence blocks, which contain four important catalytic

residues (Ding et al, 2019). As there are deletions and SNPs seen only in haplotype 2, this suggests that one or more of these changes are the cause of the difference in temperature response between haplotypes.

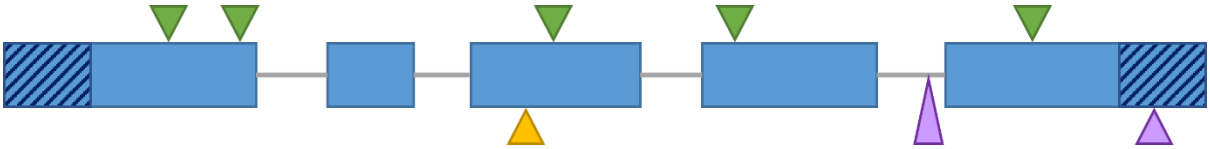

Figure 5.14: GDSL1 gene structure (5' to 3'), showing features of haplotype 2. Blue boxes indicate exons, with diagonal lines showing the UTR. Grey lines indicate introns. Green arrows above the boxes show the location of conserved blocks I-V as described in Ding et al, 2019. Yellow arrow below shows location of amino acid change. Purple arrows below show locations of deletions.

Although there are a number of SNPs present in haplotype 2 compared to haplotype 1, only one of these results in a change to the amino acid sequence, changing an alanine residue to glutamic acid in exon 3 (A148E) and therefore introducing a negatively charged amino acid in place of a non-polar residue.

There is a small deletion of 4bp (TGTG) in the 3' untranslated region of GDSL1 in haplotype 2. The largest deletion is a 12bp deletion (TAT TTG TTT CTC) in intron 4, just before exon 5 in the GDSL1 gene, although did not overlap with the splice acceptor site.

All of the active sites are conserved between haplotypes 1 and 2, with no effect of SNPs or deletions on their sequence. The transmembrane domain and signal peptide sequence is also conserved between both haplotypes, based on their predicted location from comparison to AtGDSL1 (Ding et al, 2019). While there are changes in the amino acid sequence present in haplotype 2 compared to haplotype 1, it is unclear what impact these changes might have on the function or expression of GDSL1.

5.3.3.3 Relationship between *GDSL1* haplotype and oil content

Oil content at 9% moisture was determined by NMR in six varieties. Seeds were sourced from those produced in the large glasshouse phenotyping experiment detailed in Chapter 2, from plants which had undergone a 15°C vernalisation treatment for 12 weeks. This treatment was selected due to experimental limitations and as the GWAS association with GDSL1 was identified in plants which had experienced these

conditions. Oil content was measured for seeds produced at both 18°C and 24°C to determine the effect of seed maturation temperature and *GDSL1* haplotype on oil content.

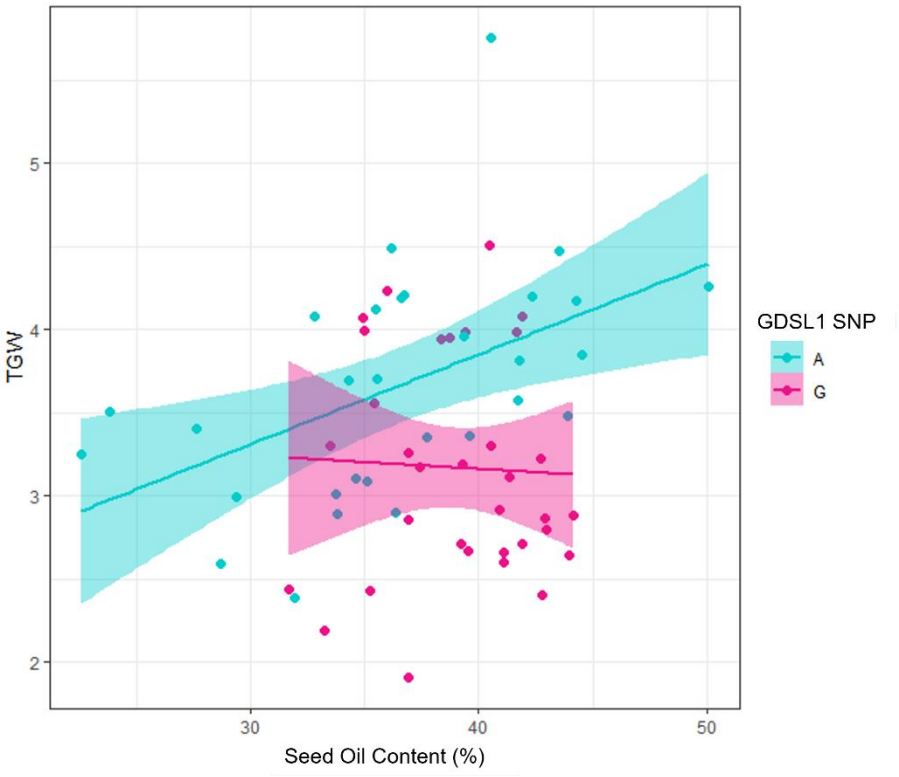

Figure 5.15: Correlation between mean TGW and seed oil content at 9% moisture for two haplotypes of GDSL1. Line of best fist fitted using a linear model. Areas around line of best fit show 95% confidence interval.

Table 5.10: Statistical correlation analysis of TGW and seed oil content between two haplotypes of GDSL1. Data shown to 3 significant figures.

Haplotype	**Spearman's rho**	**p-value**
1 / A	0.589	0.000768
2 / G	-0.0996	0.58

Figure 5.15 shows that for haplotype 1, as seed oil content increases so does TGW; the same relationship is not seen in haplotype 2, with a smaller range of TGW. This suggests that the difference in *GDSL1* sequence between the two haplotypes is affecting the relationship between seed size and oil content.

In haplotype 1 (SNP call A), which resembles the functional *GDSL1* copy in Darmor, there was a moderate correlation between TGW and oil content (Table 5.10), suggesting there is a relationship between the two. However, in haplotype 2 (SNP call G), which has a number of SNPs compared to the version of *GDSL1* found in haplotype 1, there is no statistical correlation between TGW and oil content.

5.3.4 Replicating *Brassica napus* results in *Arabidopsis thaliana*

The effects of temperature on seed size in this project have so far been identified in *Brassica napus*. Experiments were carried out to learn if these results could be replicated in the model plant *Arabidopsis thaliana*. If seed maturation temperature was found to have similar effects in both species, this would allow exploration on the function of candidate genes to be carried out in the smaller and faster growing model species.

To test the hypothesis that increased seed maturation temperature would lead to a reduction in seed weight in *Arabidopsis thaliana,* plants were grown 21°C until bolting began and then moved to either 15°C or 25°C until seed production was complete. This was designed to mimic the experimental setup of the large glasshouse phenotyping experiment in Chapter 2. As the plants used were either Columbia-0 or Landsberg *erecta*, they did not have a vernalisation requirement and so did not undergo a vernalisation treatment.

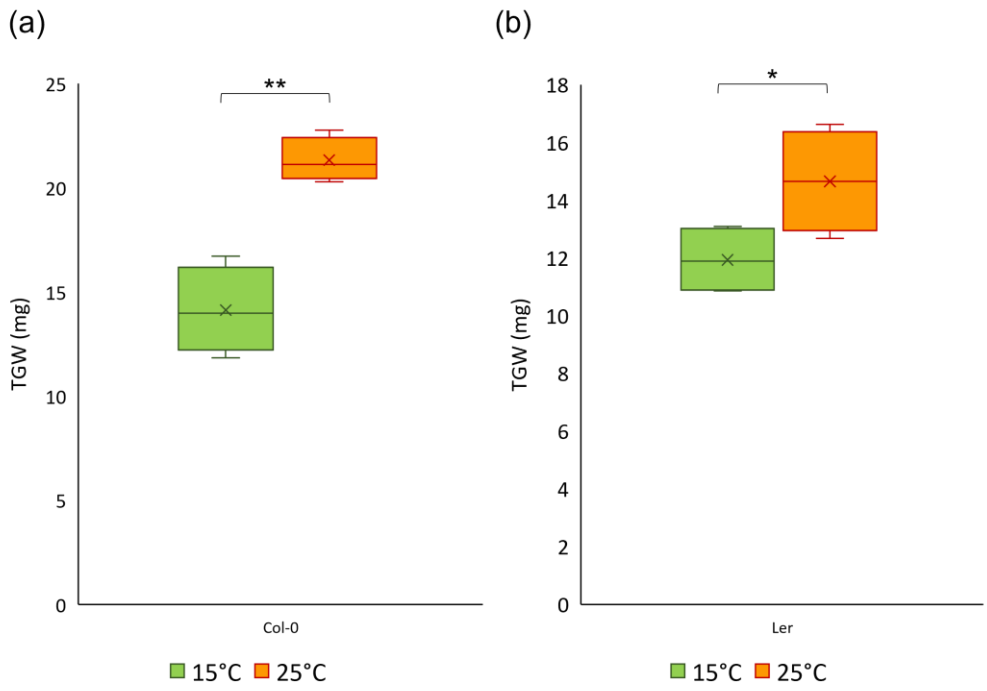

*Figure 5.16: TGW at two seed maturation temperatures in Arabidopsis thaliana. (a) Col-0, (b) Ler. Asterisks used to indicate significance, where p ≤ 0.05 is *, p<0.01 is ** and p< 0.001 is ***. Between 15°C and 24°C, p-values are (a) p =0.00234, (b) p = 0.0494, produced by Student's t-test.*

In *Arabidopsis thaliana*, higher temperatures are associated with increased seed weight in both Columbia and Landsberg ecotypes (Figure 5.16). In Columbia, increasing the temperature from 15°C to 25°C resulted in a seed weight increase of 7.19mg, while in L*er* the seed weight increase was 2.72mg. This is in contrast to the results seen in *Brassica napus*, where increasing the seed maturation temperature resulted in lower seed weight (as described in Chapter 2.3.5). From these results it could be concluded that the temperature of seed maturation affects *Arabidopsis thaliana* differently compared to *Brassica napus*. As such, the effect of seed maturation temperature on seed weight seen in *Brassica napus* could not easily be studied in the model plant.

5.4 Conclusions & Discussion

The GWAS association at BnaA01g26020D was first hypothesised to being linked to changes in *BnaA01.IKU2* as it was a known seed size gene near to the SNP. However, SNPs within *IKU2* were not present in the SNP data file used for GWAS as this data came from leaf transcriptomes (Harper et al, 2012) and *IKU2* is expressed only in seeds. As this GWAS peak was not identified within the *IKU2* gene, it is possible that another SNP nearby was responsible for the GWAS result or that the result was due to chance. It is also possible that the association was related to changes in *IKU2* that I did not uncover and that a wider analysis of *IKU2* could reveal differences. However, in this study I was unable to identify any link between *IKU2* gene sequence and temperature effects on seed size.

From the results detailed in this chapter, although a deletion was identified in *BnaC09.AN3* it was not found to have an impact on seed weight. The deletion identified was located in the 5' untranslated region of the gene. This region is important for regulation of gene translation (Floris et al, 2009), and so changes here may influence protein levels of *AN3*. The deletion was linked to the G SNP call, which was associated with smaller seeds. However, as the effect on seed size seen between varieties with contrasting SNP calls was no longer present when looking at how TGW correlated with presence of the deletion, it is unlikely that the deletion was the source of the effect of the SNP on TGW changes in response to maturation temperature.

If *AN3* protein levels were affected by this deletion, this would suggest that these were higher in these lines as plants with a loss-of-function mutation in *AN3* produce larger seeds (Meng et al, 2016). This could indicate that the deletion is impacting the ability of a repressive complex to bind to the gene. However, as the deletion itself was not found to correlate with a significant difference in TGW, this may not be the actual effect of the deletion. It is possible that this deletion does have an impact on *AN3* transcriptional regulation which could then be considered for any effects in seed size. To determine this, further experimental work using RT-qPCR could be carried out to assess changes in AN3 expression between different varieties of *Brassica napus*.

The G SNP call was here associated with lower TGW, but was also previously associated with an increase in SPP. This may indicate that seed size effects, rather than being a direct result of the action of AN3, are instead a result of the trade-off between seed size and seed number seen in *Brassica napus* (Cai et al, 2014). Additionally, if changes in seed number are being offset by changes in seed size, differences in AN3 between varieties may not have a significant impact on overall yield.

GDSL1 is a member of the GDSL esterase/lipase gene family (GELPs), which contains a range of lipolytic enzymes which hydrolyse lipids (reviewed in Shen et al, 2022). GELPs are defined by their unique and conserved GDSL motif (GDSxxDxG) at the N-terminus of the protein (Ding et al, 2019b). These genes have been shown to have a range of functions in oil metabolism throughout development. *BnGILP1*, a GELP gene in *Brassica napus*, is involved in oil catabolism during rapeseed germination (Ling et al, 2006; Ding et al, 2019b); other GDSL lipase family genes are hypothesized to regulate oil synthesis during embryo development of oilseed plants as well as oil catabolism during seed maturation (Clauss et al, 2011; Chen et al, 2012; Ding et al, 2019a). Overexpression of *GDSL1* in *Brassica napus* is associated with reduced seed oil content as well as reduced seed size (Ding et al, 2019), suggesting that this gene also has a role in determining seed oil content.

Analysis of *GDSL1* revealed two main haplotypes which corresponded with SNP calls identified in GWAS, with differences in both temperature responses and the relationship between TGW and oil content of the seed. This indicates that when *GDSL1* is functioning normally, TGW and oil content are linked and suggests that *GDSL1* allele represented in haplotype 1 may be required for increases in seed oil content to result in larger seed size. The breakdown of the relationship between seed size and oil content in haplotype 2, combined with the reduced effect of seed maturation temperature on these parameters, suggests that one of the differences in sequence present in this haplotype is responsible for altering *GDSL1* function in a manner which also affects this relationship.

Only one difference was found in the protein sequence between haplotypes, a change from an alanine to a glutamic acid at amino acid position 148. As this represents a

change in charge, it is possible that the negative charge gained from the glutamic acid could affect protein structure due to the differences in how these two residues interact with other amino acids in the protein (Biro, 2006). This change was 9 amino acid residues away from conserved block III, which contains an active catalytic site at an asparagine residue (Ding et al, 2019a), meaning changes in protein structure here could cause changes to the availability of this active site. A number of synonymous SNPs were identified Although the amino acid sequence remained the same at these locations, it has been suggested that synonymous changes in the codon can also affect protein structure (Gu et al, 2003; Saunders and Deane, 2010) or miRNA binding sites (Gu et al, 2012). Although not explored here, it is a possibility that *GDSL1* function is different between the two haplotypes due to changes not identified here, such as changes in protein structure, differences in promoter regions or variation in copy number between varieties.

Furthermore, there were two deletions identified in *GDSL1* haplotype 2 – a 12bp deletion in an intron close to the beginning of exon 5, and a 4bp deletion in the 3' untranslated region. It is possible these deletions affected regulatory sites, particularly the deletion in the 3' untranslated region as this area of the mRNA is known to have regulatory functions (Bernardes and Menossi, 2020). No SNPs or deletions present in haplotype 2 affected intron splice sites within the intron, meaning that none of the differences between haplotypes should cause a difference in splicing.

Overexpression of *BnGDSL1* leads to reduced oil content, while silencing of *BnGDSL1* leads to higher oil content (Ding et al, 2019a). This may indicate that *GDSL1* activity is either reduced or impaired in larger seeds, such as those seen in haplotype 1. This could suggest that the function of a negative regulatory element within *GDSL1* was impacted by a SNP or deletion in haplotype 2. It may be an interesting future experiment to compare expression levels of *GDSL1* between the two haplotypes to learn more about any regulatory effects of these deletions.

Results of growing *Arabidopsis thaliana* under different temperature treatments indicated that Arabidopsis is not a suitable model for exploring the effects of seed

maturation temperature seen in *Brassica napus*. Higher temperatures cause increased TGW in Arabidopsis but reduced seed weight in Brassica, suggesting the seed size response to temperature may be affected differently between the two species at the temperatures used in this experiment. It is possible that under these growing conditions, *Arabidopsis thaliana* plants produced at higher temperatures produced fewer seeds, as high temperatures are known to cause reductions in seed set in Arabidopsis; this reduction in seed number may have then been offset by a rise in seed size.

The ecotypes of Arabidopsis used in these experiments were Col-0 and L*er*, neither of which have a vernalisation requirement. This is in contrast to *Brassica napus*, where some varieties (in particular Winter OSR) have a strong vernalisation requirement. However even in Spring or Semi-winter varieties of OSR, which lack a strong vernalisation requirement, TGW decreased at higher maturation temperatures. The difference in seed size response between species is therefore unlikely to be linked directly to the vernalisation requirement. It may be the case that genetic pathways controlling seed size in *Brassica napus* have a greater degree of integration with vernalisation pathways when compared to *Arabidopsis thaliana*, and this difference in treatment may be linked to their contrasting seed size responses to temperature. Further insight could be gained here by studying the effect of a wider range of seed maturation temperatures on seed size in a vernalisation-requiring ecotype of *Arabidopsis thaliana*.

Chapter 6 - Effects of Vernalisation on Gene Expression and Development

6.1 Introduction

It is well known that vernalisation impacts gene expression within the plant. As part of the vernalisation response, *FLC* levels decrease and the expression of *FLC* target genes is impacted through floral developmental processes. The question remains as to what the impact is of these gene expression responses to vernalisation on gene expression later in reproductive development, such as in the seed.

Increased vernalisation temperature has been previously shown here to be associated with decreases in seed size in *Brassica napus,* prominently in Winter OSR varieties (see Chapter 2.3.4; Chapter 2.3.7.3). Other previous work carried out in the Penfield group (unpublished data, described in Chapter 1) also supports this, with plants from a Winter x Semi-winter OSR population which have not undergone laboratory vernalisation treatment producing smaller seeds than plants vernalised in standard 5°C conditions.

FLC has been previously described to have effects within the seed. *FLC* has been shown to affect seed dormancy and contribute to germination behaviour in *Arabidopsis* in a temperature-dependent manner (Chiang et al, 2009; Chen and Penfield, 2018). It remains unknown by what mechanism *FLC* affects dormancy, or how this process may affect seed size in *Brassica napus*. This raises the question of whether changes in *FLC* expression resulting from vernalisation could also be influencing seed development or pathways regulating seed size, as well as how these impacts could be more or less pronounced based on vernalisation temperature.

6.2 Aims & Hypotheses

The first hypothesis tested by the experiments described in this chapter is that differences in temperature during the floral transition will translate to differences in the transcriptome of the shoot apex or the gynoecia.

The second hypothesis is that differences in vernalisation temperature will cause differences in gene expression during reproductive development and in the developing seed.

As differences in seed size resulting from changes in vernalisation temperature have already been described in this thesis, another aim of this chapter is to clarify when during development these changes arise.

6.3 Results

6.3.1. Effect of floral transition occurring in the cold in a Semi-winter OSR variety

A previous experiment had been carried out prior to this project (unpublished data, Rachel Wells), sampling transcriptomes from Zhongshuang 11 (ZS11) plants with a three-week period of vernalisation. Here, a similar experiment was carried out using a twelve-week vernalisation period in order to compare changes in gene expression between the two treatments. In both experiments, weekly sampling of the shoot apex was carried out beginning at 1 week after sowing and weekly sampling continued until the plant began to bolt (Figure 6.1). After this stage, tissue samples were taken for transcriptome analysis from gynoecia and at BBCH51. The BBCH scale describes the developmental stage of *Brassica napus*, where BBCH51 corresponds to the visible appearance of floral buds and BBCH60 denotes the point when the first flowers open (Lancashire et al, 1991).

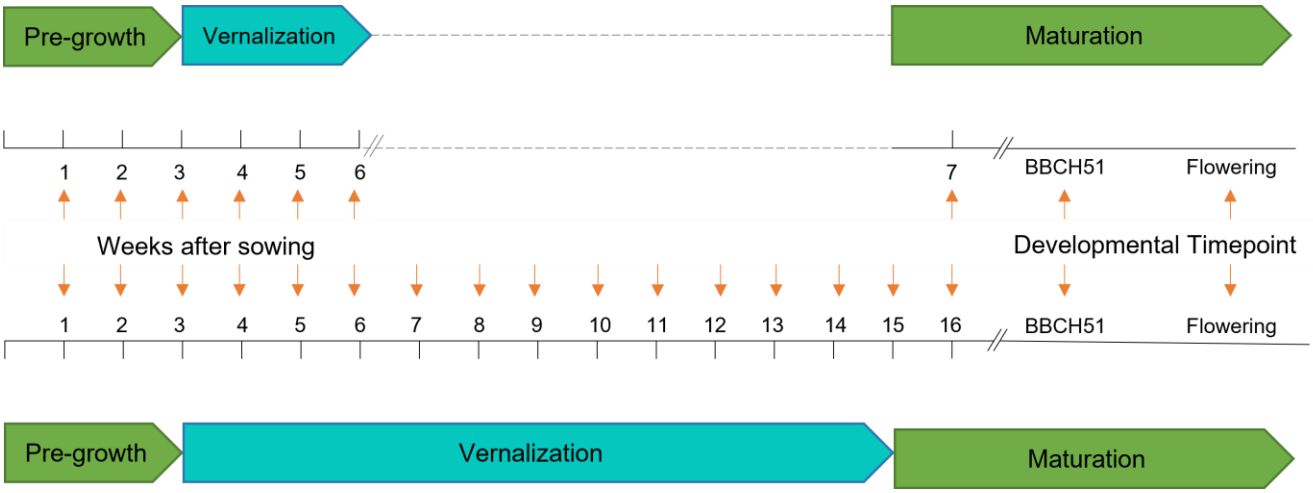

Figure 6.1: Experimental design of ZS11 transcriptome experiments. Orange arrows indicate sampling timepoints.

6.3.1.1 Floral transition occurring in the cold

Along with weekly sampling for transcriptome analysis, three plants were taken for developmental imaging on a weekly basis to determine the developmental stage of the plants. During the twelve-week cold treatment, it was found that the floral transition occurred after approximately 7 weeks of growth at 5°C, when the plant was 10 weeks old (Figure 6.2). After six weeks of growth at 5°C, the apex was still in a vegetative state, with flattened leaf primordia surrounding the shoot apex and the apex itself relatively flat (Figure 6.2a). After 7 weeks of growth at 5°C (Figure 6.2b), the apex has become more domed and rounded floral primordia are visible, indicating that the floral transition has begun. This growth continues after 8 weeks of growth at 5°C (Figure 6.2c), and after 9 weeks there are many floral primordia along with visible differentiation into bud structures beginning (Figure 6.2d). During the work carried out in this these, a study using Winter OSR was published confirming that *Brassica napus* crops in the field undergo floral transition in the cold (O'Neill et al, 2019).

As the shoot apex was still vegetative after 3 weeks of vernalisation, this meant that an additional difference between the two experiments was revealed. Along with the difference in vernalisation duration, the plants also experienced different temperatures

during the floral transition and early reproductive growth. Either or both of these effects could independently affect later development.

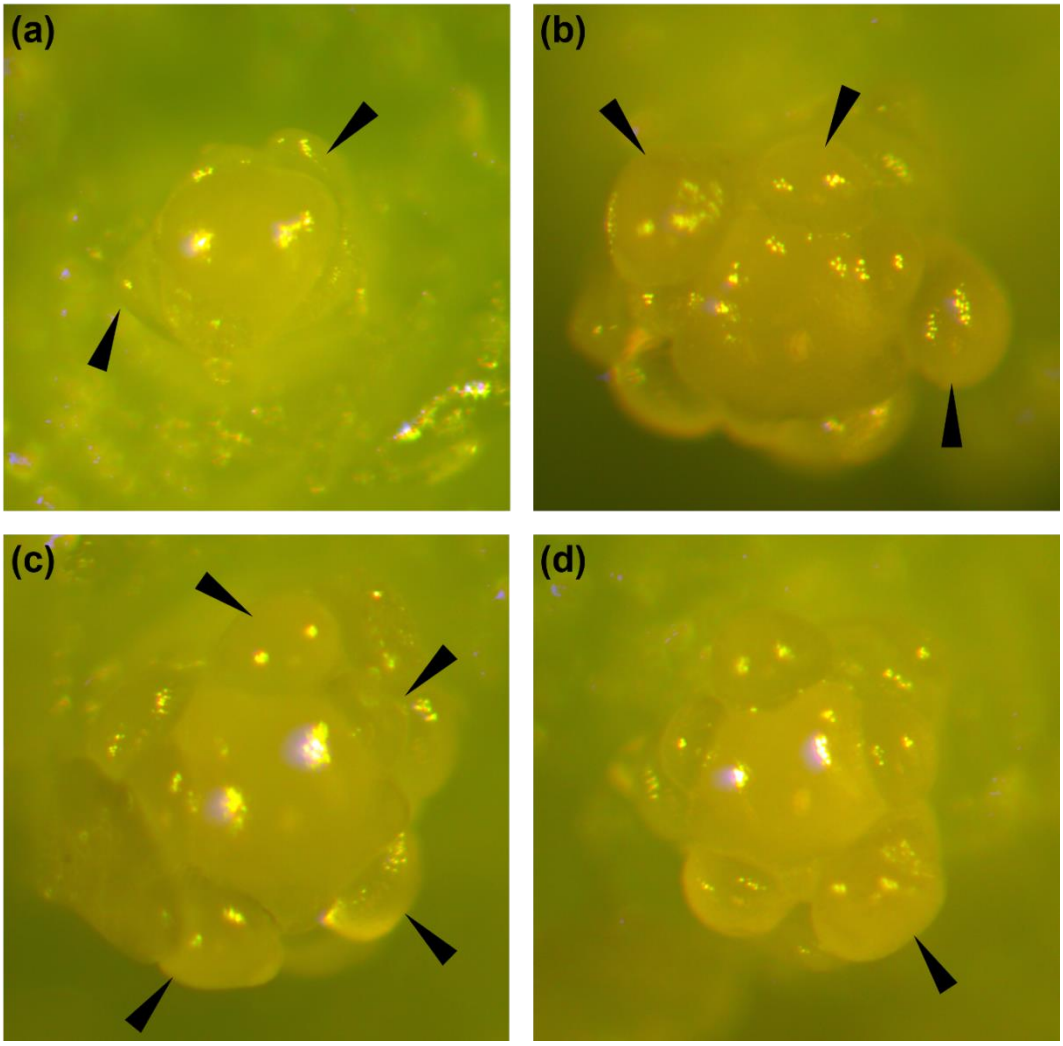

Figure 6.2: Images showing ZS11 shoot apex development in 5°C after (a) 6 weeks (b) 7 weeks (c) 8 weeks (d) 9 weeks. Black arrows indicate (a) leaf primordia, (b - c) floral primordia, (d) differentiating floral primordia.

6.3.1.2 Differential gene expression by temperature during floral transition

Gene expression was compared between tissues from Zhongshuang 11 plants that had undergone either 3 weeks or 12 weeks of cold treatment. As the floral transition was found to occur after 7 weeks in vernalisation at 5°C, this comparison was able to

examine how gene expression changed when the floral transition occurred at 5°C compared to at 18°C.

Comparisons were made for the following time points: apex of 21-day old plants, before vernalisation treatment had begun; BBCH51 (188d for 12-week treatment and 59d for 3-week treatment); and gynoecia.

Principal component analysis (Figure 6.3) showed the three repeats for each sample grouping together, with the apex samples progressing along the PC1 axis with time and samples separated along PC2 by tissue type. Additional samples were taken from the 12-week treatment from plants after vernalisation treatment had finished, but there was no direct timepoint comparison in the 3-week treatment.

Figure 6.3: Principal component analysis chart for RNA transcriptome samples from Zhongshuang 11 plants in 12-week cold experiment. Timepoints are: 21d = pre-vernalisation apex samples; 113d = post-vernalisation apex samples; 118d = BBCH51; Sanders stage 12-13 = gynoecia samples. Percentage variation explained was: PC1 = 30%, PC2 = 18%.

When comparing pre-vernalisation samples taken at 21 days after sowing between the two treatments (before plants entered vernalisation), there were a large number of differentially expressed genes. In the 3-week treatment, 6715 genes were upregulated and 2095 were downregulated.

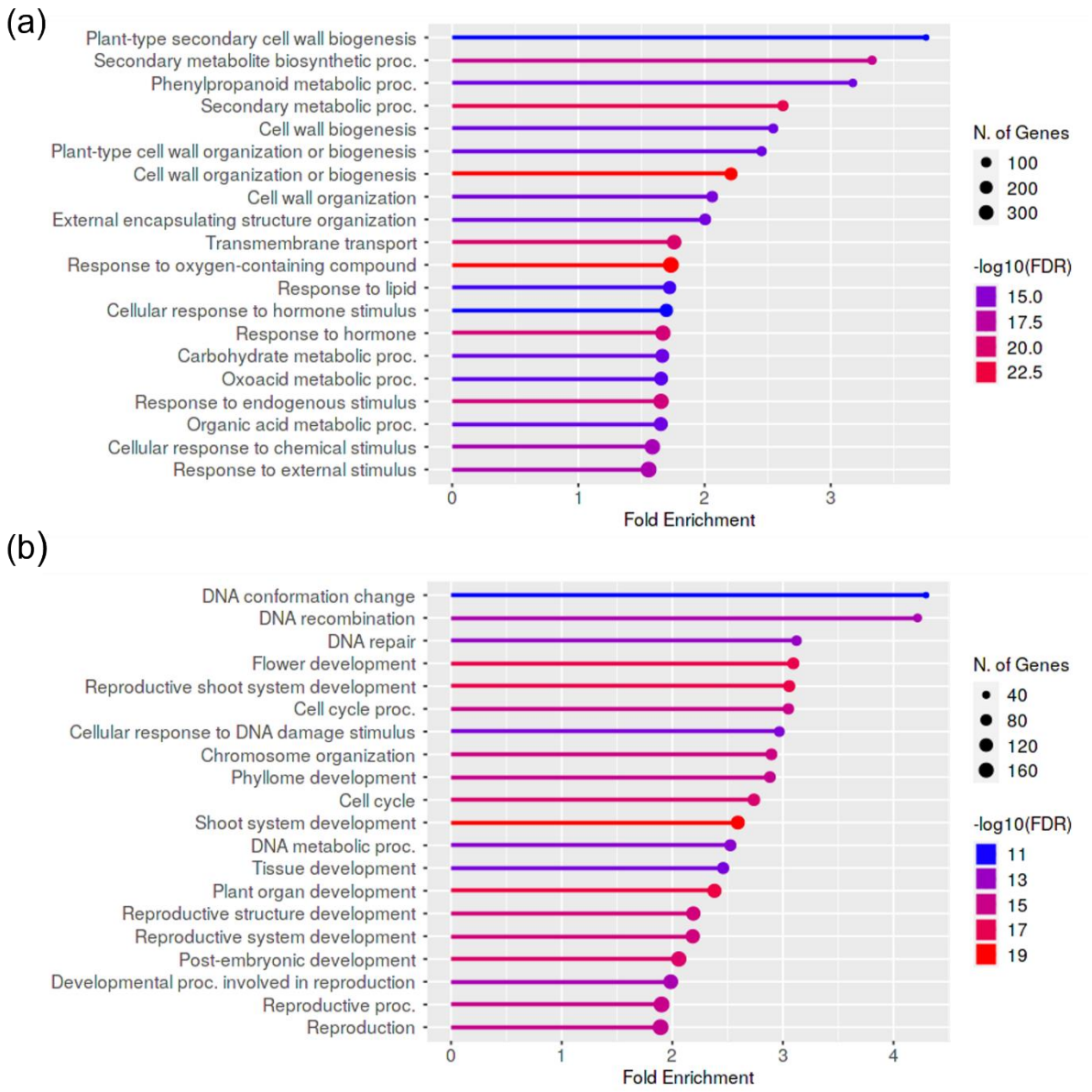

Figure 6.4: GO term enrichment for genes (a) upregulated (b) downregulated, in 3-week treatment compared to 12-week treatment using 21d pre-vernalisation samples

This was not as expected for plants which should have been grown under similar conditions and had not yet been exposed to the 5°C cold treatment. The genes which were upregulated before the 3-week vernalisation treatment were mainly metabolic processes (Figure 6.4a), while the downregulated genes compared to the recent 12-week experiment were mainly involved in cell cycle and growth (Figure 6.4b). The differences

identified from gene expression data suggest that the two datasets may not be reliably comparable. Samples seem to be developmentally different at a timepoint where they should have been similar, as they had not yet received any difference in treatment. This may be due to differences in plant care such as changes in soil composition or the nutrients available to the plant.

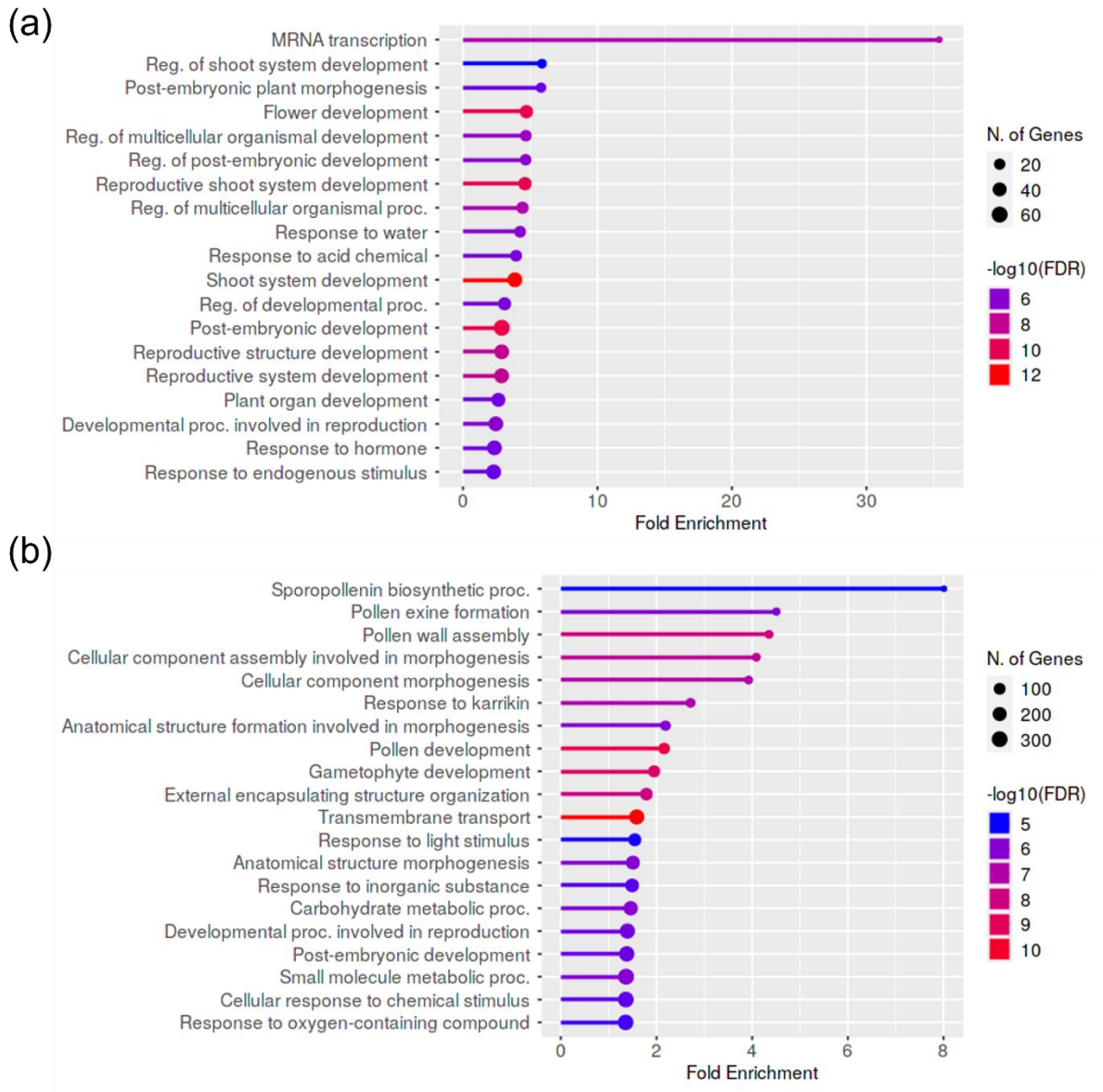

Figure 6.5: GO term enrichment for genes (a) upregulated (b) downregulated, in 3-week treatment compared to 12-week treatment when the apex reached growth stage BBCH51.

Comparison of BBCH51 samples revealed that in the plants which had been vernalised at 5°C for 3 weeks, there were 572 upregulated genes and 5422 downregulated genes compared to samples which reached BBCH51 after 12 weeks at 5°C. Upregulated genes in the 3-week treatment included in particular a large increase in a relatively low number of genes related to mRNA transcription (Figure 6.5a). GO term analysis of downregulated genes showed enrichment in pollen and gametophyte development (Figure 6.5b). Additional enriched GO terms included those related to floral and reproductive development. These differences suggest that samples from the 12-week vernalisation treatment are at meiosis while samples from the 3-week vernalisation are not, and so BBCH51 the developmental stage of flowers within the buds may not be the same between experiments.

When differential gene expression was compared between gynoecia of plants vernalised for either 3 or 12 weeks, it was found that 159 genes were upregulated in the 3-week cold treatment, with 197 genes were downregulated in the 3-week cold treatment compared to the 12-week treatment. GO term enrichment revealed that genes which were upregulated in gynoecia were enriched for responses to jasmonate signalling (Figure 6.6a). Genes which were downregulated included mainly biosynthetic and metabolic pathways, with particularly large fold-changes for a small number of genes involved in thiosulfate, succinate, and oxaloacetate transport (Figure 6.6b). This GO term analysis suggests that while vernalisation duration or floral development temperature affects gene expression in gynoecia, developmental pathways are not strongly affected.

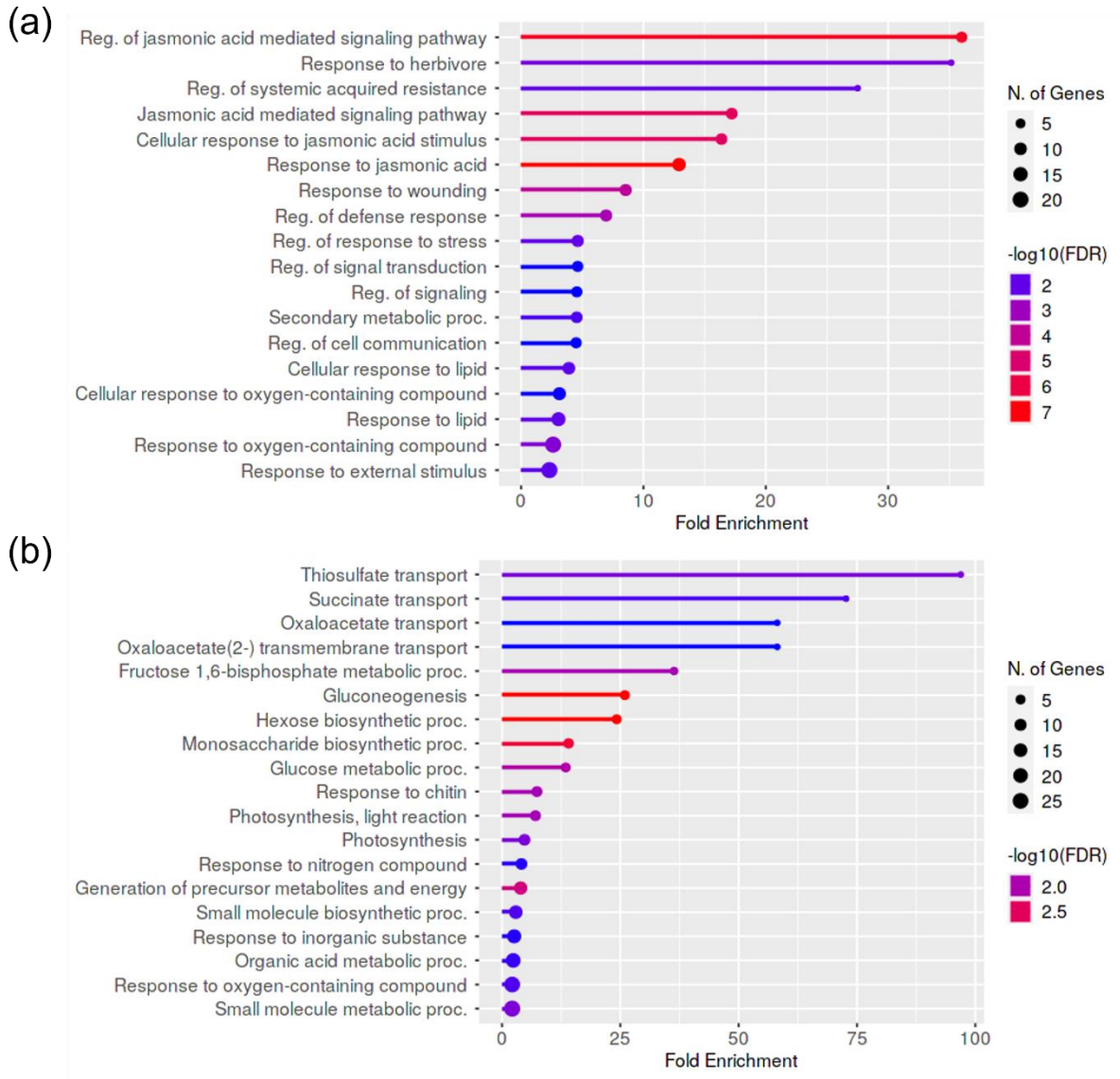

Figure 6.6: GO term enrichment for genes (a) upregulated (b) downregulated, in 3-week treatment compared to 12-week treatment using gynoecia samples

Additionally, the *IKU1* gene was found to be upregulated in gynoecia which experienced 12 weeks of cold compared to 3 weeks. *IKU1* was not expressed in gynoecia after the 3-week cold treatment but was expressed in gynoecia after the 12-week cold treatment (Figure 6.7). However, no other genes in the HAIKU pathway were seen to be differentially regulated. This indicates that while the duration of vernalisation or

temperature during floral development did not lead to overall altered expression of genes in the HAIKU seed size pathway in Zhongshuang 11.

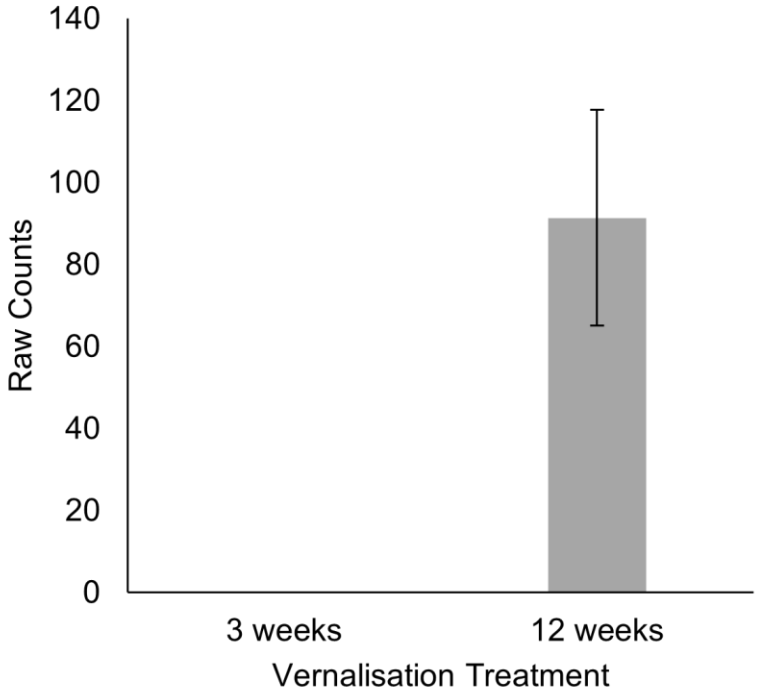

Figure 6.7: Raw count data from RNAseq showing expression of IKU1 in gynoecia of Zhongshuang 11. Error bars show standard error.

6.3.2 Quantifying the effects of vernalisation temperature in a Winter OSR variety

Analysis of results from the previous glasshouse phenotyping experiment (Chapter 2) indicated that TGW in Semi-winter OSR varieties (including Zhongshuang 11) was not significantly affected by vernalisation temperature. Following this, rather than continue to analyse a Semi-winter OSR variety another transcriptomic study was carried out to attempt to uncover the mechanism by which vernalisation affects seed size. Results from Chapter 2 of this thesis showed that increased vernalisation temperature led to smaller seed size in Winter OSR. A Winter OSR variety, Catana, showed a decrease in seed size after vernalisation at 15°C compared to 5°C. Catana was selected for this experiment as it had been previously shown to be robustly affected by temperature (Chapter 2.3.8).

Development was compared between Catana plants which experienced the same temperatures pre- and post- vernalisation, but which were vernalised separately at either 5°C or 15°C for 6 weeks (Figure 6.8). Tissue was sampled for transcriptomic analysis at five timepoints. Shoot apex samples were taken after 1 week of vernalisation; after 6 weeks of vernalisation; and at the floral transition. Samples were also taken of the gynoecia and of the whole seed 14 days after pollination (DAP).

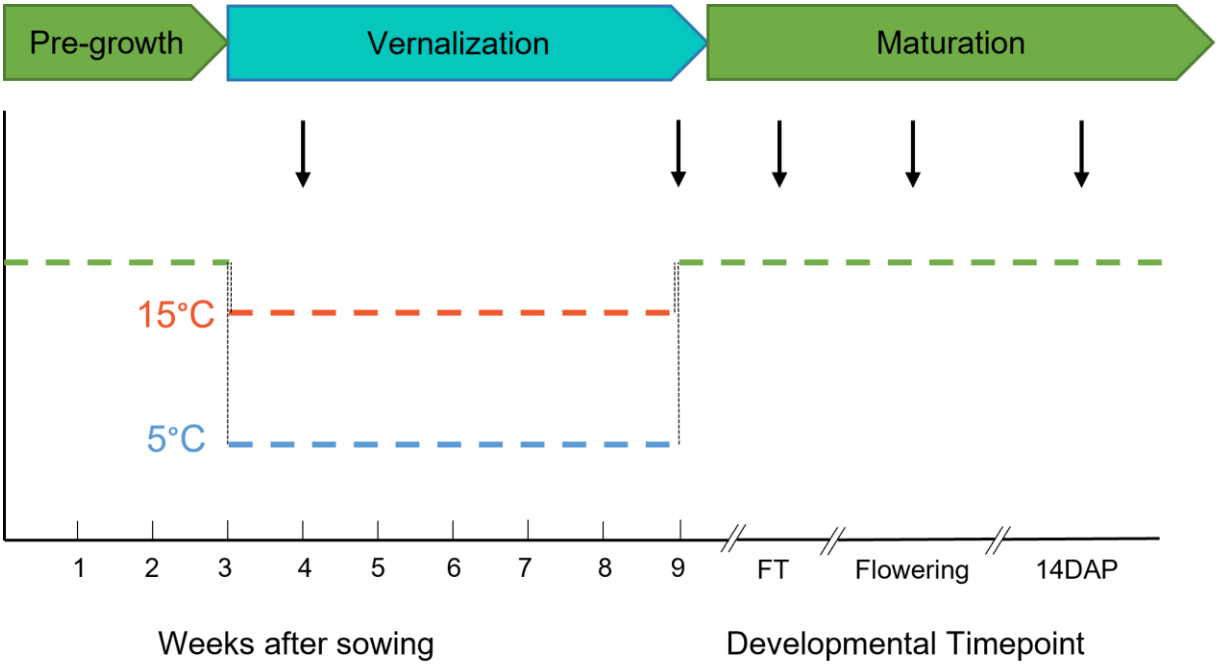

Figure 6.8: Experimental setup for comparing development between vernalisation treatments at 5°C and 15°C in Catana. Black arrows indicate sampling timepoints for transcriptome analysis. FT = floral transition; DAP = days after pollination.

6.3.2.1 Effects of different vernalisation temperatures on ovule and seed development

Differences in development between plants vernalised at 5°C and 15°C were also compared using microscopy, with same experimental setup described in Figure 6.8. Samples of the gynoecia were taken from freshly opened flowers to compare ovule size between plants vernalised at different temperatures. Seed samples were also taken at

14DAP and 21DAP to compare seed size and embryo development at these stages. This experiment aimed to reveal when developmental changes leading to differences in seed size between vernalisation temperatures occur.

Gynoecia were sampled from freshly opened flowers after plants had been removed from the vernalisation treatment and grown at 18°C. Ovule size was measured from the gynoecia and compared between plants vernalised at 5°C and 15°C.

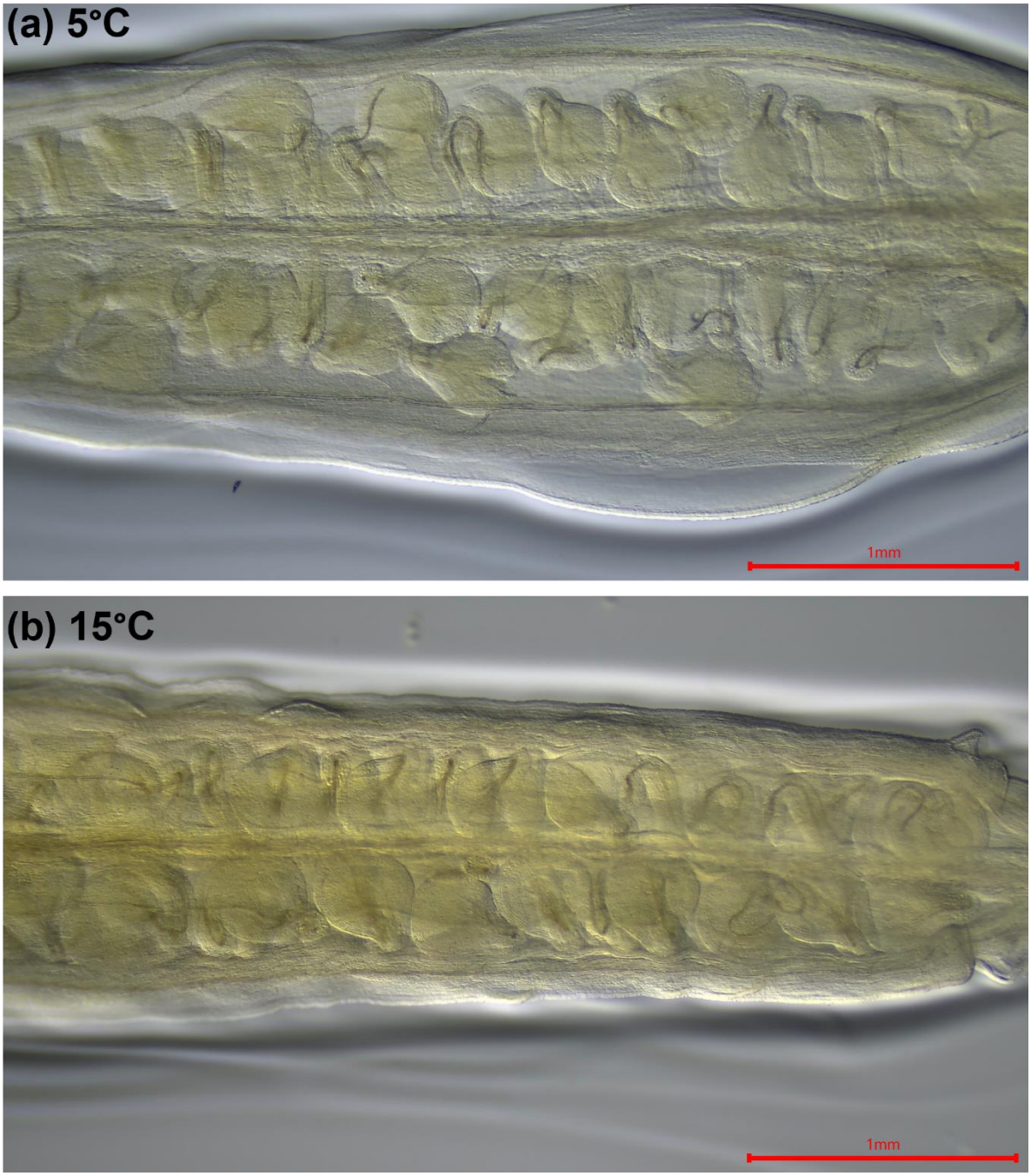

Figure 6.9: Gynoecia dissections from Catana plants vernalised at (a) 5°C (b) 15°C

Images of the gynoecia (Figure 6.9) show that ovule size was larger in gynoecia from plants vernalised at 15°C. Statistical analysis confirmed that ovules were significantly larger in plants which had been vernalised at 15°C (Figure 6.10), with an increase in

mean ovule area of 0.0097mm^2, representing an increase in ovule area of 10% of the mean at 5°C.

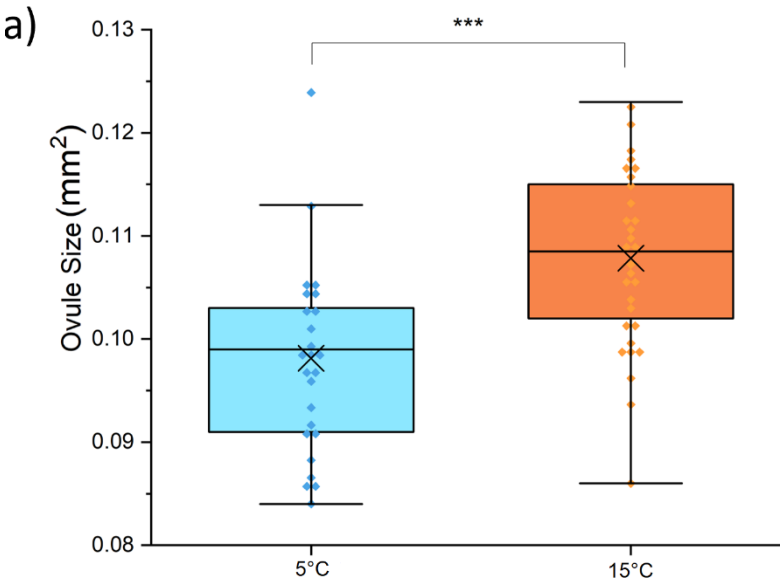

Figure 6.10: Area of ovules in mm^2 from open flowers after two different vernalisation treatments. Asterisks used to indicate significance: (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$. P-values calculated using Student's t-test, $p = 0.009$. For 5°C, n = 25; for 15°C, n = 30.

Embryos were dissected from seeds of plants grown at both vernalisation temperatures, after plants had been removed from the vernalisation treatment and grown at 18°C. Embryos from plants vernalised at 15°C were less well developed than those from plants vernalised at 5°C at 21DAP (Figure 6.11). In seeds of plants vernalised at 5°C, embryos had clearly reached the bent cotyledon stage of development (Figure 6.11a-b). In seeds of plants vernalised at 15°C, some embryos had only just begun to progress from torpedo stage to bent cotyledon (Figure 6.11c), while others had progressed clearly to bent cotyledon stage but appeared smaller than embryos from the 5°C vernalisation treatment (Figure 6.11d).

A similar trend was seen when examining the seeds at 21DAP under the microscope, with seeds from plants vernalised at 5°C appearing to be larger than those from plants vernalised at 15°C. Analysis of images revealed that at 21DAP, the seed area was

significantly larger in seeds vernalised at 5°C compared to at 15°C with an increase in mean seed area of 0.59mm^2 (Figure 6.12). At 14DAP there was no significant difference in seed area between the two vernalisation treatments. These results indicate that vernalisation at 15°C leads to smaller seeds than vernalisation at 5°C, but that this effect only becomes apparent after 14DAP as the embryo develops from torpedo to bent cotyledon stage.

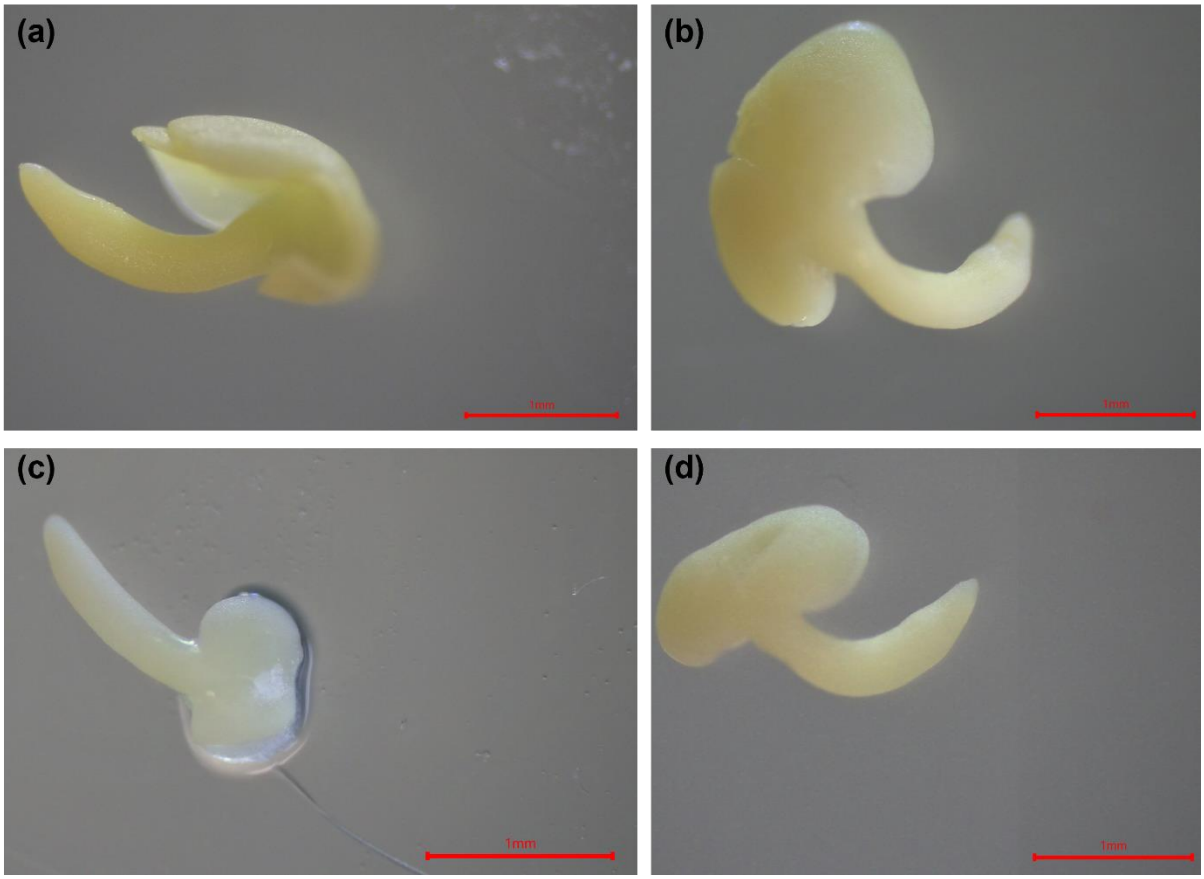

Figure 6.11: Embryos from Catana plants vernalised at (a - b) 5°C, (c - d) 15°C. Scale bars are 1mm. Embryos were dissected from seeds at 21DAP.

Before fertilization, the warmer vernalisation temperature is correlated with larger ovules. However, the larger ovules do not go on to produce larger seeds. By 14DAP, the difference in area has disappeared and by 21DAP the seeds from plants vernalised at 5°C are larger. It seems that 5°C seeds have a size advantage that arises after fertilization and continues at least through the first 21 days of seed development. As the

mature seeds are also larger in plants vernalised at 5°C, it can be presumed that this trend continues through the rest of seed development and maturation.

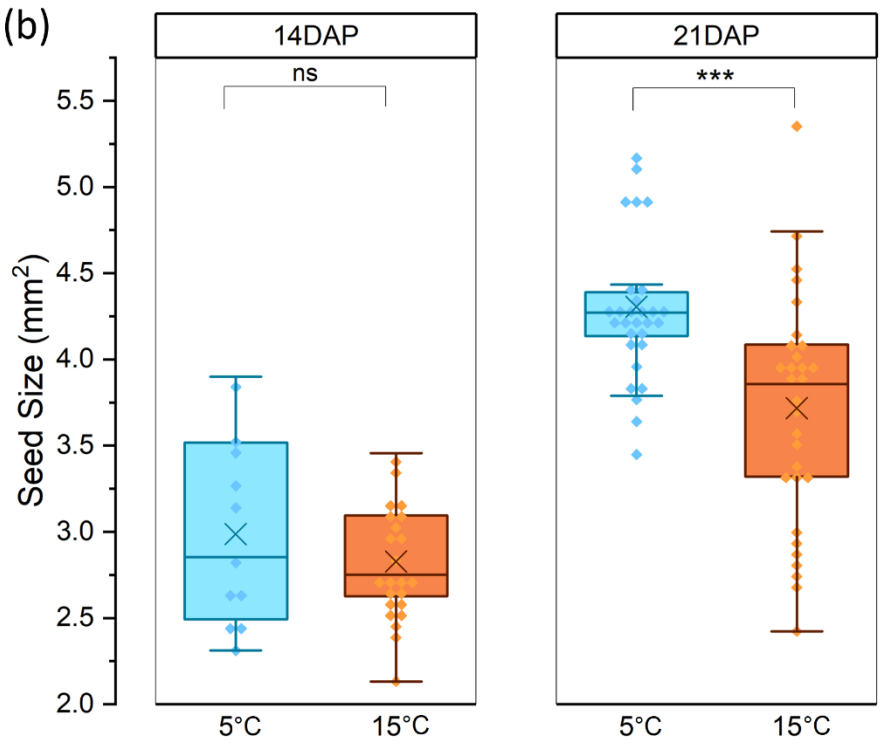

Figure 6.12: Area of seeds in mm2 after two different vernalisation treatments. Asterisks indicate significance: (*) p < 0.05, (**) p < 0.01, (***) p < 0.001. P-values calculated using Student's t-test. At 14DAP, n (5°C) = 11; n (15°C) = 23; p = 0.371. At 21DAP, n (5°C) = 29; n (15°C) = 30; p =0.000172.

6.3.2.2 Comparing effects of vernalisation on gene expression through development

To identify differences in gene expression using a developmental time-series, samples were taken for transcriptomic analysis during vernalisation and during reproductive development. In vernalisation, samples were taken of the shoot apex after one week and six weeks of vernalisation. This allows for direct comparisons based on how long the plant has been in vernalisation.

Additionally, comparisons could be made between samples with equal amounts of chilling time and thermal time (Figure 6.13). Plants were grown at 17°C for 21 days preceding the vernalisation treatment, meaning that thermal time but not chilling time was accumulated during the pre-growth period. Figure 12 shows that plants vernalised at 5°C took one week to accumulate 84 chilling units, while plants a 15°C took six weeks to accumulate 84 chilling units. For thermal units, plants vernalised at 5°C had accumulated 378 thermal units by day 63, after six weeks in vernalisation; plants vernalised at 15°C accumulated 378 thermal units by 28 days, after one week in vernalisation.

Thermal time was defined using a base temperature of 3°C (Hodgson et al, 1978; Habekotté et al, 1997).

$$\text{Thermal time} = (\text{Temp} - 3°\text{C}) \times \text{days}$$

Chilling time was defined using a maximum vernalisation temperature of 17°C (Tommey and Evans, 1991; Habekotté et al, 1997).

$$\text{Chilling time} = (17°\text{C} - \text{Temp}) \times \text{days}$$

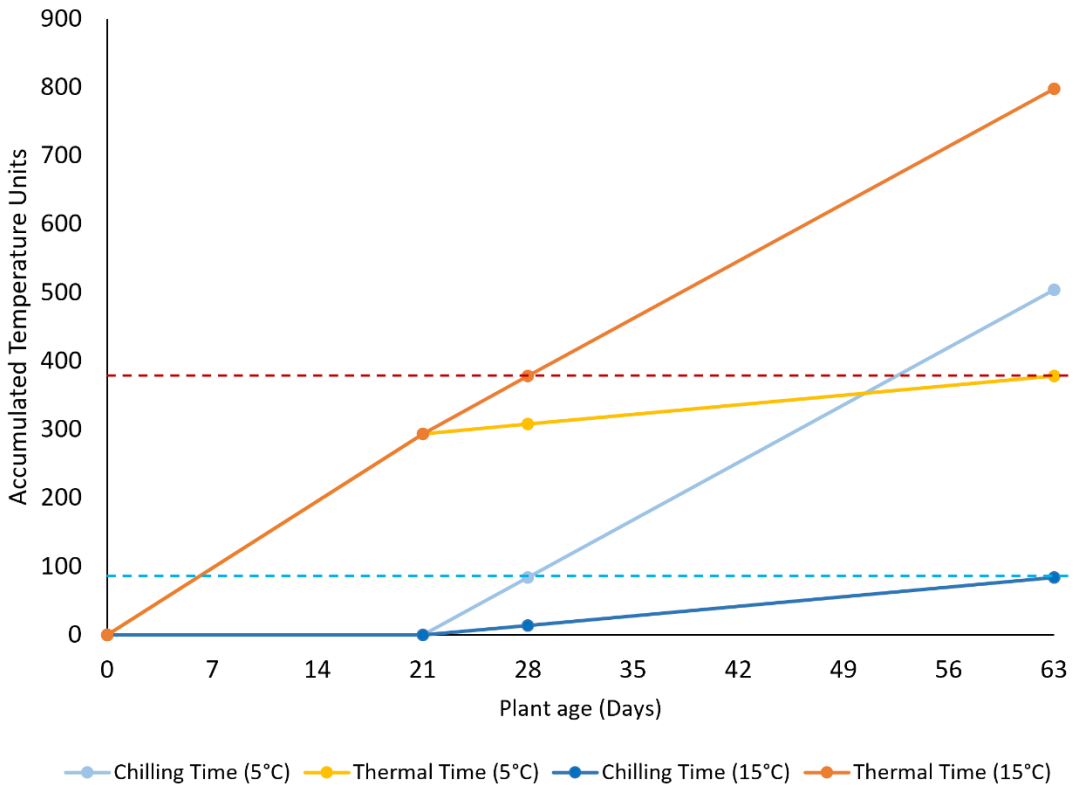

Figure 6.13: Graph showing accumulation of thermal and chilling time for plants vernalised at 5°C or 15°C. Vernalisation treatment began at 21 days. Red dashed line indicates equivalent thermal time between samples (378 thermal units). Blue dashed line indicates equivalent chilling time (84 chilling units).

Table 6.1: Number of differentially regulated genes for each transcriptome comparison. Upregulated and downregulated refers to fold-change compared to the 5°C treatment. DEGs = differentially expressed genes.

Comparison	***Total DEGs***	**Upregulated genes**	**Downregulated genes**
Shoot apex - 1 week 5°C vs 6 weeks 15°C	*10577*	5611	4966
Shoot apex - 6 weeks 5°C vs 1 week 15°C	*6113*	1842	4271
Shoot apex - 1 week vern, 5°C vs 15°C	*7511*	3596	3915
Shoot apex - 6 weeks vern, 5°C vs 15°C	*5215*	1966	3249
Shoot apex - Floral transition, 5°C vs 15°C	*8228*	5116	3112
Gynoecia, 5°C vs 15°C	*133*	94	39
14 DAP seeds, 5°C vs 15°C	*527*	243	284

The comparison with the least differentially expressed genes was the gynoecia after either 5°C or 15°C vernalisation, while the comparison with the greatest number of differentially expressed genes was the comparison between plants with equal chilling time, comparing the shoot apex 1 week at 5°C against 6 weeks at 15°C (Table 6.1). These results also suggest that gene expression in the gynoecia is not strongly affected by vernalisation temperature.

Clustering of all DEGs identifies some patterns in the overall gene expression (Figure 6.14). The apex samples from 15°C are more similar to each other than to the apex samples from 5°C at either the 1-week or 6-week timepoint. There is also a notable cluster of genes which are switched on in the seeds.

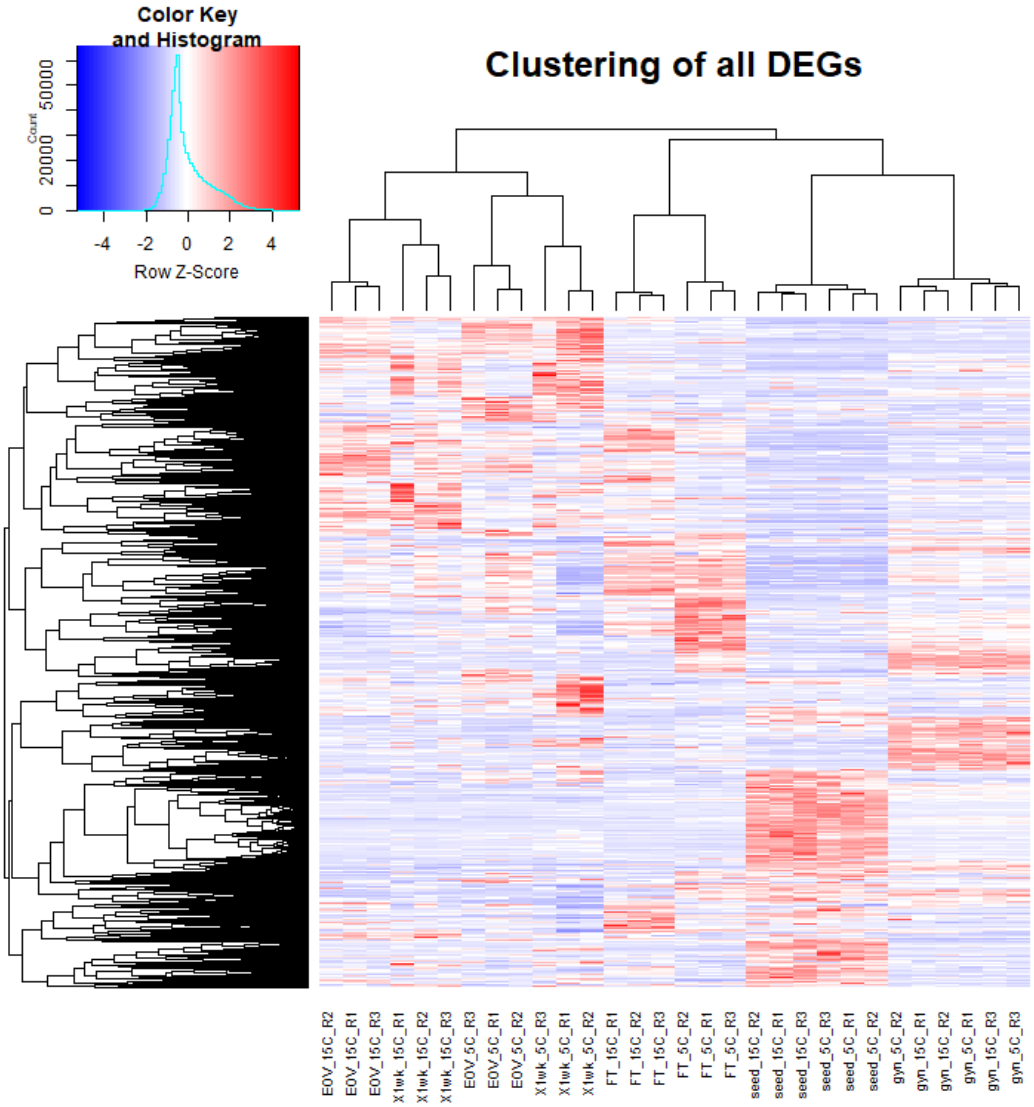

Figure 6.14: Clustering heatmap for 22,540 differentially expressed genes across all timepoints.

6.3.2.3 Differences in gene expression between two vernalisation temperatures

Comparing gene expression in the shoot apex between plants which had received equal amounts of chilling hours at either 5°C or 15°C yielded the largest amount of DEGs in this experiment (10,577). In comparison, comparing plants with equal thermal time produced 6,113 DEGs.

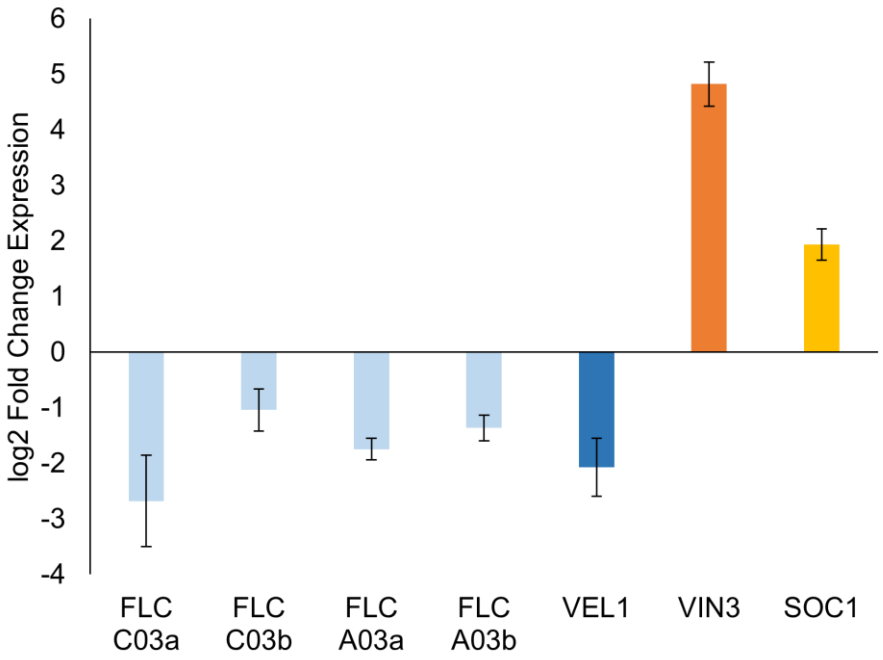

Figure 6.15: Log fold-change for gene expression in equivalent chilling time samples, relative to samples from 5°C vernalisation treatment. Error bars show standard error.

For chilling time, the A03 and C03 (BnaC03g04170D, BnaC03g16530D, BnaA03g02820D, BnaA03g13630D) copies of *FLC* were differentially regulated, with higher expression in plants vernalizing at 15°C for 6 weeks compared to plants vernalizing at 5°C for 1 week (Figure 6.15). *VEL1* was also upregulated in plants vernalizing at 15°C. In plants vernalised at 5°C, higher expression of *VIN3* and *SOC1* was found. Genes included in this graphical analysis were based on known vernalization-related and seed development genes alongside candidate gene selection.

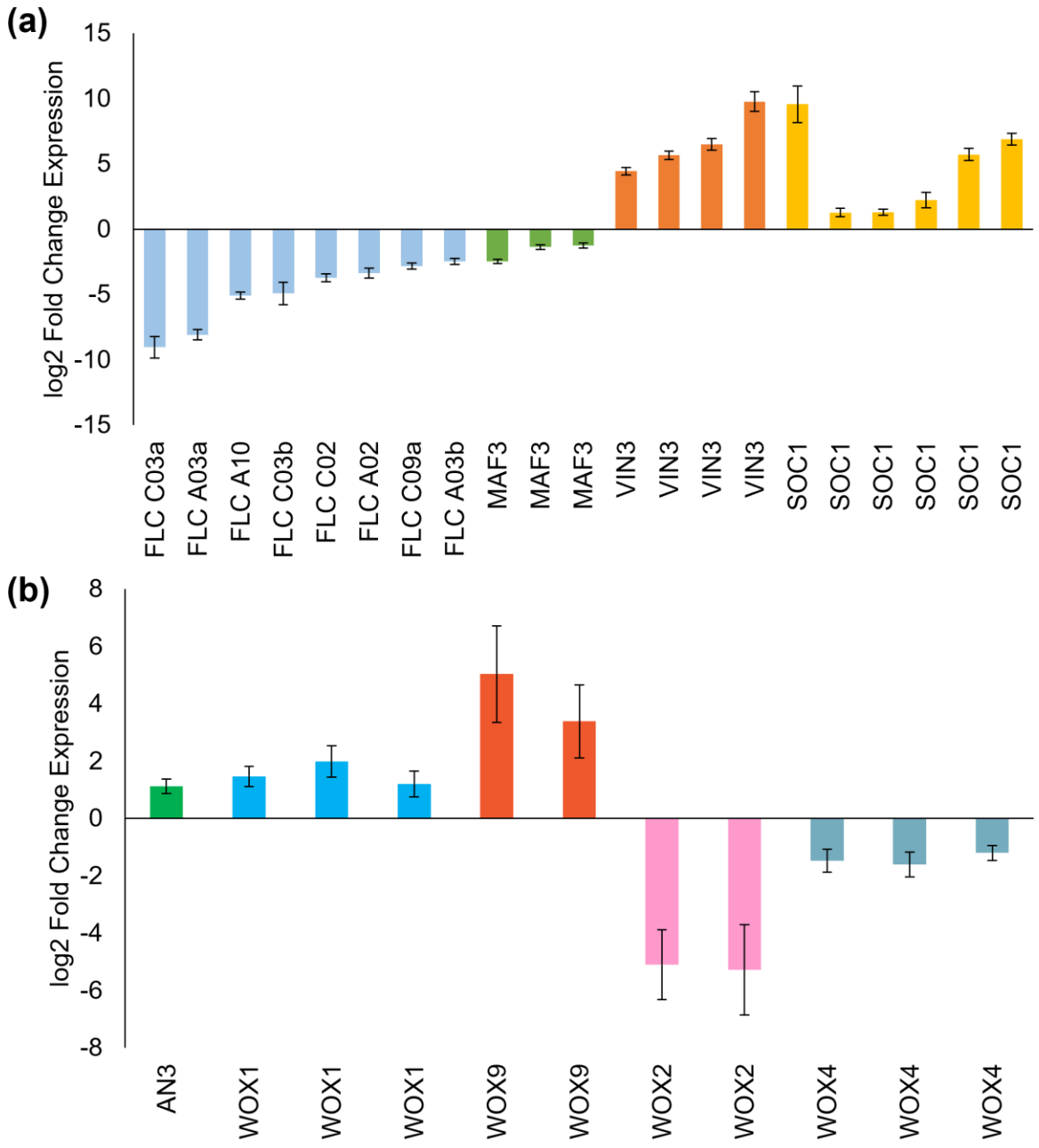

Figure 6.16: Log fold-change for gene expression in equivalent thermal time samples, relative to samples from 5°C vernalisation treatment. Error bars show standard error. (a) shows vernalisation pathway genes (b) shows AN3 and WOX organ identity genes.

In comparing plants with equal thermal time, shoot apex samples from plants vernalised at 5°C for 6 weeks were compared to plants vernalised at 15°C for 1 week, meaning that 5°C plants had a substantially higher amount of accumulated chilling time as well as a longer overall time period in the vernalisation treatment. Eight *FLC* copies

were upregulated in plants vernalizing at 15°C (Figure 6.16a). Additionally, three copies of *MAF3* were upregulated (BnaC02g43420D, BnaC02g43470D, BnaA02g34510D). In plants which were vernalizing at 5°C for six weeks, five copies of *VIN3* were found to be upregulated (BnaC03g12980D, BnaA03g10310D, BnaA02g08140D, BnaC02g11530D) along with six copies of *SOC1* (BnaA03g56880D, BnaC04g50370D, BnaCnng36880D, BnaA04g26320D, BnaA05g05010D, BnaC04g53290D). These differences in vernalisation response genes are essentially what would be expected when comparing a strongly vernalised plant to one which has not yet been vernalised.

Additionally, one copy of *AN3* was found to be upregulated in the 5°C shoot apex (BnaC02g39370D). This was accompanied by upregulation of several copies of *WOX1* and *WOX9*; meanwhile, in the same plants several copies of *WOX2* and *WOX4* were downregulated (Figure 6.16b). *AN3* is involved in maintenance of the shoot apical meristem and functions with *WOX* transcriptional regulators to control floral organ development (Zhang et al, 2019), suggesting that floral development is at different stages between samples with different thermal time.

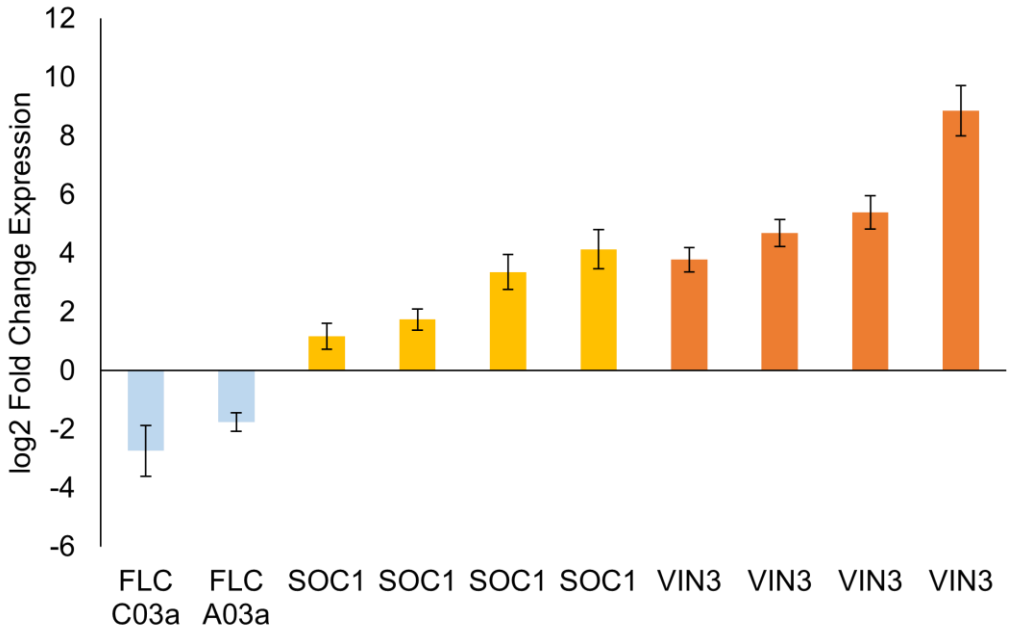

Figure 6.17: Log fold-change for gene expression in samples taken after 1 week of vernalisation, relative to samples from 5°C vernalisation treatment. Error bars show standard error.

When comparing apex samples taken after 1 week of vernalisation at either 5°C or 15°C, reduced levels of *FLC* in 5°C plants alongside increased levels of *SOC1* and *VIN3* indicates that the vernalisation process is occurring more quickly at 5°C (Figure 6.17). Two *FLC* genes were more highly expressed in plants vernalised at 15°C (BnaC03g04170D, BnaA03g02820D). In plants vernalizing at 5°C, *SOC1* (BnaCnng36880D, BnaC04g50370D, BnaA05g05010D, BnaC04g53290D) and *VIN3* (BnaC03g12980D, BnaA03g10310D, BnaA02g08140D, BnaC02g11530D) were both more highly expressed compared to 15°C plants.

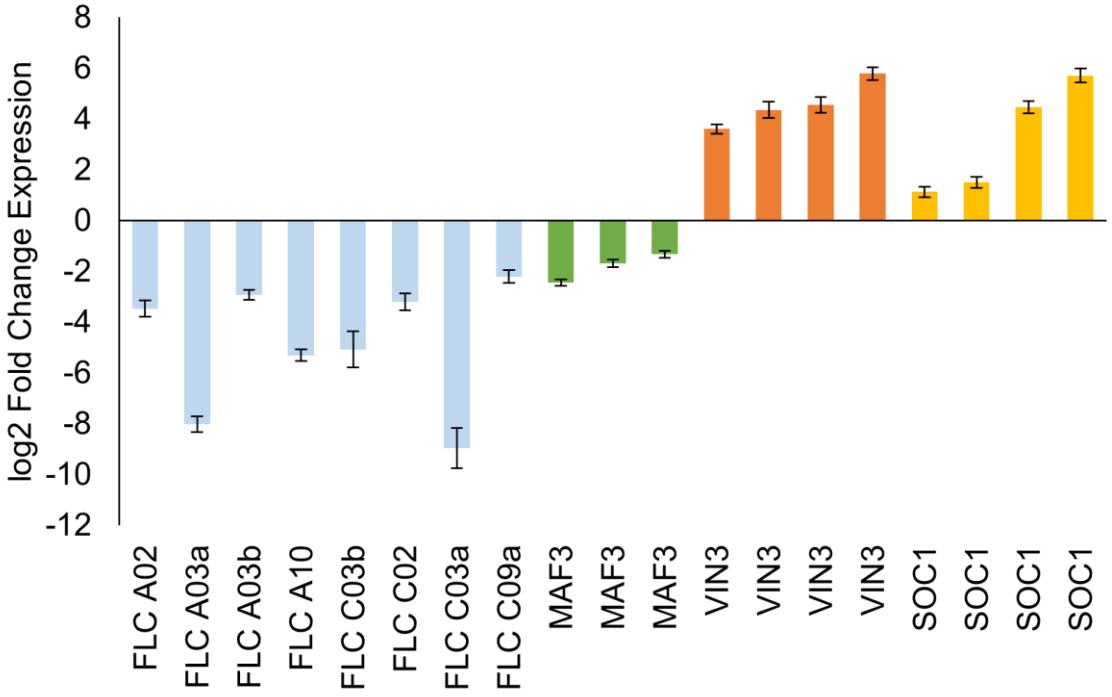

Figure 6.18: Log fold-change for gene expression in samples taken after 6 weeks of vernalisation, relative to samples from 5°C vernalisation treatment. Error bars show standard error.

By the end of the six-week vernalisation treatment, the difference in vernalisation related genes has increased. Eight copies of *FLC* were more highly expressed in plants vernalised at 15°C (Figure 6.18); as well as *FLC*, multiple copies of *MAF3, VIN3* and *SOC1* were differentially regulated at this point in development.

The result that plants are less deeply vernalised at the end of the 15°C 6-week vernalisation treatment compared to after six weeks of vernalisation at 5°C is logical. However, even when taking chilling time into account, plants which have been

vernalised at 5°C show lower expression levels of *FLC*. The differential gene expression patterns of vernalisation response genes seen in the chilling time comparison indicates that even when chilling time is equal, plants will be more strongly vernalised at 5°C than at 15°C.

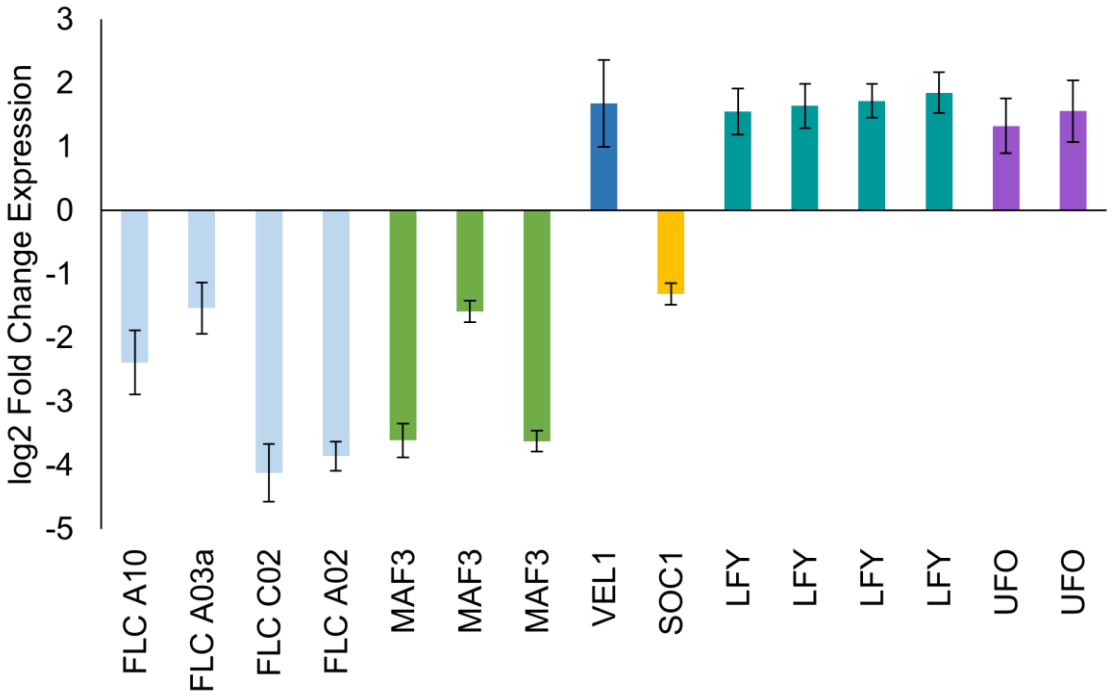

Figure 6.19: Log fold-change for gene expression in the shoot apex at the floral transition, relative to samples from 5°C vernalisation treatment. Error bars show standard error.

At the point of the floral transition there were 8,288 differentially expressed genes. This included four copies of *FLC* (BnaA10g22080D, BnaA03g02820D. BnaC02g00490D, BnaA02g00370D), three copies of *MAF3* (BnaA06g23990D, BnaC02g43420D, BnaA02g34510D), *VEL1* (BnaA06g25100D) and *SOC1* (BnaA04g26320D); all were found to be more highly expressed in plants vernalised at 15°C compared to 5°C (Figure 6.19). In plants which had been vernalised at 5°C, there was found to be increased expression of four copies of *LFY* (BnaCnng78420D, BnaA06g39070D, BnaCnng24550D, BnaCnng78500D) and two copies of *UFO* (BnaC08g09370D, BnaAnng11240D). These genes positively influence flower development (Wilkinson and Haughn, 1991; Lee et al, 1997), suggesting that a 5°C vernalisation temperature encourages more rapid floral

development more than a 15°C vernalisation treatment even at a comparable developmental stage.

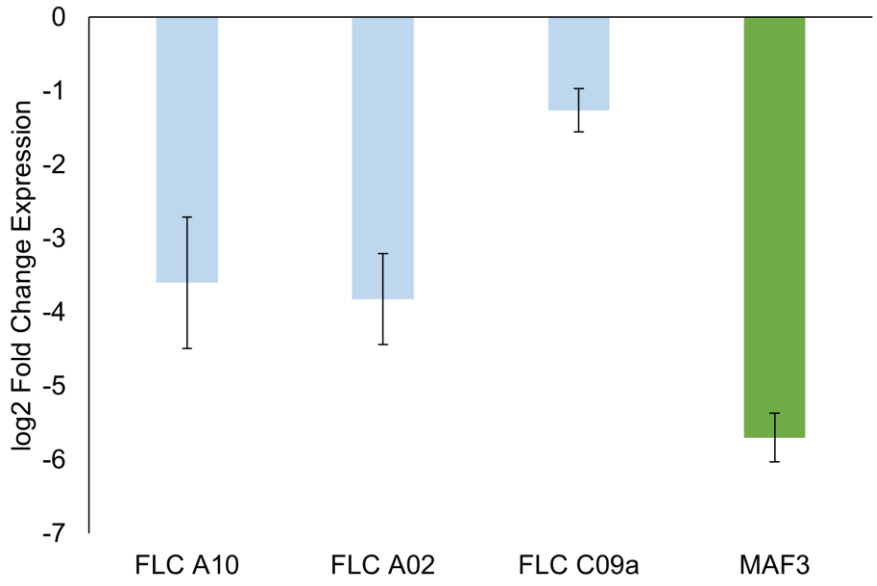

Figure 6.20: Log fold-change for gene expression in gynoecia, relative to samples from 5°C vernalisation treatment. Error bars show standard error.

In gynoecia, only 133 genes were differentially regulated when comparing between the vernalisation temperatures. Several *FLC* genes as well as *MAF3* were upregulated in gynoecia from plants vernalised at 15°C (Figure 6.20). This highlights that differences in vernalisation pathway genes are still differentially expressed during reproductive development after the vernalisation treatment has ended.

6.3.2.4 Optimizing of clustering for seed DEGs based on expression patterns

This experiment found a total of 527 differentially expressed genes in seeds at 14 days after pollination (DAP) between plants from the two vernalisation treatments. To understand how and when these genes respond to vernalisation, they were clustered based on their expression profiles through development. Genes were clustered separately based on whether they were upregulated in 15°C vernalised samples or in 5°C vernalised seed samples, using a fuzzy C-means clustering algorithm. Fuzzy clustering allows for genes to belong to more than one cluster, which can be helpful for identifying

different patterns of regulation (Futschik and Carlisle, 2005). To split the group of genes into clusters effectively, the distance between cluster centroids should be low, balanced with preventing the number of clusters from being too large to contain useful information.

6.3.2.4.1 Clustering of genes upregulated in seeds from 15°C-vernalised plants

For genes upregulated in seeds produced by plants vernalised at 15°C, the best number of clusters using this method was identified to be eight (Figure 6.21).

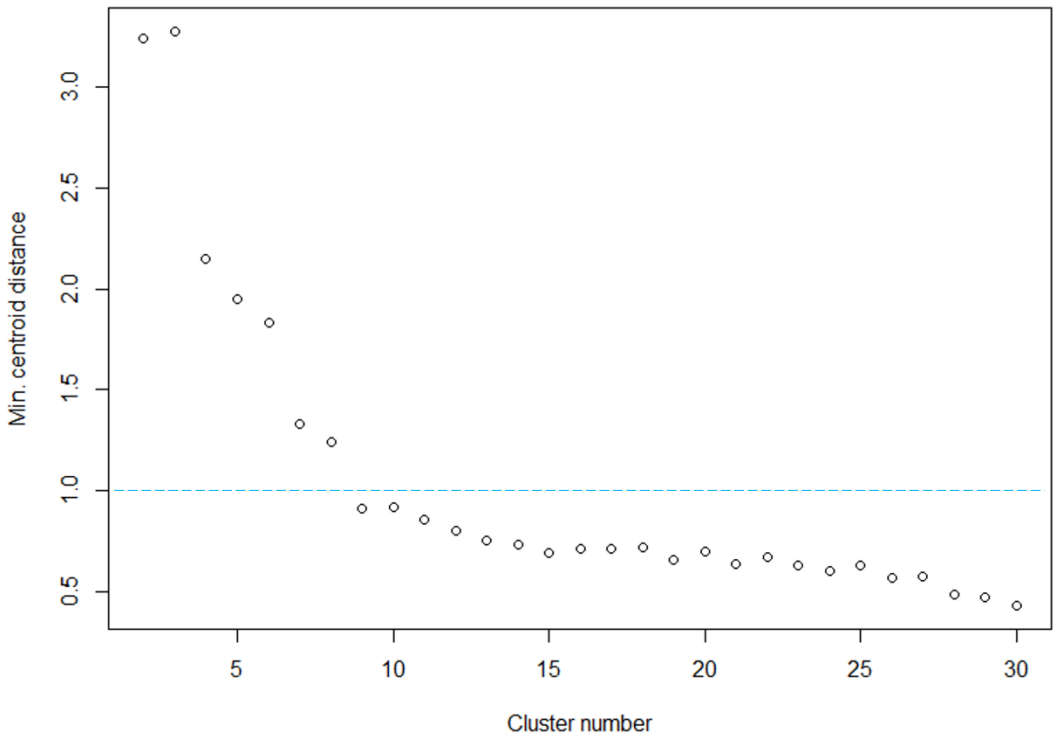

Figure 6.21: Minimum centroid distance for clustering genes upregulated in seeds of 15°C vernalised plants. Blue dashed line indicates best fitting minimum centroid distance.

In Figure 6.22, cohesion amongst gene expression within the clusters is indicated by colour, with reds indicating many genes following a similar expression pattern and greens indicating genes with more distinct gene expression patterns. Clusters 3 and 5 are most cohesive, containing genes that are expressed mainly in seed and gynoecia; cluster 4 appears to be the least cohesive, with varying gene expression mainly around

the floral transition stage. Clusters 1 and 2 have higher expression in seeds, with less cohesive expression at earlier timepoints particularly in Cluster 2. Clusters 7 and 8 indicate gene expression which falls across the sampling timepoints.

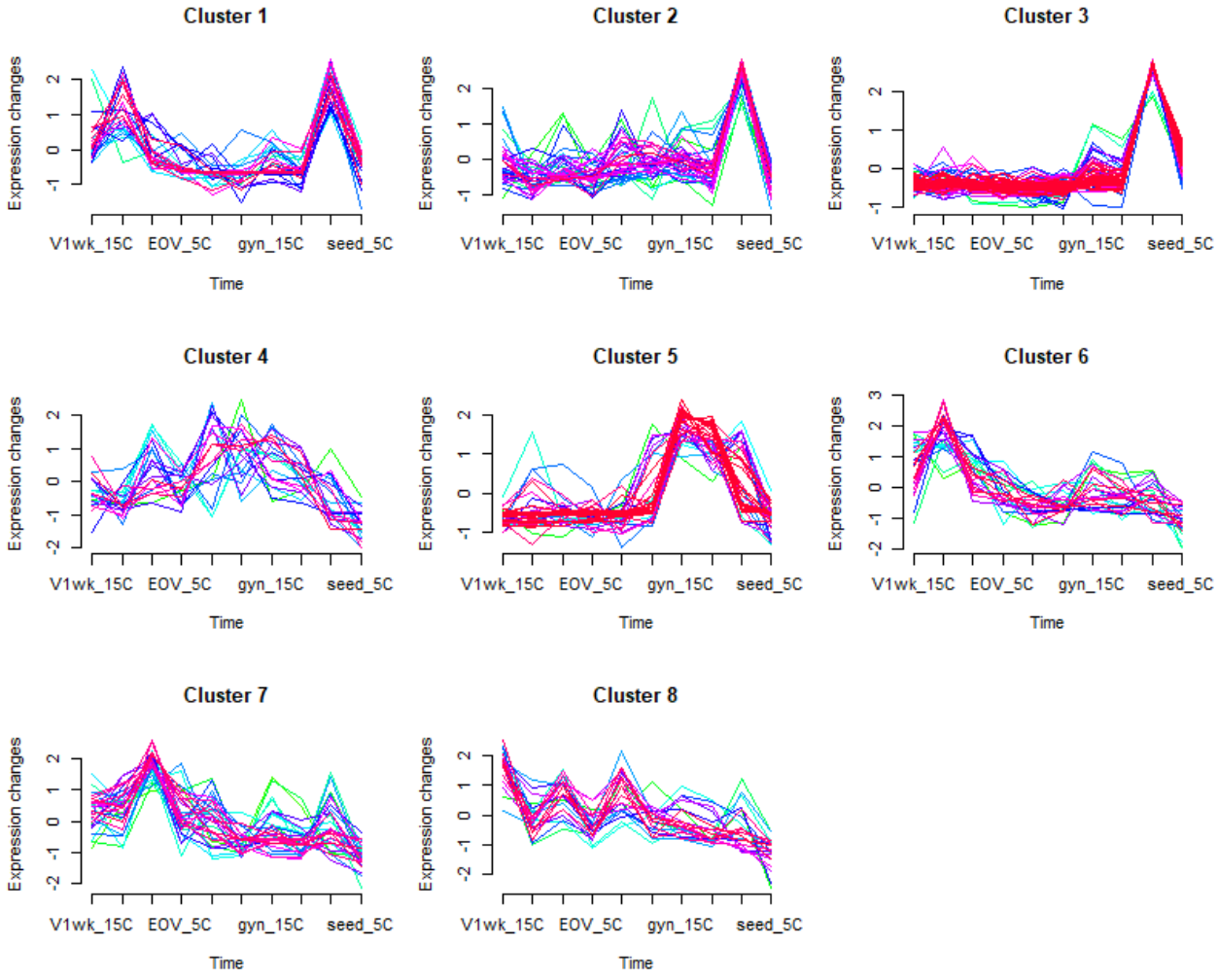

Figure 6.22: Clustering of eight groups of genes upregulated in seeds from 15°C vernalised plants, produced using fuzzy clustering. X-axis shows timepoints ordered by chronological timepoints and then vernalisation temperature, 15°C and 5°C.

6.3.2.4.2 Clustering of genes upregulated in seeds from 5°C-vernalised plants

For genes upregulated in 5°C-vernalised plants, Figure 6.23 shows that the minimum centroid distance did not decrease steadily with increasing cluster number. To optimize

the number of clusters selected, the cohesion of clusters was comparing using six, eleven, and thirteen clusters.

When splitting genes upregulated in 5°C seeds into six clusters, it seems that cluster 1 is highly cohesive, with genes that are mainly expressed in seeds (Figure 6.24). Cluster 4 is also highly cohesive, with a range of differential gene expression indicated across timepoints. Clusters 3 and 6 are the least cohesive, with genes in Cluster 6 generally decreasing in expression after the vernalisation treatment.

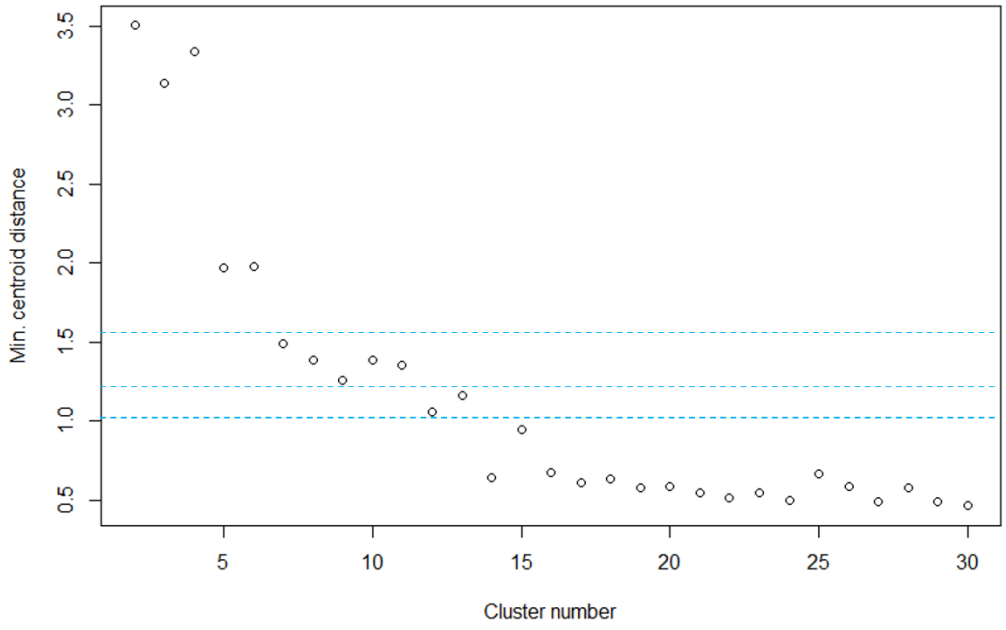

Figure 6.23: Minimum centroid distance between clusters for seed DEGs upregulated in 14DAP seeds after vernalisation at 5°C. Lines indicate possible cut-off points for number of clusters, from top to bottom: 6 clusters, 11 clusters or 13 clusters.

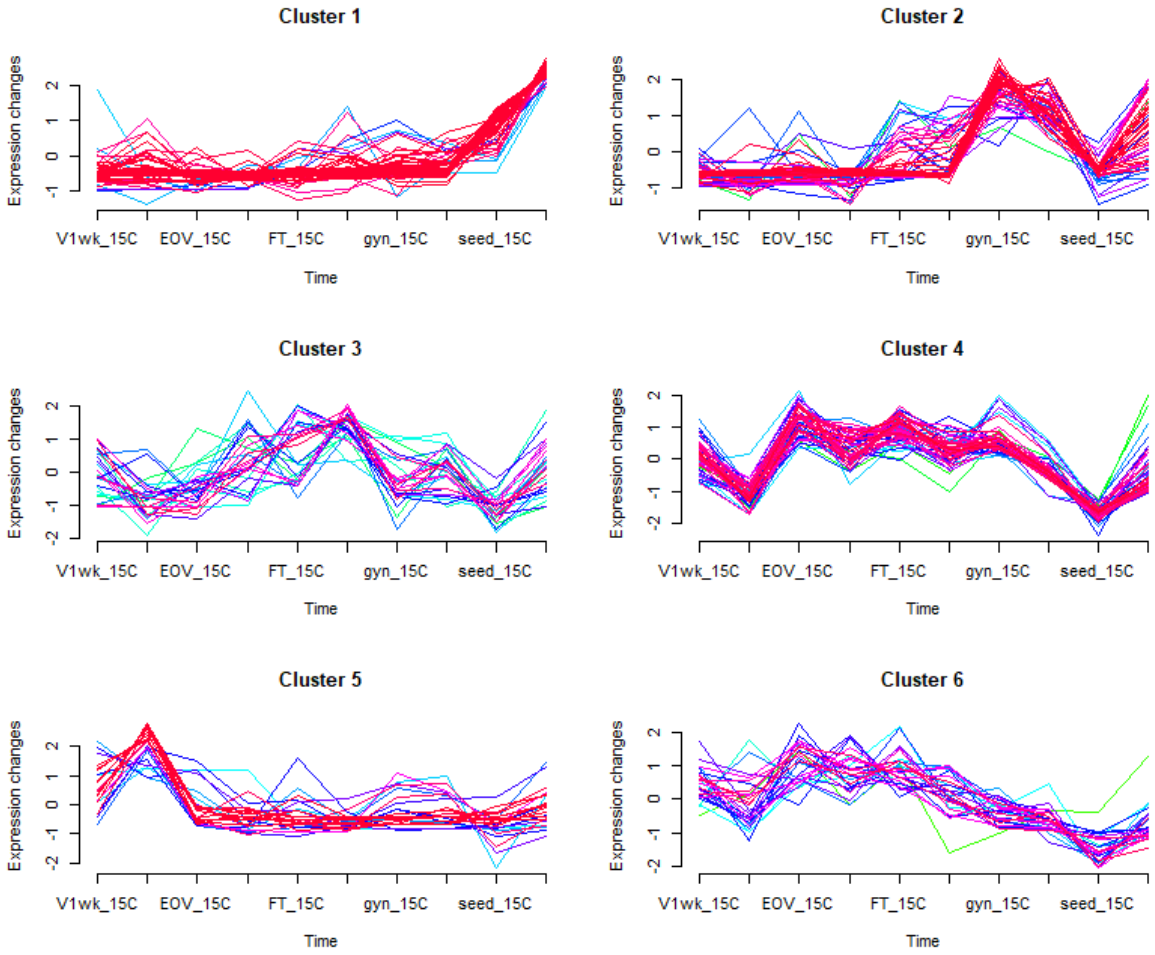

Figure 6.24: Clustering of genes upregulated in seeds of 5°C vernalised plants, with six clusters producing by fuzzy clustering. X-axis shows timepoints ordered by chronological timepoints and then vernalisation temperature, 15°C and 5°C.

When the number of clusters in increased to 11 (Figure 6.25), several of the clusters – particularly Clusters 5 and 6 – can be seen to have very few genes, with the differences in gene expression patterns between individual genes being apparent. Cluster 5 contained 7 genes while Cluster 6 contained 8.

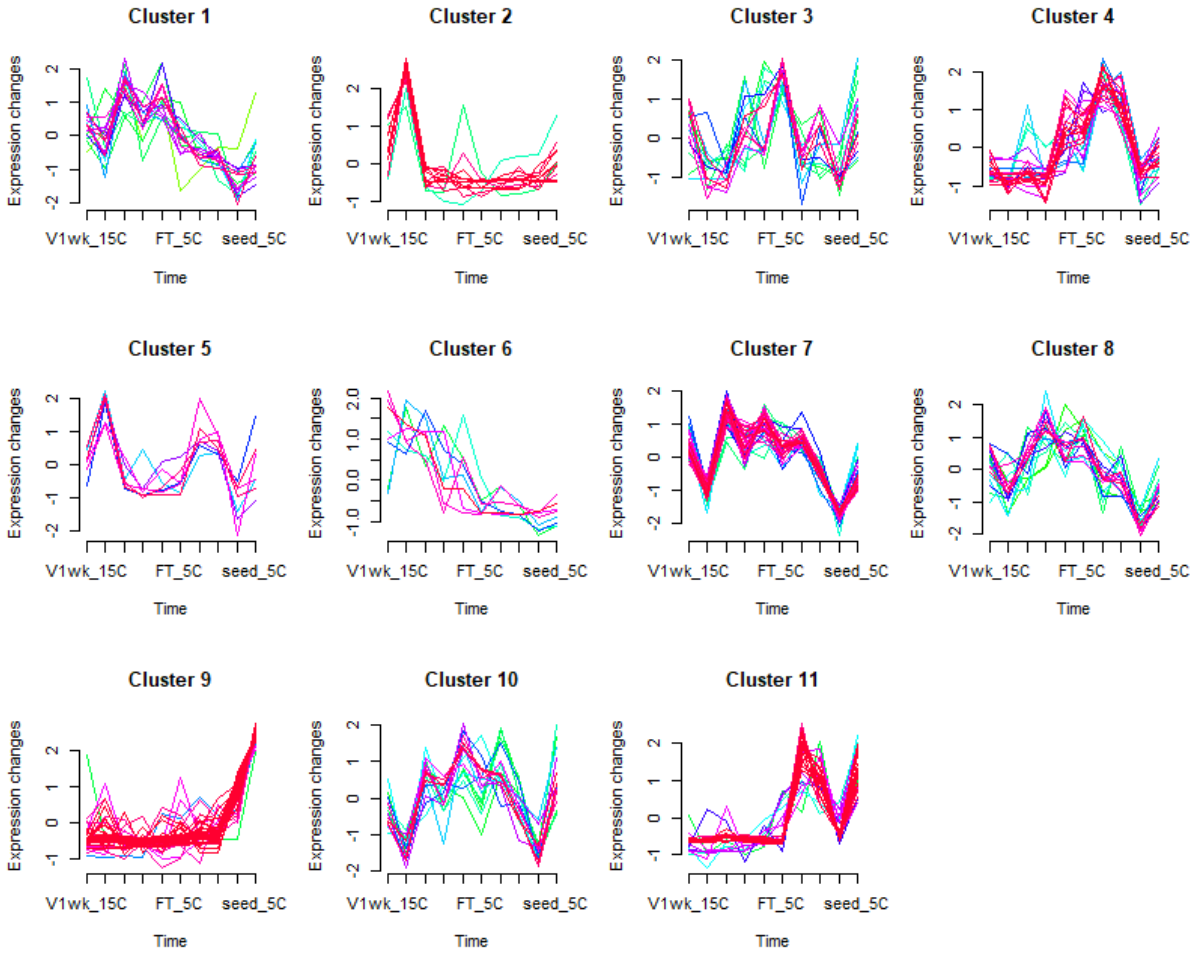

Figure 6.25: Clustering of genes upregulated in seeds of 5°C vernalised plants, with eleven clusters producing by fuzzy clustering. X-axis shows timepoints ordered by chronological timepoints and then vernalisation temperature, 15°C and 5°C.

Although increasing the number of clusters again to 13 (Figure 6.26) lowers the minimum centroid distance (as seen in Figure 6.23), some of the clusters have very few genes with more distinct expression patterns. Clusters 5 and 9 contains only 7 genes. The most cohesive clusters in this analysis have close analogues in the six-cluster analysis. Cluster 3 shown in Figure 6.19 corresponds closely in expression pattern to Cluster 1, while Cluster 10 has a similar pattern to Cluster 2 in the six-cluster analysis. From these results it appears that while the minimum centroid distance is lower when

using an increased number of clusters, the results produced clusters that were smaller and less consistent.

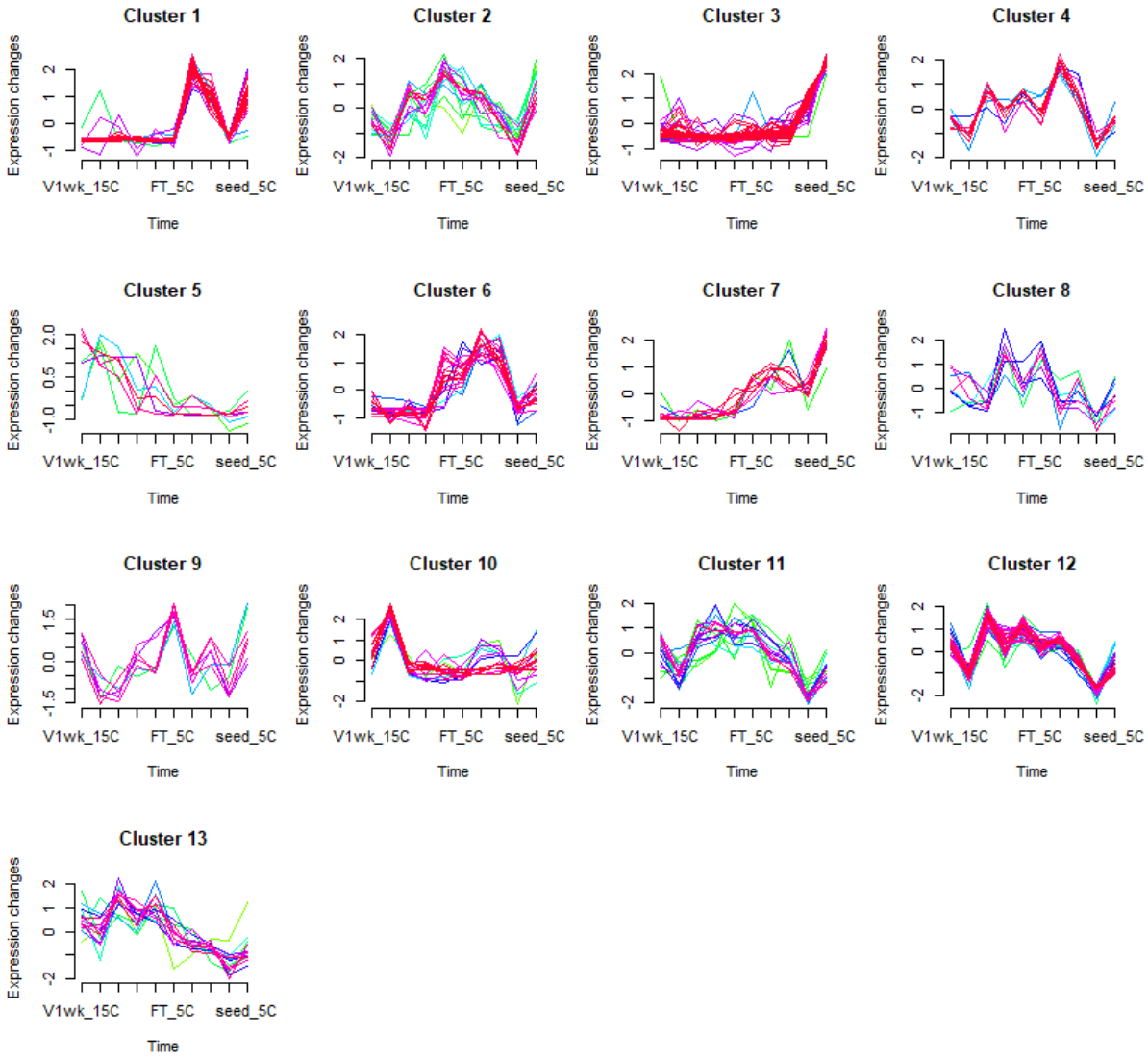

Figure 6.26: Clustering of genes upregulated in seeds of 5°C vernalised plants, with thirteen clusters producing by fuzzy clustering. X-axis shows timepoints ordered by chronological timepoints and then vernalisation temperature, 15°C and 5°C.

Selecting an ideal number of clusters for genes upregulated from the 5°C vernalisation treatment was less straightforward as the minimum centroid distance did not decrease steadily with increasing cluster number (Figure 6.23). As the number of clusters

increased, the number of genes in each cluster decreased. The smallest cluster using 6 clusters contained 23 genes; when the genes were separated into either 11 or 13 clusters, the smallest number of genes per cluster was 7. The information from clusters with smaller numbers of seeds may be useful in identifying genes with highly similar expression patterns within the larger clusters. To maximize cohesion of gene expression within clusters and reduce the number of clusters containing low numbers of genes with inconsistent expression patterns, the group of genes upregulated in seeds from the 5°C vernalisation treatment were separated into 6 clusters.

6.3.2.4.3 Examining expression profiles of gene clusters produced using a fuzzy C-means algorithm

The graphs used to identify cohesion show expression across all timepoints regardless of vernalisation temperature. While these are helpful for understanding how well the clustering algorithm is working, they do not show an information profile for comparing gene expression patterns between the two treatments. To visualise this, the mean FPKM of genes at each timepoint in each cluster was graphed for the two vernalisation treatments (Figures 6.27 and 6.28).

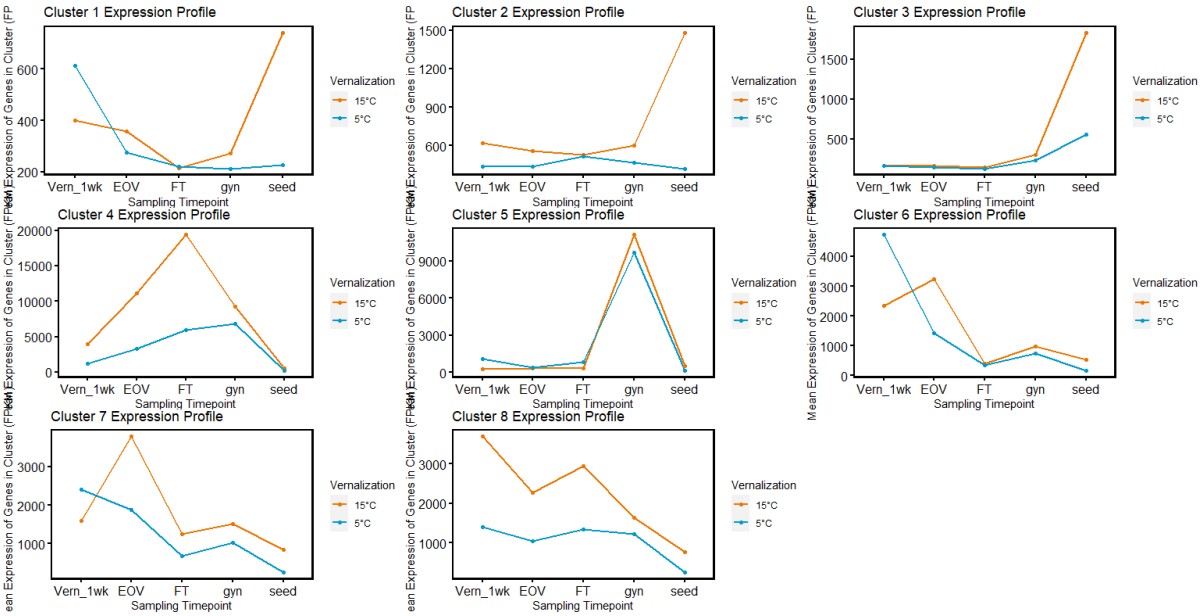

Figure 6.27: Expression profiles over time of genes upregulated in 15°C seeds by cluster.

A range of expression patterns were identified in both sets of differentially regulated genes (Figures 6.27 and 6.28). These clusters can be used to identify genes which have similar expressions to genes of interest.

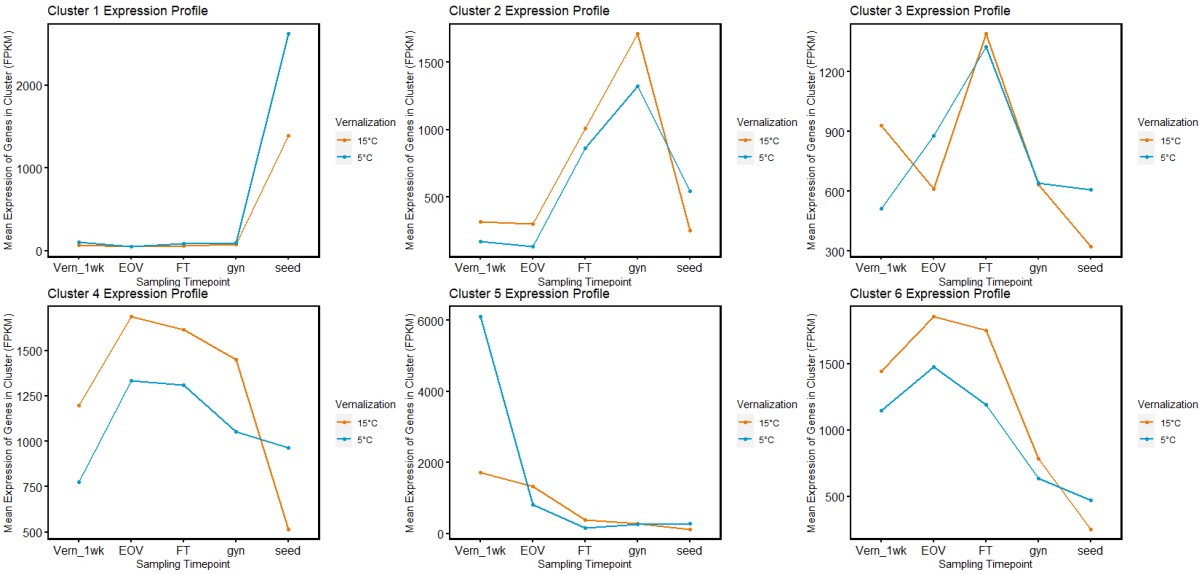

Figure 6.28: Expression profiles over time of genes upregulated in seeds produced after 5°C vernalisation by cluster.

Cluster 3 of genes upregulated in 15°C seeds displays an overall pattern of genes switching on in seeds (Figure 6.27). This cluster includes *MYB77, MYB44, TCH4, PAPP2C, TFL1, PBP1, MERI-5* and *GATL8*. GO term enrichment for this cluster showed enrichment for positive regulation of auxin signalling, suggesting that auxin signalling activity is higher in seeds produced from plants vernalised at 15°C compared to 5°C.

6.3.2.5 Difference in overall gene expression in seeds between vernalisation temperatures

Gene expression was compared in seeds produced by plants vernalised at either 5°C or 15°C, 14 days after pollination. In total, there were 527 differentially expressed genes, with 46% of these being upregulated in plants vernalised at 5°C (Table 6.1).

In seeds produced from plants vernalised at 5°C, upregulated GO terms related to seed and embryo development and nutrient transport (Figure 6.29a). GO terms analysis of genes which were downregulated in seeds produced at 5°C compared to those produced at 15°C indicated that seeds from 15°C-vernalised plants has higher expression of genes related to hormone and stress responses (Figure 6.29b).

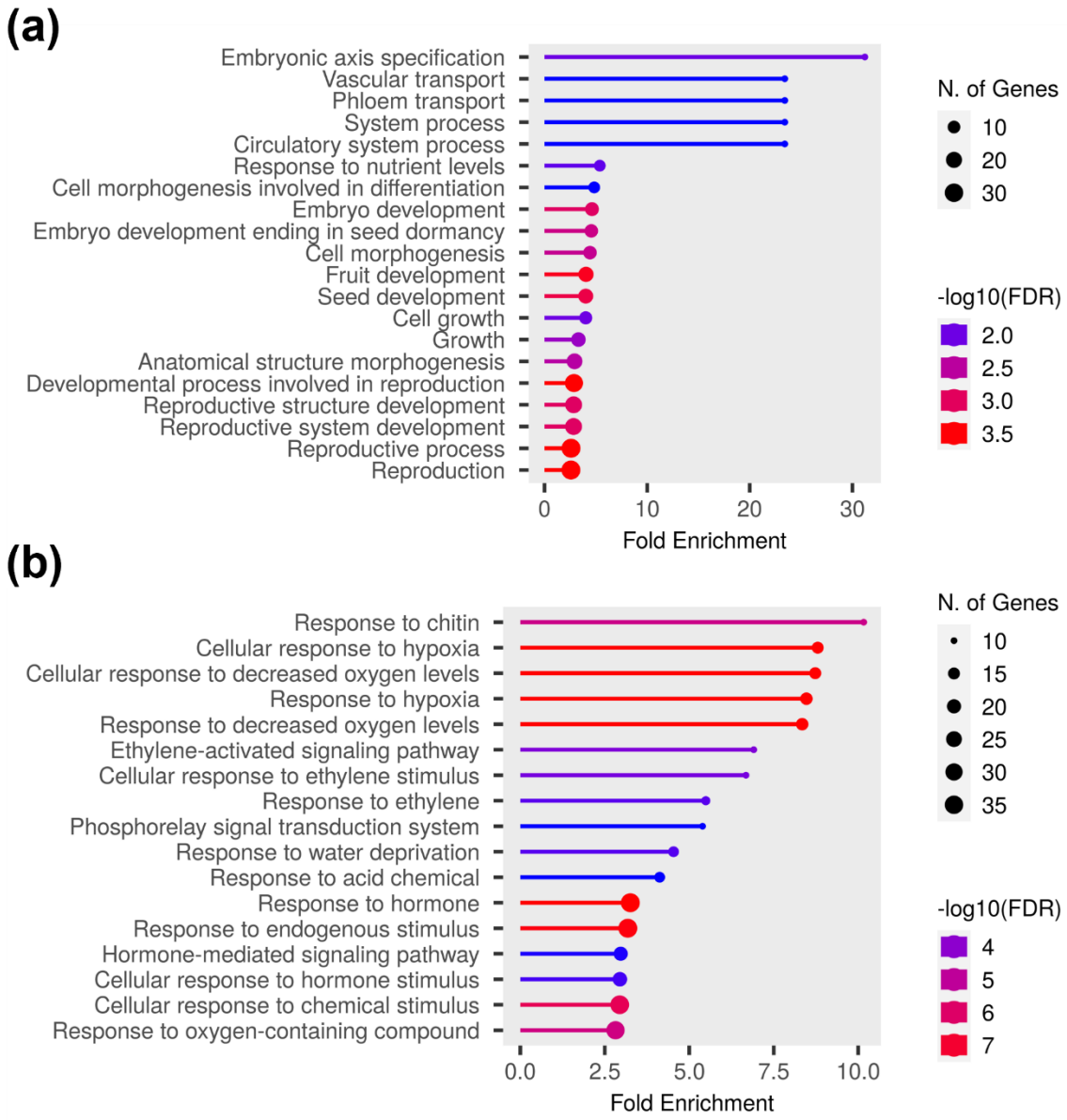

Figure 6.29: GO term enrichment for genes (a) upregulated (b) downregulated, in 14DAP seeds produced at 5°C compared to 14DAP seeds produced at 15°C.

Genes upregulated in seeds produced from plants vernalised at 5°C included several EMB genes, including BnaC0406520D and BnaA04g18070D (both orthologues of *EMB14/SUS2*), BnaA08g10450D (*EMB1507/BRR2a*) and BnaA08g29250D (*EMB1011/LNO1*). EMB genes are classified by their loss-of-function mutants resulting in embryo-defective phenotypes, highlighting the importance of these genes for embryo

development (Meinke, 2019). *EMB1507/BRR2a* is involved in splicing of FLC transcripts (Mahrez et al, 2016), while EMB14 is also involved in *FLC* repression and splicing of COOLAIR transcripts (Marquardt et al, 2014). EMB1011 affects cell division during embryo development, with mutants arresting during embryogenesis (Braud et al, 2012).

This may suggest that although a difference in size is not yet visible at 14DAP, seeds produced after cooler vernalisation are more highly expressing developmental genes which could lead to size differences seen at 21DAP (Figure 6.12). Differences in expression of EMB genes which also affect *FLC* regulation between the two temperature treatments suggests that changes in the seed could be related to *FLC*.

Genes which were downregulated in seeds produced from plants vernalised at 5°C compared to 15°C included multiple copies of *IAA3* (on chromosomes A09 and C08). *IAA3* is involved in auxin signalling and has been shown to function as both an activator and repressor of different auxin responses in roots (Tian and Reed, 1999; Oono et al, 2002). This suggests that auxin signalling may be altered in seeds produced from plants vernalised at 15°C compared to 5°C. The list of genes which were more highly expressed in seeds from plants vernalised at 15°C also included multiple *FLC* genes, which was further explored.

6.3.2.6 Differences in *FLC* expression in seeds between vernalisation temperatures

The response of *FLC* expression changes in response to temperature is well characterised (reviewed in Whittaker and Dean, 2017). To learn more about how changes in *FLC* expression could be affecting seed size, expression patterns of all nine *Brassica napus FLC* copies were compared across the five timepoints studied in this transcriptome experiment (Figure 6.30).

Most *FLC* genes are downregulated after 6 weeks of vernalisation at 5°C, with significantly higher expression in the shoot apex after 6 weeks of 15°C vernalisation

compared to at 5°C. This pattern is seen to varying degrees amongst seven of the nine *FLC* copies. In the C09a and C09b copies, which lack a vernalisation response (Schiessel et al, 2019), expression of *FLC* rises again in the floral meristem.

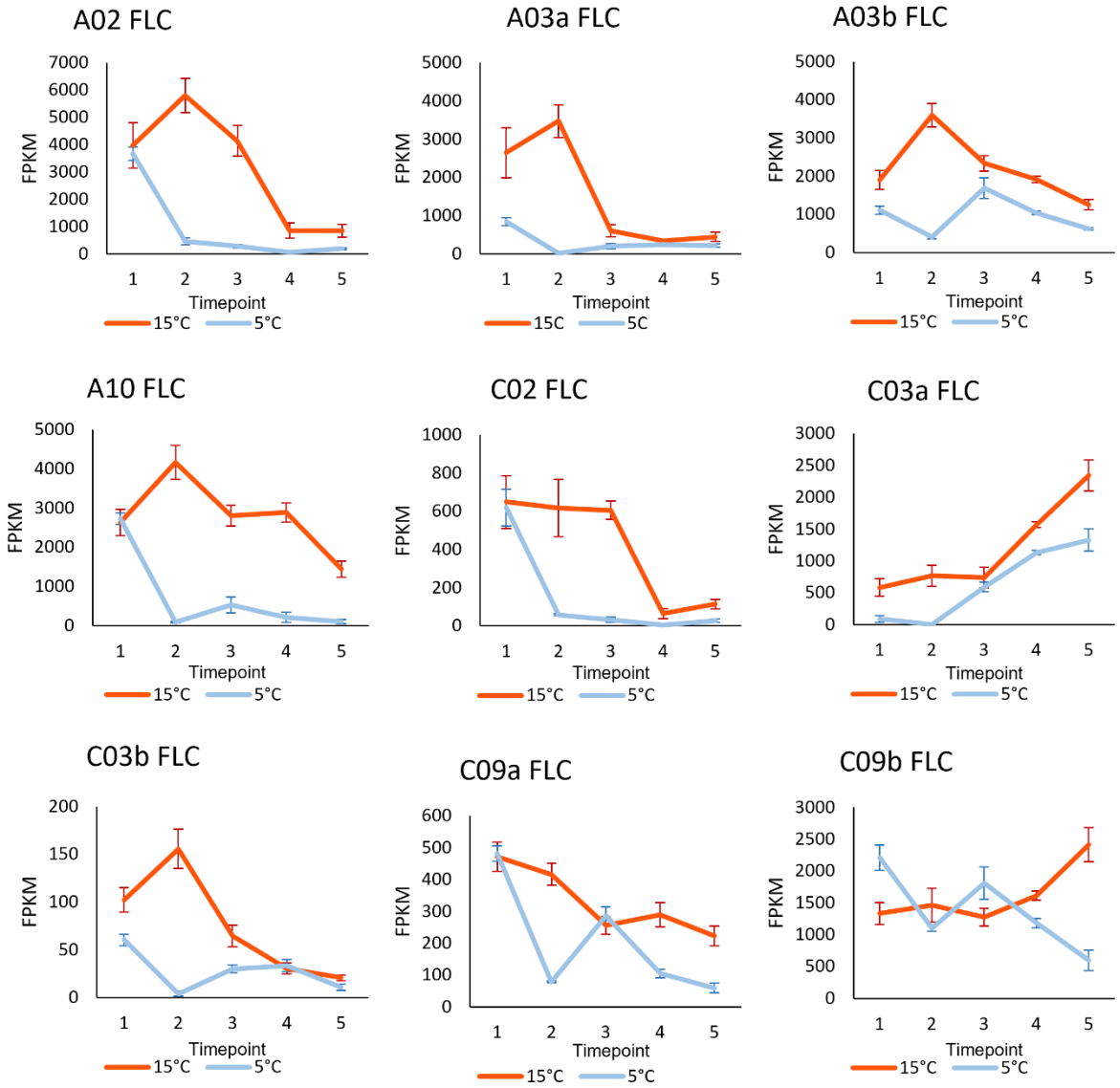

Figure 6.30: Expression patterns of 9 Brassica napus FLC copies across five sampled timepoints in Catana. Timepoints on x-axis are: 1 = 1 week of vernalisation; 2 = 6 weeks of vernalisation; 3 = floral transition; 4 = gynoecia; 5 = 14DAP seed.

In five copies of FLC (A03a, A03b, A10, C03b, C09a), expression had decreased by the point of the floral transition from levels at the end of the 15°C vernalisation treatment. Excepting the C09 copies, expression remained higher than in the 5°C vernalised plants. For both C09 copies of *FLC*, there was an increase in expression between the end of vernalisation and the floral transition.

FLC expression for all copies in the seed was higher in plants that had been vernalised at 15°C compared to 5°C, showing that the effects of vernalisation temperature on *FLC* expression are retained during seed development.

FLC copies *Bna.A10* and *Bna.A02* (two functional copies which were differentially expressed in seeds) were found in cluster 7 of the genes upregulated in 15°C vernalised plants. In this cluster of genes, the general trend for gene expression was to decrease from vernalisation through to the seed in plants vernalised at 5°C, compared to being more highly expressed in plants vernalised at 15°C and showing increased expression after 6 weeks of vernalisation.

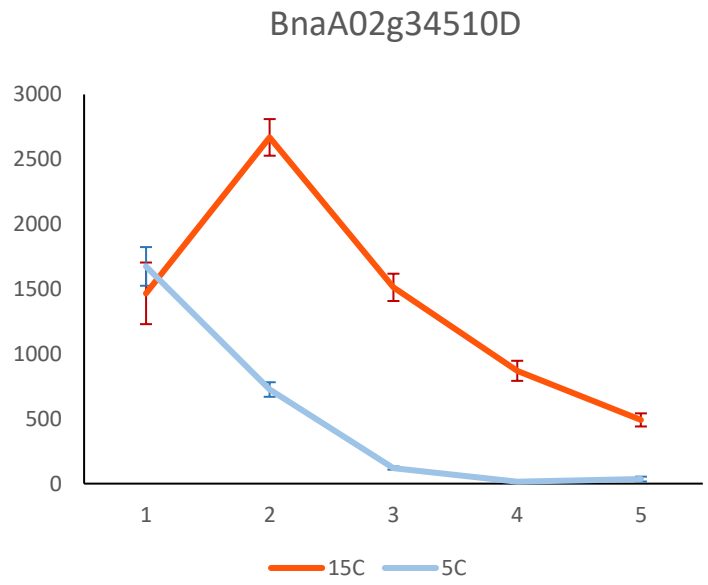

Figure 6.31: Expression of MAF3 (BnaA02g34510D) across five sampled timepoints in Catana. Timepoints on x-axis are: 1 = 1 week of vernalisation; 2 = 6 weeks of vernalisation; 3 = floral transition; 4 = gynoecia; 5 = 14DAP seed.

Cluster 7 contained 31 genes overall. This cluster of genes also contained *MAF3*, a flowering repressor gene which is closely related to *FLC* (Ratcliffe et al, 2003). The expression profile of *MAF3* over time was similar to that of the functional *FLC* genes (Figure 6.31). In plants vernalised at 5°C, *MAF3* expression decreased through vernalisation and remained low from the floral transition onwards. In plants vernalised at 15°C, *MAF3* expression was high by the end of vernalisation treatment and remained higher than in 5°C treated plants through the rest of development. These results suggest that the differences in gene expression seen after vernalisation are still impacting gene expression in the seed.

6.4 Conclusions & Discussion

From the experiments detailed in this chapter comparing gene expression between plants of the *Brassica napus* variety Zhongshuang 11, In the Semi-winter OSR *Brassica napus* variety Zhongshuang 11, the temperature experienced by the plant during the floral transition was not found to impact the expression of vernalisation-related genes using by GO term analysis, and there was not found to be any differential gene expression of *FLC* genes. Differences were identified in patterns of stress-related gene expression in the shoot apex at BBCH51, alongside differences in reproductive gene expression in the gynoecia.

The first of these was the increased activity of the jasmonate signalling in the gynoecia of plants which underwent the floral transition at 17°C. Jasmonic acid and jasmonate signalling pathways have known roles in increasing tolerance to high temperature stress (Balfagón et al, 2019; reviewed in Raza et al, 2021). Although these plants were not exposed to high temperature stress, increased jasmonic acid signalling suggests that these samples were displaying an increased stress response compared to samples which underwent the floral transition at 5°C.

A number of genes with GO terms highlighting involvement in pollen development and formation of pollen structures were comparatively upregulated in the shoot apex of plants at BBCH51 which underwent the floral transition at 5°C. It is possible there are paternal effects on seed development via the pollen, suggesting that pollen development

could have been delayed in plants which went through the floral transition at 17°C. Changes in pollen development could also affect overall seed production due to changes in fertility.

Seed size was not compared between plants vernalised for 3 weeks and for 12 weeks at 5°C. Seed size results from the glasshouse phenotyping experiment (Chapter 2) were analysed after the transcriptomic experiment had been carried out; these results showed that increasing the duration of the vernalisation treatment from 6 weeks to 12 weeks was not found to have a significant effect on seed size in Zhongshuang 11 (ANOVA p = 0.729 as seen in Table 6.2). As the floral transition was here shown to occur after 7-8 weeks at 5°C, the 6-week to 12-week vernalisation comparison could have similar effects of temperature on the floral transition as seen when comparing 3-weeks to 12-weeks. This could indicate that the floral transition occurring in the cold, and the resulting changes in gene expression described here, had no significant impact on seed size in the Semi-winter OSR variety Zhongshuang 11.

Table 6.2: ANOVA results assessing effect of temperature treatments on TGW in Zhongshuang 11.

	Degrees Freedom	**Sum of Squares**	**Mean of Squares**	**F value**	**Pr(> F)**
Vernalisation Duration	1	0.107	0.107	0.122	0.729
Vernalisation Temperature	2	0.237	0.119	0.135	0.874
Maturation Temperature	1	5.583	5.583	6.375	0.017
Residuals	31	27.150	0.876		

In the Winter OSR variety Catana, the results of ovule and seed size analysis shows that seeds produced from plants vernalised at 5°C are larger than seeds produced from plants vernalised at 15°C. The results suggest that this seed size increase is due to differences in development that arise after fertilization and increase throughout the first 21 days of seed development. The seed itself is a much larger structure than the ovule. As there is no difference in seeds at 14DAP, I conclude that that the difference in seed

size between vernalisation temperatures seen at later development timepoints is not related to early seed development. Instead, this may be due to changes in the seed occurring between 14 and 21 days after pollination. In this experiment, no samples were taken in the time in between these growth stages. Although differences in seed size and development do not become apparent until 21DAP, it is alternatively possible that seed development occurs more rapidly in seeds produced at 5°C before this point. As ovules were found to be smaller after a 5°C vernalisation treatment, the fact that seed size is equivalent between seeds produced at both temperatures by 14DAP suggests that seeds produced at 5°C may have "caught up" in size due to an increased growth rate. However, a more thorough examination of how seeds develop during this window could reveal more about how this difference in seed size arises.

In both the Semi-winter OSR variety Zhongshuang11 and the Winter OSR variety Catana, there were notably fewer differentially expressed genes in gynoecia compared to other timepoints. However, in Catana the genes that were differentially expressed at this point included known vernalisation genes including *FLC* and *MAF3*.

Although *FLC* expression remained high after six weeks of vernalisation at 15°C in the Winter OSR variety Catana, plants were able to go through the floral transition to flowering and seed production. This could indicate that levels of *FLC* expression were still able to decrease below the threshold for the floral transition, or that the cause of the floral transition is due to the activity of other genes than the fall in *FLC* expression. This could be ascribed to the action of FLC-independent vernalisation responses, such as those seen in *Arabidopsis* via the activity of the MADS-box genes *AGL19* and *AGL24* (Shönrock et al, 2006; Michaels et al, 2003; Alexandre and Hennig, 2007). Genes in the *Arabidopsis* flowering pathway largely have orthologues in *Brassica* species, including *AGL24* and *AGL19* (Scheissel et al, 2020; Shah et al, 2018).

At the end of vernalisation, *Bna.FLC.C09b* (BnaC09g46540D) was not found to be differentially expressed between vernalisation treatments. This *FLC* gene does not have a cold response (Schiessel et al, 2019). However, it was one of the four *FLC* genes to be differentially expressed in seeds. Eight of 9 *FLC* copies were downregulated by vernalisation in a previous study (Leijten et al, 2018).

Differences in *FLC* between plants vernalised at 5°C and 15°C were seen throughout development, including in the gynoecia and in the seed, highlighting the persistent changes in expression induced by vernalisation temperature. As this effect also included *MAF3*, this suggests that the effect is not solely focused on *FLC* but that overall regulatory pathways responding to vernalisation may be having an effect on seed development.

Transcriptomic analysis also suggests that there is differential auxin signalling between seeds from plants which experienced different vernalisation temperatures, with auxin signalling being positively regulated following a warmer vernalisation treatment. High levels of auxin in the seed have been shown to delay endosperm cellularization (Batista et al, 2019); delayed endosperm cellularization is known to contribute to reduced seed size (Zhou et al, 2009; Zhang et al, 2020).

FLC is reactivated in the seed to ensure there is a vernalisation requirement across generations. Resetting of *FLC* in *Arabidopsis* begins in the globular embryo and levels increase through embryo development until *FLC* is fully reset in the mature seed (Choi et al, 2009; Crevillén et al., 2014). At 14DAP in *Brassica napus*, embryos have usually reached the heart stage of development. The majority of FLC resetting in the embryo is initiated by LEC1 (Niu et al, 2019; Tao et al. 2019). Many of the genes which are differentially regulated between plants vernalised at 5°C and 15°C were expressed in the endosperm. In the endosperm, *FLC* resetting is prevented by the activity of *FVE* (Choi et al, 2009). This could suggest that differences in *FLC* between seeds vernalised at different temperatures may be due to maintenance of differences in *FLC* expression in endosperm. More tissue-specific analysis of *FLC* expression in seeds produced from plants vernalised at different temperatures could reveal more details of which genetic pathways are likely to be affected by differences in *FLC* expression and affect seed development as a result.

In addition to differences in *FLC* regulation between seeds produced after the two vernalisation treatments, seeds also showed differences in EMB gene including *EMB14* (also known as *SUS2* or *PHP8*). Loss-of-function mutants in *EMB14* show increased *FLC* expression, and *EMB14* has been shown to repress *FLC* expression through

regulation of COOLAIR splicing (Marquardt et al, 2014). As this embryo development gene which represses *FLC* is more highly expressed in seeds from plants vernalised at 5°C seeds, which are bigger at 21DAP, this could indicate a link between embryo development and *FLC* regulation.

The results of transcriptome analysis described in this chapter include increased expression of vernalisation-responsive genes such as *FLC* and *MAF3* at 14DAP in seeds produced from plants which experienced a warmer vernalisation treatment. However, at this point in development there was no significant difference in seed area between the two vernalisation temperatures; differences in seed area were instead seen later, at 21DAP. This indicates that the phenological effects of differential gene expression between the temperature treatments took a longer time to become apparent. As gene expression was not studied at 21DAP in this experiment, it would be an interesting future experiments to learn if differential gene expression changes when the differences in development have become apparent.

Chapter 7 - Methods & Materials

7.1 Phenotyping

7.1.1 Plant Material for Glasshouse Phenotyping Experiment

Seeds for *Brassica napus* experiments were sourced from the *Brassica napus* DFFS (Diversity Fixed Foundation Set) (Harper et al, 2012; Latunde-Dada et al, 2004; data available at https://www.brassica.info/resource/plants/diversity_sets.php)

7.1.2 Glasshouse Phenotyping

Plant growth for the seed phenotyping experiment consisted of a diversity set of 96 *Brassica napus* lines grown under 12 different conditions, to a total of 3, 456 plants. The conditions were: vernalisation in a CER at either 5°C, 10°C or 15°C; for either six or twelve weeks; then growth until maturation in a glasshouse at a temperature of either 18°C or 24°C. Temperatures in the glasshouse were cooler at night and warmer during the day. Ambient daylight hours in the UK between January and August were present in the glasshouse. Three plants of each line under each specific condition were grown for biological replicates. Sowing was staggered so that the plants would all come out of vernalisation at the same time. As the plants started to dry out and become ready for harvesting, water levels were reduced, and the most matured plants were raised up on pots to allow the plants to fully dry out. When plants were fully dry, 20 pods were harvested from the main raceme of the plant.

Seed pods were threshed, and the size and number of seeds was analysed using a MARVIN seed analyser. This counted seed number; seed number per pod was then determined by dividing total seed number by the number of pods. The MARVIN also produced figures for total seed weight, thousand grain weight, area, width, and length. Pre-harvest sprouting was determined manually and graded on a scale of 1 to 6, with 1 indicating the greatest degree of pre-harvest sprouting.

7.2.2 CER Phenotyping

Brassica napus plants of varieties Catana, Palmedor, Castille and Expert were grown at 17°C for three weeks of pre-growth in a Panasonic growth cabinet and were transplanted into 24-cell trays at 7 days old. After this, plants were placed into vernalisation at a constant temperature of 5°C or 15°C on an 8-hour day cycle. Vernalisation lasted for 6 weeks.

Post-vernalisation, plants were grown in CERs on 16h day/8h night cycles. The temperature regime of the cooler CER was a daytime temperature of 20°C and a night-time temperature of 16°C. The temperature regime of the cooler CER was a daytime temperature of 26°C and a night-time temperature of 22°C. After harvesting, samples were immediately frozen in liquid nitrogen and stored at -80°C.

7.2 Germination

7.2.1 Seed germination experiment

Seeds were sent to IBERS in Aberystwyth to be germinated. Seed batches were randomized, sown in pots with 4 seedlings per pot and left to germinate on a conveyor belt system at temperatures of 22°C during a 14-hour day and 18°C at night.. Two belt runs were used due to the large volume of seedlings and two technical replicates were grown per run, with biological replicates randomized. Watering treatments were grown next to each other to reduce environmental variation between treatments. The pots were imaged through the conveyor belt system multiple times a day and the image analysis software was used at IBERS to provide data on plant area. Germination was also manually scored for days after sowing, as time taken for the seedling to be fully emerged and the hypocotyl hook unfurled.

7.3.1 Associative Transcriptomics

The SNP data used for associative transcriptomics is described in Harper et al, 2012. This data set contains 355,536 SNPs from leaf transcriptome data before filtering; removing SNPs found in fewer than 5 varieties left 177,725 SNPs across the dataset.

GWAS on all temperature treatments and lines was run as a GLM in TASSEL 5 (Bradbury et al, 2007). The Q matrix for kinship analysis was derived from STRUCTURE by Guanyuang Lu at JIC. Minor allele frequency was set to 0.05. Meta-analysis of GWAS results was carried out using the METAL tool (Willer et al, 2010; available at http://csg.sph.umich.edu/abecasis/Metal).

GWAS was run firstly on each of the individual 12 treatments. Comparison tables between treatments were then analysed. Comparisons for maturation temperature were created by comparing plants grown at 18°C against plants grown at 24°C, from the same vernalisation treatment. Comparisons for vernalisation treatments were created by comparing plants vernalised at 5°C against plants vernalised at 15°C, for the same duration (6 or 12 weeks) and matured at the same temperature.

7.4 Genetic Analysis

7.4.1 General DNA Extraction Protocol

DNA from *Brassica napus* for sequencing and PCR was extracted using the Edwards method (Edward's buffer: 200 mM Tris-HCl to pH 7.5, 250mM NaCl, 25mM EDTA to pH 8.0, 0.5% SDS). Tissue was ground on ice and 400µl of Edward's buffer added. Samples vortexed at room temperature for 5 seconds before centrifuging for 1 minute at maximum speed. An equal amount of the supernatant and isopropanol were mixed in a fresh microcentrifuge tube. Samples centrifuged at maximum speed for 5 minutes until formation of a pellet. Supernatant discarded and pellet allowed to air dry before being resuspended in 50µl of distilled water. Samples were stored short-term at 5°C and long-term at -20°C.

7.4.2 DNA Extraction for *Brassica napus* DFFS

DNA from the 96 lines of the diversity set was carried out by Richard Goram of the Genotyping Platform at JIC. Leaf tissue was harvested and stored in a 96-well collection microtubule box. Extraction buffer was (0.1M Tris-HCl pH 7.5, 0.05 EDTA pH 8.0,

1.25% SDS) heated to 65°C. For each 96-well plate, 37ml extraction buffer with 100µl RNaseA was prepared. To each extraction tube, 333µl of extracted buffer was added and plates sealed before shaking in a Genogrinder for 2 minutes at 1750rpm. Spin plate was then pulsed until reaching 3000rpm using a Sigma 4-15 centrifuge. Plate was incubated at 65°C for 45 minutes before cooling to room temperature. Samples were pulse spun to 6000rpm and 167µl of 6M ammonium acetate added before sealing plates and shaking for 15 seconds. Samples were then stored fat -20°C for 15 minutes before centrifuging for 8 minutes at 6000rpm to precipitate proteins/plant tissue. Supernatant transferred into cold 1.2ml storage plates containing 240µl of isopropanol per well. Plates were once again sealed and shaken for 15 seconds before pulse spinning. DNA was then left to precipitate for 20 minutes at -20°C. Samples centrifuged for 12 minutes at 5200rpm to pellet DNA and supernatant tipped off before washing pellet in 350µl of 70% ethanol. Plates centrifuged again for 12 minutes at 5200rpm to pellet DNA and supernatant tipped off before plates were left to dry for 30 minutes at 65°C. Pellet was resuspended in 200µl of water. Plates were then sealed and DNA allowed to dissolve at room temperature for 15 minutes before vortexing and centrifuging plates for 10 minutes at 5200rpm.

7.4.3 Gene sequencing of candidate genes in *Brassica napus*

Primers were designed using Primer3 (https://primer3.ut.ee/) and the gene of interest was amplified using PCR; presence of DNA in the sample was confirmed using electrophoresis on a 1% agarose gel and a Nanodrop to check concentration.

DNA was also purified after PCR; some sampled were extracted from gels and then purified. DNA purification was carried out using a NucleoSpin Gel and PCR Clean Up kit according to the manufacturer's instructions. Purified DNA was sent to Genewiz for sequencing.

7.4.4 Standard PCR Protocol

Unless otherwise specified, PCR samples were made up to 25µL using the following recipe:

12.5µL GoTaq G2 Green Master Mix + 9.5µL H2O + 1µL forward primer + 1µL reverse primer + 1µL template DNA

All PCR reactions began with a hot start of 110°C. The standard PCR protocol then used of an initial denaturation at 95°C for 3 minutes. The PCR cycle then consisted of 30 seconds of denaturation at 95°C, followed by a standard annealing step of 30 seconds at 60°C and completing with an extension step of 60 seconds at 72°C. This cycle was repeated 30 times. A final extension step was carried out for 5 minutes at 72°C.

Annealing temperatures were adjusted to account for the T_m of the primers used and time was adjusted to account for the length of the gene fragment being amplified. After completion of the reaction, samples would be held in the machine at 10°C.

All samples were run on a GStorm ThermoCycler.

7.4.5 Gel Electrophoresis

Unless otherwise specified, samples were run on a 1% agarose gel at 100V using a BioRad gel electrophoresis tank powered by a Kikusui power pack, model PAB. For every 100ml of agarose gel, 5µL of ethidium bromide was added. Unless otherwise specified, ladder used was 1bk Plus DNA Ladder from New England Biolabs.

7.4.6 PCR Product Cloning

Individual copies of AN3 from *Brassica napus* were cloned into the Promega pGEM-T Easy Vector System. Colony selection was carried out using ampicillin/carbenicillin.

7.4.7 Seed Oil Content

Seed oil content of *Brassica napus* seeds was determined using NMR at Rothamsted Research, using a Bruker 600 MHz NMR machine with attached Samplejet autosampler. All oil measurements were normalised to 9% moisture content.

7.4.8 Exome Capture Data

Exome capture data was used to analyse *GDSL1* sequences in IGViewer (Robinson et al, 2011). This data was generated by Woodhouse et al, 2021 and is available in the NCBI SRA, accession number PRJNA309368 (data available at URL: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA309368). Details of bait data used to generate this dataset is available in Steuernagel et al, 2021 (URL: https://zenodo.org/record/4473283#.Y4n6An3P1D-).

7.4.8 Arabidopsis Seed Phenotyping

Seeds for *Arabidopsis thaliana* plants were sourced from NASC. Seed number was counted by imaging seeds using a dissection microscope and counting the number of seeds using ImageJ. The same seeds were then weighed on an electronic scale in milligrams. Three technical repeats were carried out for each plant and four biological replicates were used for each line.

7.5 Transcriptomics & Vernalisation

7.5.1 Transcriptome Sampling for Vernalisation Duration Effect

Zhongshuang 11 plants were sown and grown in a CER with a 16- hour day, an 18°C day temperature and 16°C night temperature, at 70% humidity. They remained in these conditions for three weeks of pre-growth and were transplanted into 24-cell trays at 7 days old. After this, plants were placed into vernalisation at a constant temperature of 5°C on an 8-hour day cycle. Vernalisation lasted for 12 weeks.

Post-vernalisation, plants were placed back into the original conditions (16-hour day, 18/16°C temperature cycle) and transplanted into 1L pots. Plants grew for another 4 weeks during which time floral samples were taken.

Samples were taken of whole 7-day seedlings, then of leaf and apex on a weekly basis thereafter through to flowering. Bud, anther, and gynoecia samples were taken once the plants were flowering. Seed samples were also taken at the heart, torpedo, green seed,

and mature stage. Seed samples were of the embryo, endosperm, and seed coat; at the mature seed stage, endosperm and seed coat were sampled together.

7.5.2 Transcriptome Sampling for Vernalisation Temperature Effect

For transcriptome sampling and gynoecia analysis, *Brassica napus* plants of variety Catana were grown at 17°C on a 16-hour day cycle for three weeks of pre-growth and were transplanted into 24-cell trays at 7 days old, grown in John Innes Cereal soil mix. After this, plants were placed into vernalisation at a constant temperature of 5°C or 15°C on an 8-hour day cycle with light levels of 60 - 90 μmol m^2 s^{-1}. Vernalisation lasted for 6 weeks.

Post-vernalisation, plants were grown in a CER on 16h day/8h night cycle, with a daytime temperature of 20°C and a night-time temperature of 16°C, until ready to be harvested. One week after being transferred to the CER, plants were transplanted from 24-cell trays into 1L pots using John Innes Cereal soil mix. Samples were immediately frozen in liquid nitrogen and stored at -80°C.

7.5.3 RNA Extraction

RNA from samples was extracted using the EZNA Plant RNA Kit from Omega Bio-tek, according to the manufacturer's instructions. Extracted RNA was sent to Novogene for sequencing and completed samples were analysed for differential gene expression.

7.5.4 RNAseq Analysis

Transcriptome samples from Novogene were run through an RNAseq pipeline. Raw sequence reads were trimmed used Trimmomatic (Bolger at al, 2014). Quality control of reads was carried out using FastQC (Andrews, 2010) on raw and trimmed reads. Sequences were aligned to the reference sequence using HiSat2 (Kim et al, 2015). The reference sequence used was the Darmo-bzh v4.1 (Chalhoub et al, 2014). Sequences

were sorted, indexed, and filtered using Samtools (Li et al, 2009). StringTie was used to assemble and quantify reads (Pertea et al, 2015). An in-house python script was then used to extract read count matrices for genes and transcripts to produce final matrices for analysis. For analysis of DEGs in Catana, DESeq2 was used (Love et al, 2022).

To determine significant DEGs, p value was adjusted to $p < 0.05$ and absolute log fold change was adjusted to $\log_2 FC > 2$. Clustering of genes with similar expression patterns was carried out using a fuzzy C-means clustering method with the Mfuzz package in R (Futschik, 2022). GO terms were visualized using ShinyGo 0.76 (Ge et al, 2020; tool available at http://bioinformatics.sdstate.edu/go/)

7.5.5 Gynoecia Sampling and Imaging

Gynoecia were sampled and vacuum infiltrated with FAA (50% ethanol, 10% formalin, 5% acetic acid) on ice. FAA was replaced with 70% ethanol and samples were stored like this until ready for clearing. For clearing, samples were dehydrated with an ethanol series of 70%, 80%, 90% and absolute ethanol washes. Samples were then placed in Hoyer's solution (75% chloral hydrate, 20% water, 5% glycerol) and were left to clear for 24 hours before being imaged with a dissection microscope. Microscope used was a Leica GX with a GXCAM HiChrome HR4 camera attachment and GXCAM software. Images were analysed using ImageJ.

7.6 General methods

7.6.1 Plant growth

7.6.1.1 Sowing and germination of seeds

Arabidopsis thaliana lines were sown to agar (1/2 MS agar, 1% sucrose) and germinated in a Panasonic cabinet at 22°C 16h day/8h night, before transplanting to soil. *Brassica napus* plants were sown directly to John Innes Cereal Mix soil and germinated at 17°C, 16h day/8h night in a Panasonic cabinet with light levels of 80 – 100 $\mu mol\ m^2\ s^{-1}$.

7.6.2 Data Analysis

7.2.2.1 Data Analysis in R

R version used for production of final statistics and graphs was R 4.0.2 “Taking Off Again”, from the R Foundation for Statistical Computing (R Core Team, 2020).

7.6.2.2 Linear modelling

Linear modelling was carried out in R. For mixed linear models, the package lmerTest was used, with linear mixed models fit by REML and t-tests using Satterthwaite's method. For general linear models, the glm function in base R was used, fit using a Gaussian family function unless otherwise specified. Pairwise comparisons were run using the emmeans package. Results of linear models were shown graphically using either the ggplot2 package in R, or Microsoft Excel.

7.6.2.3 Analysis of Variance

ANOVA was run using R. Statistical significance groups were calculated using Tukey’s post-hoc test, with results generated and graphed using the agricolae package in R.

7.6.2.4 Crop type analysis

When analysing the effects of environmental conditions on crop type, crop type groups were considered for analysis if they contained more than five varieties to ensure sufficient genetic variety within the group. This resulted in the removal of the crop types Winter fodder and Spring fodder from these analyses.

7.6.2.5 Main R Packages Used

lmerTest - for mixed linear modelling (Kuznetsova et al, 2017)
emmeans - for linear modelling (Lenth et al, 2022)
ggplot2- for producing graphs (Wickham, 2016)
dplyr - for data cleaning and organization (Wickham et al, 2022)
agricolae - for post-hoc testing (Mendiburu et al, 2022)
edgeR - for RNAseq analysis and cluster normalization (Yunshun et al, 2022)

DESeq2 - for DEG analysis (Love et al, 2022)
Dendextend - for visualizing DEG clusters (Galili et al, 2022)
Mfuzz - for fuzzy clustering (Futschik, 2022)

Chapter 8 – Discussion

8.1 Chapter Summaries

8.1.1 Exploring the effect of temperature on *Brassica napus* yield

It has been previously shown that temperature during winter is correlated with yield in *Brassica napus* (Brown et al, 2019), and that temperature during seed maturation affects yield in *B. napus* (Rashid et al, 2018). To better understand how the *B. napus* crop is affected by changes in temperature throughout development, I tested the hypotheses that seed weight and seed number per pod would be affected by the temperature of vernalisation, the duration of vernalisation, and the temperature during seed maturation. To achieve this, the *B. napus* Diversity Fixed Foundation Set (Latunde-Dada et al, 2004) was phenotyped for seed traits across a range of temperature treatments. Combining these temperature treatments allowed for detailed statistical analysis to determine differences between each treatment factor. As the DFFS contains multiple different crop types of *B. napus*, differences could be quantified between individual varieties and between crop types as a group.

The results of this analysis showed that temperature during vernalisation, as well as during seed maturation. is correlated with seed weight and seed number per pod. The results indicated that higher temperatures are associated with reduced seed weight and fewer seeds per pod. Using seed weight per pod as an indicator of yield confirmed that these effects were additive, and that yield was overall lower in varieties grown at warmer temperatures during both vernalisation and seed maturation. However, duration of vernalisation was not found to be correlated with seed weight or seed number per pod. Responses to vernalisation temperature differed significantly between crop types. Winter OSR varieties showed the largest reduction in seed weight and seed number per

pod in response to increased vernalisation temperature, while Spring and Semi-winter OSR were affected to little or no degree. Meanwhile, all crop types showed decreases in one or both of the tested yield-related seed traits when grown at higher seed maturation temperatures. The results produced from this experiment formed the basis for much of the further work in this thesis. As this experiment has identified varieties and crop types which are more resistant to increases in vernalisation temperature, the results detailed in this chapter could contribute to development of improved *Brassica napus* crops breeding, as well as selection of more appropriate varieties for growth under expected changes in UK climate.

The results detailed in this chapter indicate that the effect of seed set temperature on seed traits is greater in magnitude than the effect of vernalisation temperature. Although gene expression in the seed can be affected from changes in gene expression during vernalisation, these effects have a longer developmental timescale during which they can be mitigated by other developmental processes. In comparison, changes in gene expression occurring during seed development may be having a more immediate effect on seed development, resulting in this difference of magnitude. This thesis goes on to explore the mechanism by which long-lasting developmental changes in the seed could result from earlier changes resulting from vernalisation temperature effects.

Previous studies exploring how temperature affects yield in *Brassica napus* and other crops have often focused on the effects of temperature stress (Kourani et al, 2022). However, the large phenotyping experiment described in this chapter used growth temperatures which are likely to be experienced by OSR crops in the field in the UK. The highest vernalisation temperature was 15°C, and Met Office data indicates that the average temperature in the East of England during October between 1991 – 2020 was 15°C. Maturation temperature in this experiment was compared between 18°C and 24°C. Met Office data indicates that the average May temperature in the East of England from 1991-2020 was 17.1°C. With the UK Climate predicted to warm in the future, this experiment can provide insight as to how rising summer temperatures could affect OSR yield.

However, *Brassica napus* experiences heat stress at a threshold of around 29.5°C (Morrison and Stewart, 2002), with studies using heat stress treatments often exceeding 30°C (Brunel-Muguet, 2015; Rahaman et al, 2018; Huang et al, 2019; Mácová et al, 2021). The results presented here indicate that even below the threshold for the plant to experience heat stress responses, increases in temperature can cause reductions in yield. This is important information for ensuring that *Brassica* crops grown the UK can be bred and selected to reduced yield loss from changes in the UK climate.

This chapter also details how yield is impacted different by vernalisation temperature in different crop types of *Brassica napus*, which have different vernalisation requirements (Calderwood et al, 2021). If differences in vernalisation response have later impacts for yield, this could have implications for future production of *Brassica napus* crop varieties. *Brassica napus* contains nine *FLC* genes and these *FLC* states already differ between crop types (Yin et al, 2020), which could provide a basis for diversifying vernalisation responses and *FLC* states in future crops to improve yield.

8.1.2 Environmental effects on germination and emergence

Germination, vigour and establishmentare key traits in crops and understanding how these are affected by the environment is important for breeding future high-performing OSR crops (Finch-Savage and Bassel, 2016). To achieve this, the hypotheses were tested that rate of germination and early seedling vigour would be impacted by soil water content and the temperature at which the seed matured. However, as vigour testing was not carried out under low temperatures, these experiments instead tested emergence and early seedling growth.

To test these hypotheses, 84 OSR varieties from the DFFS (Latunde-Data et al, 2014) were germinated and grown as seedlings on a phenotyping platform at IBERS in Aberystwyth. The seeds were grown under contrasting treatments for soil water content to test the effect of drought, and seeds which had been matured at both 18°C and 24°C were included in the experiment to understand the impact of seed maturation temperature. Use of the phenotyping platform enabled the germinating seeds and young seedlings to be imaged, and so image analysis was used to evaluate the effect of each

treatment. Germination was scored by assessing the time in days taken for the hypocotyl hook to unfurl, while early seedling growth was assessed using seedling area in mm^2.

The results of this experiment indicated that the speed of germination was not strongly affected by either the difference in soil water content or the seed maturation temperature. Although the effect size was small, the differences between treatments aligned with previous germination studies in *A. thaliana* and *B. napus* (Springthorpe and Penfield, 2015; Zhang et al, 2014). Early seedling growth and emergence was evaluated using the time taken to reach $1000mm^2$. This metric was significantly affected by both soil water content and seed maturation temperature; however, the effect size of seed maturation temperature was extremely small and would likely have no real-world impact.

Additionally, I tested the hypothesis that early seedling growth, as well as responses to temperature and drought, would vary between crop types. This was done using a linear modelling to assess the results of the experiment. The results of this analysis indicated that different crop types had different rates of early seedling growth, with Semi-winter OSR growing most vigorously as it took the shortest amount of time to reach $1000mm^2$. The effect of seed maturation temperature varied between crop type, with most crop types showing no difference early seedling growth between temperatures. The only crop type which showed a statistically significant effect was Winter OSR, and here the effect was minor as seen previously when assessing the full diversity set. All crop types were found to be significantly negatively affected by reduced soil water content.

A correlation analysis was also carried out to assess the relationship between seed size and seedling vigour. The results of this analysis suggested that TGW was negatively correlated with time taken to reach $1000mm^2$, indicating that larger seeds reached this stage more quickly than smaller seeds. Seed size has been previously linked to seedling growth in *Brassica* crops (Bettey et al, 2000), although this experiment was not able to assess any mechanism behind this relationship.

The results described in this chapter indicate the importance of developing *Brassica napus* crops which can grow effectively at low soil water content. Additionally, analysis suggests that producing larger seeds is not only beneficial for crop yield, but for

producing plants in late generations which will grow more effectively in the field. In this experiment, seedlings were grown for a maximum of 12 days after sowing. It would have been interesting to continue to grow the plants to learn more about how the yield would have been affected by these environmental parameters; however, this experiment was not possible due to time limitations when using the phenotyping platform. Further experiments to assess the effects of drought and seed size on crop performance could increase our understanding of how to produce high-yielding OSR crops under drought conditions.

8.1.3 Using associative transcriptomics to identify the genetic basis of temperature responses in *Brassica napus*

After identifying differences between varieties and crop types in response to temperature during vernalisation and maturation, associative transcriptomics was used to identify genetic variation in *B. napus* associated with phenotypic differences in seed traits.

Associative transcriptomics has been successfully used to explore genetic basis of a wide range of traits in *B. napus*, and a number of studies have been published both before and during this study. A Genome Wide Association Study (GWAS) has been successfully used to identify QTLs and candidate genes for plant architecture (Li et al, 2016; He et al, 2017), flowering traits (Xu et al, 2016; Helal et al, 2021) and disease resistance (Wei et al, 2016; Wu et al, 2016; Roy et al, 2021). Traits related to seed quality and yield have also effectively been explored using GWAS methods in *B. napus* (Li et al, 2014; Gajardo et al, 2015; Körber et al, 2016; Dong et al, 2018; Khan et al, 2019; Pal et al, 2021).

A GWAS was used to locate regions in the genome associated with differences in TGW and SPP. Data from the phenotyping experiment detailed in Chapter 2 was used. As twelve separate temperature treatments were used in the phenotyping experiment, the data was analysed separately for each treatment. Comparative data was also used, such as the ratio of TGW compared between 18°C and 24°C maturation temperature in plants which underwent the same vernalisation treatment. Along with the GWAS analysis, a

meta-analysis was also produced using the full dataset to identify any additional candidate genes.

Several associations with SNPs and TGW/SPP were identified that had nearby genes with plausible developmental function which could be affecting seed traits. These were narrowed down to three main candidate genes which were then studied further to identify if they were likely to be the genetic basis for the SNP association, and what their role in determining TGW in response to temperature might be. The genes selected for further study were *IKU2, AN3/GIF1*, and *GDSL1*. *IKU2* and *AN3* both have known roles in determining seed size (Garcia et al, 2005; Meng et al, 2016), and *GDSL1* has also been associated with changes in seed size in *Brassica napus* (Ding et al, 2019a).

8.1.4 Exploring the genetic basis of temperature responses in *Brassica napus* seeds

In this project, genetic differences between varieties were identified with were associated with changes in TGW and SPP related to temperature, with candidate genes assigned to each. This chapter aimed to test if these candidate genes contained genetic differences which could be causing the effects seen on seeds. To achieve this, the candidate genes were sequenced from *Brassica napus* varieties with contrasting SNP calls.

IKU2 is a gene which is known to affect seed size via changes in endosperm cellularization (Luo et al, 2005). Sequencing *IKU2* in *Brassica napus* did not reveal any genetic differences which were consistently linked to the SNP call identified in the GWAS, or which were likely to impact the function of *IKU2* between variety groups. This could mean that *IKU2* was not the gene responsible for the GWAS association. Alternatively, differences may exist between *IKU2* genes in different varieties which were not uncovered during this study. No link was found here between *IKU2* sequence and differences in temperature response.

AN3 is a gene which affects seed number per silique in Arabidopsis and has been shown to regulate seed size through control of seed expansion (Meng et al, 2016). *AN3* was first

identified as a candidate gene here through associative transcriptomics suggesting differences in *AN3* were related to changes in SPP. ANOVA was also used to show that varieties with different SNP calls in the region associated with *AN3* showed differences in TGW. Sequencing *AN3* revealed a deletion in the 5' untranslated region of the gene, which was found mainly in varieties with one particular SNP call. However, further analysis could not associate this deletion with differences in temperature responses. Future experiments to test *AN3* expression between different haplotype groups could reveal more about if this deletion had any effect in seeds.

GDSL1 is a GDSL-family lipase, which has been previously shown to have a role in oil content and seed size in *Brassica napus* (Ding et al, 2019). Sequence data for *GDSL1* was available from previously created exome capture data (Woodhouse et al, 2021), which allowed comparison of the gene across the *Brassica napus* DFFS. Two main haplotypes of GDSL1 were identified, which showed differences in response to temperature. The first haplotype was similar to the *Brassica napus* Darmor reference sequence. Varieties with this haplotype produced seeds with comparatively high TGW, and seed size was sensitive to temperature. This haplotype was also associated with a strong link between TGW and seed oil content. The second haplotype contained a number of SNPs relative to the Darmor reference sequence, as well as deletions towards the 3' end of the gene. Varieties with this haplotype produced seeds which had a comparatively lower TGW, but TGW was not affected by changes in seed maturation temperature. Varieties with this haplotype also did not show a strong link between TGW and seed oil content.

Exploration of these candidate genes has revealed that there may be a link between different *GDSL1* haplotypes and both seed oil content and stability of seed size in increasing seed maturation temperature. Further experiments to determine the role of *GDSL1* and how the relationship between *GDSL1* and seed size functions could provide an interesting basis for producing higher-yielding OSR crops.

8.1.5 Effects of vernalisation on gene expression and development

While exposure to cold is known to affect the expression of vernalisation-related genes such as *FLC, VIN3* and *VRN1* (Whittaker and Dean, 2017), this chapter aimed to determine the impact of gene expression changes resulting from vernalisation on the seed.

An experiment was carried out which tested the hypotheses that duration of vernalisation or the temperature during floral transition would impact gene expression at the shoot apex and the gynoecia. In this case, a Semi-winter OSR variety was vernalised for either 3 or 12 weeks. Gene expression was examined weekly throughout the vernalisation treatment and afterwards during reproductive development. This experiment showed that Semi-winter OSR would undergo the floral transition in the cold when exposed to a long-term vernalisation treatment. Similar developmental timelines have been identified in Winter OSR (O'Neill et al, 2019).

To examine the effect of vernalisation temperature on development, hypotheses were tested that differences in temperature during vernalisation would cause differences in gene expression and phenotypic differences in the shoot apex, gynoecia, and seed of a Winter OSR variety. To accomplish this, *Brassica napus* plants were given vernalisation treatments at either 5°C or 15°C before completing development under the same conditions. Gene expression was examined through the vernalisation treatment, the floral transition, and through reproductive development of the gynoecia and seed. Phenotypic differences in the gynoecia and seed were also quantified. This enabled gene expression differences between plants vernalised at different temperatures to be linked to changes in reproductive development.

The results of this experiment indicated that ovules were larger in the gynoecia of plants vernalised at 15°C compared to those vernalised at 5°C. However, when looking at the developing seed, by 21 days after pollination (DAP) plants which had been vernalised at 5°C produced larger seeds. These phenotypic differences could then be linked to changes in the transcriptome, which indicated differences in gene expression in the seed at 14DAP. GO term analysis suggested that changes in auxin regulation and ABA gene regulatory pathways were present between seeds vernalised at different temperatures. The nine *FLC* genes in *Brassica napus*, along with a number of other genes associated

with vernalisation responses such as *MAF3, VEL1* and *VIN3* displayed differential regulation throughout the developmental timepoints analysed. It is possible that differences in *FLC* regulation in the endosperm resulting from differences in vernalisation temperature are related to differences in seed size seen at 21DAP. These differences could also occur through changes in auxin and ABA regulation within the seed. Further experiments to determine how changes in seed development, gene expression and hormone profiling between 14DAP and 21DAP could reveal more about how these processes may be occurring and increase our understanding of how vernalisation affects seed development.

8.2 Implications of Research & Limitations

A key limitation of the phenotyping experiment which provided the basis for much of the data and further work carried out in this project was the size of the diversity panel used to produce phenotyping data. Many studies use larger panels for GWAS (Shi et al, 2009; Basunanda et al, 2010; Ding et al, 2012); the panel of 96 varieties used here was relatively small and contained mainly Winter OSR varieties. Due to the combination of twelve temperature treatments used, each variety in the panel added 36 plants to be grown. As such, the size of panel was limited by the number of plants which could be grown in the glasshouse facilities available. In the results described in this thesis, associative transcriptomics revealed almost no statistically significant associations with TGW and SPP. Instead, associations were identified manually using the most significant regions available from the data. Use of a larger panel with a greater diversity of Spring and Semi-winter OSR may have yielded clearer results from associative transcriptomics.

Throughout this project, the effect of vernalisation temperature as well as seed maturation temperature on seeds was explored. Vernalisation temperature was seen to affect seed number and seed size. However, GWAS produced limited evidence for a genetic basis behind this change, with most GWAS SNPs showing differences due to maturation temperature. In the phenotyping, maturation temperature also had a larger effect. This may imply that differences in seed traits resulting from vernalisation temperature are due to gene expression differences, rather than genetic differences. This

corresponds to the differences in *FLC* expression identified in the seed. . Additionally, the results of the GWAS may have been improved by using alternative methods for multiple test corrections or to remove highly correlated markers, which would have improved the statistical threshold used for identifying significant markers. This may have allowed for more statistically robust selection of genes compared to manual identification of peaks on GWAS Manhattan plots. Further exploration of the impact of gene expression on seed traits could have been quantified using GEM analysis; however, the data produced from the phenotyping experiment was not sufficient to show any significant effects from GEM analysis.

More specific sequencing data may also have improved identification and analysis of candidate genes. The SNP array used for GWAS was generated from transcriptomic data using leaf samples; this meant that a wider range of SNPs from non-coding regions and genes not expressed in the leaf were not available to analyse. Although the data provided had good coverage and enabled identification of genetic regions of interest, it may have been possible to easily identify specific candidate genes affecting seed traits by using a SNP array from whole genome sequencing or from seed transcriptomics. Confirmation of candidate genes which were linked to the SNP associations identified from GWAS could also have been improved using linkage disequilibrium analysis to ensure that candidate genes which were followed up were plausible candidates. Despite this lack of region specificity, candidate genes were identified close to a number of the identified SNPs.

In GWAS analysis, *IKU2* and *GDSL1* were identified as candidate genes controlling seed size in response to seed maturation temperature. These genes showed no significant difference in TGW between vernalisation temperatures. They also did not appear as differentially expressed in the vernalisation transcriptome comparison experiment. The same was true when analysing *AN3,* which did not show differences in SPP or TGW between vernalisation treatments. This suggests that the effect of these genes is based on seed maturation temperature rather than resulting from the effect of vernalisation.

AtGDSL1 and *BnGDSL1* overexpression lines showed an increased seed germination rate and improved seedling establishment compared with wild type (Ding et al, 2019a).

In this project, different haplotypes of *GDSL1* were also shown to affect TGW, responses to seed maturation temperature, and oil content. The *GDSL1* haplotype 1, which was similar to the Darmor reference sequence, appears to be required for increases in oil content to result in increased seed size; meanwhile, this relationship is lost in haplotype 2. Future experiments could also explore if these haplotypes show differences in germination rate and seedling establishment, and if these differences related to seed size and oil content. Preliminary analysis to answer this question could be carried out using the results of the germination study detailed in Chapter 3 and the genetic analysis of *GDSL1* detailed in Chapter 5, although including a range of varieties to assess these traits specifically would produce more useful results. Haplotype 2 was associated with improved stability of TGW across maturation temperatures; yield stability is a key trait in OSR and establishing a relationship between *GDSL1* haplotypes and yield stability could be an important contributor to future breeding of *Brassica napus* crops.

Changes in ABA gene regulatory pathways in seeds based on vernalisation temperature were identified from analysis of the transcriptome analysis in Winter OSR. Although the effect of seed vernalisation temperature on germination was not tested during this project, it is possible that changes in ABA levels could results in changes to seed germination and seedling vigour. Further experiments could be carried out to compare ABA levels between seeds produced from plants under different vernalisation temperatures. Any seeds identified with differences in ABA level could be compared using germination assays, as well as compared by TGW to examine the relationship between ABA levels and seed size.

Crop types with different vernalisation requirements were found to have different responses to vernalisation temperature when looking at TGW. Spring OSR and Semi-winter OSR, which have little to no vernalisation requirement, were resistant to reduced TGW resulting from warmer vernalisation temperature. In Winter OSR, where there is a strong vernalisation requirement, increased vernalisation temperature significantly reduced TGW. This may indicate that the effect of vernalisation temperature is mediated through differences in *FLC* dynamics. This is also supported by the differences in *FLC* expression seen in the seeds of a Winter OSR variety when vernalised at two different vernalisation temperatures. Further experiments testing differences in *FLC* expression

in Spring and Semi-winter OSR varieties could reveal more about whether *FLC* dynamics in vernalisation affect seed size.

The transcriptomic experiment design to test if the duration of vernalisation had an effect on gene expression during reproductive development in a Semi-winter OSR variety was designed and started before further results became available. The first result which had an impact on this analysis was the finding that in Semi-winter OSR, as well as in Winter OSR, the floral transition occurs in the cold. This added an additional layer to the experiment where any effects of vernalisation duration could not be separated from effects of temperature during the floral transition. Additionally, the results from the phenotyping experiment showed no effect of vernalisation duration on TGW or SPP in Zhongshuang 11. Seed size in Semi-winter OSR was found to be stable across all three vernalisation temperatures studied in the large glasshouse phenotyping experiment, so it is plausible that there was also little or no effect on seed size based on temperature during the floral transition.

The relationship between ovule size and seed size identified from phenotyping the Winter OSR variety of Catana may suggest that improvements in ovule development inversely relate to seed size. It could be possible that increased ovule growth leads to thicker integuments which exert a greater tensile pressure on the developing seed, causing a reduction in seed expansion (Creff et al, 2015). Further experiments could be carried out to test the hypothesis that there is a relationship between ovule size and integument tensile stress, as well as if this impacts seed size.

The effect of vernalisation temperature on seed size was considered, showing that there was differential expression of vernalisation genes such as *FLC* and *MAF3* in the seed. These genes are differentially regulated as a result of differences in vernalisation temperature and are seen through development including through to the seed. This suggests that the effect of vernalisation on seed size is not necessarily a result of some unknown genetic pathway or genes which control seed size also responding to vernalisation temperature; more likely it is due to the impact of differential *FLC* expression within specific seed tissues. Further experiments could be carried out to test

the hypothesis that *FLC* expression in the seed coat or endosperm is linked to changes in seed size.

8.3 Concluding Remarks

The overall aim of this PhD project was to expand upon existing knowledge of *Brassica napus* development to learn more about how temperature affects development and seed traits.

A phenotyping experiment successfully identified effects of temperature on seed size and seed number per pod in *Brassica napus*, and I was able to use statistical modelling to quantify the extent of these effects and how they differed between varieties and crop types. This provides an in-depth, novel analysis of how seed traits in *Brassica napus* are affected by temperatures, particularly within ranges which are likely to be experienced in the UK.

The data from this experiment then provided a basis for using associative transcriptomics to identify candidate genes. Although GWAS has been used to study seed traits in Brassica previously (Li et al, 2014; Gajardo et al, 2015; Körber et al, 2016; Dong et al, 2018; Khan et al, 2019; Pal et al, 2021), and other recent studies have used phenotypic methods to assess the effect of temperature on reproductive development (Mácová et al, 2021; Magno Massuia de Almeida et al, 2021), the results of this analysis provide a unique insight into the genetic basis of how temperature affects reproductive development and seed traits in *Brassica napus*. Candidate genes were identified and explored, and a possible mechanism involving *GDSL1* in development responses to seed maturation temperature was identified.

The impact of vernalisation on reproductive development was also quantified using phenotypic and transcriptomic approaches. The data produced here included an examination of how contrasting vernalisation temperatures cause differences in gene expression not only in the shoot apex but through reproductive development in the gynoecia and the developing seed. These transcriptomic changes were also linked to phenotypic changes, with changes in ovule size and contrasting differences in seed size

linked to shifts in vernalisation temperature. The results of these experiments showed that warmer vernalisation temperatures led to increased seed size in a process which occurred after fertilization.

Some key questions have been raised by the results of experiments described in this thesis. Although this thesis describes many- the mechanism by which vernalisation temperature affects seed size in Winter OSR is not yet known. It is apparent that in Winter OSR, increased vernalisation temperature causes reductions in seed size; it is also apparent that several key genetic pathways are differentially expressed in seeds produced from plants vernalised at different temperatures. Differential expression was identified in ABA and auxin genes; however, their role in determining seed size in response to temperature remains unclear. Additionally, it is apparent that *FLC* is differentially regulated within the seed as a result of differences in vernalisation temperature, raising the possibility that *FLC* is impacting seed size. However, further experiments would need to be carried out to learn what mechanism could be behind this.

Analysis of the genetic basis of seed size response to maturation temperature also revealed that there is a relationship between *GDSL1* haplotype, seed size and oil content. However, this relationship and its implications for crop performance has not been fully characterized here. The *Brassica napus GDSL1* gene has been previously expressed in Arabidopsis (Ding et al, 2019a), suggesting that a functional *GDSL1* gene from *Brassica napus* should be able to restore function to *gdsl1* mutants in *Arabidopsis*. Expressing the two main *GDSL1* haplotypes identified in *Brassica napus* in Arabidopsis *gdsl1* knockout mutants could be used to compare seed phenotypes and show see if the two haplotypes are functionally different. If differences in function could not be identified, the action of *GDSL1* may instead be due to differences in gene expression levels between haplotypes, which could be tested using qPCR.

In summary, the work described in this thesis has provided greater insight into how temperatures during vernalisation and seed maturation affect seed traits, as well as producing new hypotheses which could be tested in the future to improve our knowledge in this area.

Appendices

Appendix A: Supplementary Tables for Chapter 1

Table A1: List of varieties in the Brassica napus diversity set used in glasshouse phenotyping experiment.

Line	Crop Type
ABUKUMANATANE	Exotics
Altasweet	Swede
Apex	Winter OSR
APEX-93_5XGINYOU_3	Winter OSR
Aphidresistantrape	Winter fodder
Baltia	Winter OSR
BIENVENUDH4	Winter OSR
BRAUNERSCHNITTKOHL	Exotics
Bronowski	Spring OSR

Cabernet	Winter OSR
Cabriolet	Winter OSR
CANARD	Winter fodder
CANBERRAXCOURAGE	Winter OSR
Capitol	Winter OSR
Castille	Winter OSR
Catana	Winter OSR
CeskaKrajova	Spring OSR
CHEMBEREDZAGUMHANA	Exotics
Chuanyou2	Semiwinter OSR
Coriander	Winter OSR
COUVENABICA	Exotics
CRESOR	Spring OSR
CUBSROOT	Spring OSR
Dimension	Winter OSR
Dippes	Winter OSR
Drakkar	Spring OSR
Duplo	Spring OSR
DwarfEssex	Winter fodder
EnglishGiant	Winter fodder
ERGLU	Spring OSR
EUROL	Winter OSR
Excalibur	Winter OSR
Expert	Winter OSR
Flash	Winter OSR
GROENEGRONINGERSNIJMOES	Exotics
HANSENXGASPARD	Winter OSR
HELIOS	Spring OSR
HUGUENOT	Swede
HuronxNavajo	Winter OSR
IncaxContact	Winter OSR
JanetzkisSchlesischer	Winter OSR
JAUNEACOLLETVERT	Swede

KARAT	Spring OSR
KAROO-057DH	Spring OSR
Kromerska	Winter OSR
LEMBKESMALCHOWER(LENORA)	Winter OSR
Lesira	Winter OSR
LICROWNXEXPRESS	Winter OSR
Liho	Spring fodder
MADRIGALXRECITAL	Winter OSR
Matador	Winter OSR
MAZOWIECKI	Spring OSR
MOANAMOANARAPE	Winter fodder
MONTY-028DH	Spring OSR
N01D-1330	Spring OSR
N02D-1952	Spring OSR
Ningyou7	Semiwinter OSR
NORIN	Winter OSR
Palmedor	Winter OSR
POH285Bolko	Winter OSR
Q100	Exotics
Quinta	Winter OSR
RAFALDH1	Winter OSR
RAGGEDJACK	Exotics
Ramses	Winter OSR
RAPIDCYCLINGRAPE(CrGC5)	Exotics
Rocket	Winter OSR
Samourai	Winter OSR
SENSATIONNZ	Swede
SHANNONXWINNER	Winter OSR
Shengliyoucai	Semiwinter OSR
SIBERISCHEBOERENKOOL	Exotics
SLAPSKASLAPY	Exotics
SlovenskaKrajova	Winter OSR
STELLARDH	Spring OSR

SURPASS400-024DH	Spring OSR
SWUChinese1	Semiwinter OSR
SWUChinese2	Semiwinter OSR
Taisetsu	Exotics
TANTAL	Spring OSR
TAPIDORDH	Winter OSR
Temple	Winter OSR
Tina	Swede
Topas	Spring OSR
Tribune	Spring OSR
Verona	Winter OSR
VIGEDH1	Swede
Vision	Winter OSR
WEIHENSTEPHANER	Spring OSR
WESTAR	Spring OSR
Wilhelmsburger	Swede
Willi	Spring OSR
Xiangyou15	Semiwinter OSR
YORK	Swede
ZhongshuangII	Semiwinter OSR
Zhouyou	Semiwinter OSR

Appendix B: Supplementary Tables for Chapter 4

Table B8.1: Table showing genes within 150kb of SNP marker in BnaA01g25880D associated with TGW on chromosome A01. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	**Arabidopsis orthologue**	**Annotated gene function**
Cab026548		
Cab026549		
Cab026550		
Cab026551	AT3G19553.1	Amino acid permease family protein
Cab026552	AT3G19570.2	Family of unknown function (DUF566)
Cab026553		
Cab026554	AT3G19590.1	Transducin/WD40 repeat-like superfamily protein
Cab026555	AT3G19595.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
Cab026556	AT3G19615.1	
Cab026557	AT3G19640.1	magnesium transporter 4
Cab026558		
Cab026559		

Cab026560		
Cab026561		
Cab026562	AT3G19680.1	Protein of unknown function (DUF1005)
Cab026563	AT3G19700.1	Leucine-rich repeat protein kinase family protein haiku2 Iku2
Cab026564	AT3G19720.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Cab026565	AT3G19740.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Cab026566	AT3G19760.1	eukaryotic initiation factor 4A-III
Cab026567	AT3G19800.1	Protein of unknown function (DUF177)
Cab026568	AT3G19820.3	cell elongation protein / DWARF1 / DIMINUTO (DIM)
Cab026569		
Cab026570		
Cab026571		
Cab026572	AT3G19870.1	
Cab026573	AT3G19895.1	RING/U-box superfamily protein
Cab026574	AT3G19900.1	
Cab026575	AT3G19910.1	RING/U-box superfamily protein
Cab026576		
Cab026577	AT3G19930.1	sugar transporter 4
Cab026578	AT3G19940.1	Major facilitator superfamily protein
Cab026579	AT3G19940.1	Major facilitator superfamily protein
Cab026580	AT3G19940.1	Major facilitator superfamily protein
Cab026581	AT3G19940.1	Major facilitator superfamily protein
Cab026582	AT3G19940.1	Major facilitator superfamily protein
Cab026583		
Cab026584	AT3G19990.1	
Cab026585	AT3G20000.1	translocase of the outer mitochondrial membrane 40
Cab026586		
Cab026587	AT3G20010.1	SNF2 domain-containing protein / helicase domain-containing protein / zinc finger protein-related
Cab026588	AT3G20050.1	T-complex protein 1 alpha subunit
Cab026589	AT3G20060.1	ubiquitin-conjugating enzyme19

Table B8.2: Table showing genes within 150kb of SNP marker in BnaA03g26160D associated with TGW on chromosome A03. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	Arabidopsis orthologue	Annotated gene function
Cab003775	AT4G00620.1	Amino acid dehydrogenase family protein
Cab003776	AT4G02880.2	
Cab003777	AT4G02850.1	phenazine biosynthesis PhzC/PhzF family protein
Cab003778	AT4G02850.1	phenazine biosynthesis PhzC/PhzF family protein
Cab003779	AT4G02830.1	
Cab003780		
Cab003781	AT4G02790.1	GTP-binding family protein
Cab003782	AT4G02780.1	Terpenoid cyclases/Protein prenyltransferases superfamily protein
Cab003783		
Cab003784	AT4G02740.1	F-box/RNI-like superfamily protein
Cab003785		
Cab003786	AT1G56220.4	Dormancy/auxin associated family protein
Cab003787		
Cab003788		
Cab003789	AT4G02680.1	ETO1-like 1
Cab003790	AT1G03060.1	Beige/BEACH domain ;WD domain, G-beta repeat protein
Cab003791	AT4G02660.1	Beige/BEACH domain ;WD domain, G-beta repeat protein
Cab003792		
Cab003793	AT4G02590.2	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
Cab003794	AT4G02580.1	NADH-ubiquinone oxidoreductase 24 kDa subunit, putative
Cab003795	AT4G02570.4	cullin 1
Cab003796	AT2G02930.1	glutathione S-transferase F3
Cab003797	AT2G02930.1	glutathione S-transferase F3
Cab003798	AT4G02500.1	UDP-xylosyltransferase 2
Cab003799		
Cab003800	AT4G02480.1	AAA-type ATPase family protein
Cab003801	AT4G02450.2	HSP20-like chaperones superfamily protein
Cab003802	AT4G02410.1	Concanavalin A-like lectin protein kinase family protein
Cab003803	AT4G02420.1	Concanavalin A-like lectin protein kinase family protein
Cab003804	AT4G02405.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Cab003805	AT4G02390.1	poly(ADP-ribose) polymerase
Cab003806		
Cab003807		
Cab003808	AT4G02230.1	Ribosomal protein L19e family protein
Cab003809		

Cab003810		
Cab003811	AT4G02170.1	
Cab003812	AT4G02150.1	ARM repeat superfamily protein
Cab003813	AT1G02720.2	galacturonosyltransferase 5
Cab003814	AT2G25920.1	
Cab003815		
Cab003816	AT4G02110.1	transcription coactivators
Cab003817		
Cab003818	AT4G19750.1	Glycosyl hydrolase family protein with chitinase insertion domain
Cab003819	AT4G02100.1	Heat shock protein DnaJ with tetratricopeptide repeat
Cab003820		

Table B8.3: Table showing genes within 150kb of SNP marker in BnaA03g28950D associated with TGW on chromosome A03. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	**Arabidopsis orthologue**	**Annotated gene function**
Cab004031	AT3G04510.1	Protein of unknown function (DUF640)
Cab004032	AT3G04620.1	Alba DNA/RNA-binding protein
Cab004033	AT3G04610.1	RNA-binding KH domain-containing protein
Cab004034		
Cab004035	AT3G04605.1	transposable element gene
Cab004036		
Cab004037	AT5G59940.1	Cysteine/Histidine-rich C1 domain family protein
Cab004038	AT3G04600.3	Nucleotidylyl transferase superfamily protein
Cab004039	AT3G04630.3	WVD2-like 1
Cab004040	AT3G04650.1	FAD/NAD(P)-binding oxidoreductase family protein
Cab004041	AT3G04670.1	WRKY DNA-binding protein 39
Cab004042	AT3G04680.2	CLP-similar protein 3
Cab004043	AT3G04700.1	Protein of unknown function (DUF1685)
Cab004044	AT3G04710.3	ankyrin repeat family protein
Cab004045		
Cab004046		
Cab004047	AT3G04720.1	pathogenesis-related 4
Cab004048	AT3G04810.2	NIMA-related kinase 2

Cab004049	AT3G04820.1	Pseudouridine synthase family protein
Cab004050	AT3G04830.1	Protein prenylyltransferase superfamily protein
Cab004051	AT3G04840.1	Ribosomal protein S3Ae
Cab004052	AT3G04850.1	Tesmin/TSO1-like CXC domain-containing protein
Cab004053	AT3G04890.1	Uncharacterized conserved protein (DUF2358)
Cab004054	AT3G04910.1	with no lysine (K) kinase 1
Cab004055		
Cab004056		
Cab004057	AT3G04940.1	cysteine synthase D1
Cab004058		
Cab004059		
Cab004060	AT3G04950.1	
Cab004061	AT5G27200.1	acyl carrier protein 5
Cab004062	AT3G05070.1	
Cab004063	AT3G05090.2	Transducin/WD40 repeat-like superfamily protein
Cab004064	AT3G05100.1	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Cab004065	AT3G05140.1	ROP binding protein kinases 2
Cab004066	AT3G05200.1	RING/U-box superfamily protein
Cab004067		
Cab004068	AT3G05220.1	Heavy metal transport/detoxification superfamily protein
Cab004069	AT3G05230.1	Signal peptidase subunit
Cab004070	AT2G28280.1	
Cab004071		
Cab004072		
Cab004073		
Cab004074	AT3G05280.1	Integral membrane Yip1 family protein
Cab004075	AT3G05300.1	Cytidine/deoxycytidylate deaminase family protein
Cab004076		
Cab004077	AT3G05470.1	Actin-binding FH2 (formin homology 2) family protein
Cab004078	AT3G05520.1	Subunits of heterodimeric actin filament capping protein Capz superfamily
Cab004079	AT3G05530.1	regulatory particle triple-A ATPase 5A

Cab004080	AT3G05560.3	Ribosomal L22e protein family
Cab004081	AT3G05580.1	Calcineurin-like metallo-phosphoesterase superfamily protein
Cab004082	AT3G05590.1	ribosomal protein L18
Cab004083		
Cab004084	AT3G05625.1	Tetratricopeptide repeat (TPR)-like superfamily protein
Cab004085	AT3G05670.1	RING/U-box protein
Cab004086	AT5G18350.1	Disease resistance protein (TIR-NBS-LRR class) family
Cab004087		
Cab004088		
Cab004089	AT3G05680.1	embryo defective 2016
Cab004090	AT3G05690.1	nuclear factor Y, subunit A2
Cab004091		

Table B8.4: Table showing genes within 150kb of SNP marker in BnaA08g17820D associated with TGW on chromosome A08. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	**Arabidopsis orthologue**	**Annotated gene function**
Cab039086	AT1G30050.1	
Cab039087	AT1G30040.1	gibberellin 2-oxidase
Cab039088	AT1G30020.1	Protein of unknown function, DUF538
Cab039089	AT1G30000.1	alpha-mannosidase 3
Cab039090	AT2G34480.1	Ribosomal protein L18ae/LX family protein
Cab039091	AT1G29951.1	conserved peptide upstream open reading frame 35
Cab039092	AT1G29930.1	chlorophyll A/B binding protein 1
Cab039093	AT1G29850.1	double-stranded DNA-binding family protein
Cab039094	AT1G29820.2	Magnesium transporter CorA-like family protein
Cab039095		
Cab039096	AT1G29790.2	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Cab039097	AT1G29785.1	other RNA
Cab039098	AT1G29780.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
Cab039099	AT1G29770.1	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein

Cab039100		
Cab039101	AT1G29750.2	receptor-like kinase in flowers 1
Cab039102	AT1G29720.1	Leucine-rich repeat transmembrane protein kinase
Cab039103	AT1G29720.1	Leucine-rich repeat transmembrane protein kinase
Cab039104	AT1G29690.1	MAC/Perforin domain-containing protein
Cab039105	AT1G29670.1	GDSL-like Lipase/Acylhydrolase superfamily protein
Cab039106	AT1G29660.1	GDSL-like Lipase/Acylhydrolase superfamily protein
Cab039107	AT4G15570.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Cab039108	AT1G29640.1	Protein of unknown function, DUF584
Cab039109		
Cab039110	AT1G29470.2	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
Cab039111	AT1G29380.1	Carbohydrate-binding X8 domain superfamily protein
Cab039112	AT1G29370.1	Kinase-related protein of unknown function (DUF1296)
Cab039113	AT1G29340.1	plant U-box 17
Cab039114	AT1G29330.1	ER lumen protein retaining receptor family protein
Cab039115	AT1G29320.1	Transducin/WD40 repeat-like superfamily protein
Cab039116	AT1G29300.1	Plant protein of unknown function (DUF641)
Cab039117		
Cab039118	AT1G29280.1	WRKY DNA-binding protein 65
Cab039119	AT2G32090.2	Lactoylglutathione lyase / glyoxalase I family protein
Cab039120	AT1G29230.1	CBL-interacting protein kinase 18
Cab039121	AT1G29240.1	Protein of unknown function (DUF688)
Cab039122		
Cab039123		
Cab039124	AT1G29170.3	SCAR family protein

Table B8.5: Table showing genes within 150kb of SNP marker in BnaA06g31400D associated with TGW on chromosome A06. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	Arabidopsis orthologue	Annotated gene function
Cab035487	AT2G03060.2	AGAMOUS-like 30
Cab035488		

Cab035489		
Cab035490	AT2G02990.1	ribonuclease 1
Cab035491	AT2G02970.1	GDA1/CD39 nucleoside phosphatase family protein
Cab035492	AT2G02960.3	RING/FYVE/PHD zinc finger superfamily protein
Cab035493	AT1G06220.1	Ribosomal protein S5/Elongation factor G/III/V family protein
Cab035494		
Cab035495	AT2G02950.1	phytochrome kinase substrate 1
Cab035496	AT2G02870.3	Galactose oxidase/kelch repeat superfamily protein
Cab035497	AT2G02860.1	sucrose transporter 2
Cab035498		
Cab035499	AT2G02820.2	myb domain protein 88
Cab035500	AT2G02800.2	protein kinase 2B
Cab035501	AT2G02790.1	IQ-domain 29
Cab035502		
Cab035503		
Cab035504	AT2G02780.1	Leucine-rich repeat protein kinase family protein
Cab035505	AT2G02720.1	Pectate lyase family protein
Cab035506	AT2G02710.1	PAS/LOV protein B
Cab035507	AT2G02590.1	
Cab035508	AT2G02570.4	nucleic acid binding;RNA binding
Cab035509	AT2G02560.2	cullin-associated and neddylation dissociated
Cab035510	AT2G02480.1	AAA-type ATPase family protein
Cab035511	AT2G02470.2	alfin-like 6
Cab035512	AT2G02450.2	NAC domain containing protein 35
Cab035513	AT2G02390.1	glutathione S-transferase zeta 1
Cab035514	AT2G02230.1	phloem protein 2-B1
Cab035515	AT2G02170.2	Remorin family protein
Cab035516		
Cab035517		
Cab035518	AT2G02150.1	Tetratricopeptide repeat (TPR)-like superfamily protein
Cab035519	AT2G02100.1	low-molecular-weight cysteine-rich 69
Cab035520	AT2G02070.1	indeterminate(ID)-domain 5
Cab035521	AT2G02061.1	Nucleotide-diphospho-sugar transferase family protein
Cab035522		
Cab035523	AT2G02050.1	NADH-ubiquinone oxidoreductase B18 subunit, putative

Table B8.6: Table showing genes within 150kb of SNP marker in Bo2g024660 associated with TGW on chromosome C02. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	Arabidopsis orthologue	Annotated gene function
Bo2g024480	AT5G22860.2	Serine carboxypeptidase S28 family protein
Bo2g024490	AT5G22860.2	Serine carboxypeptidase S28 family protein
Bo2g024500		
Bo2g024510	AT5G22870.1	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
Bo2g024520		
Bo2g024530	AT5G22890.1	C2H2 and C2HC zinc fingers superfamily protein
Bo2g024540	AT5G22890.1	C2H2 and C2HC zinc fingers superfamily protein
Bo2g024550	AT5G22900.1	cation/H+ exchanger 3
Bo2g024560	AT5G22950.1	SNF7 family protein
Bo2g024570	AT5G22980.1	serine carboxypeptidase-like 47
Bo2g024580		
Bo2g024590	AT5G23000.1	myb domain protein 37
Bo2g024600		
Bo2g024610	AT5G60780.1	nitrate transporter 2.3
Bo2g024620		
Bo2g024630	AT5G60760.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein
Bo2g024640	AT5G60750.1	CAAX amino terminal protease family protein
Bo2g024650	AT5G60730.1	Anion-transporting ATPase
Bo2g024660	AT5G60710.1	Zinc finger (C3HC4-type RING finger) family protein
Bo2g024670	AT5G60700.1	glycosyltransferase family protein 2

Bo2g024680	AT5G60690.1	Homeobox-leucine zipper family protein / lipid-binding START domain-containing protein
Bo2g024690		
Bo2g024700		
Bo2g024710	AT5G06900.1	cytochrome P450, family 93, subfamily D, polypeptide 1
Bo2g024720	AT2G46520.1	cellular apoptosis susceptibility protein, putative / importin-alpha re-exporter, putative
Bo2g024730		
Bo2g024740		
Bo2g024750		
Bo2g024760	AT5G60660.1	plasma membrane intrinsic protein 2;4
Bo2g024770	AT5G60640.1	PDI-like 1-4
Bo2g024780	AT5G65010.2	asparagine synthetase 2

Table B8.7: Table showing genes within 150kb of SNP marker in BnaC06g21160D associated with TGW on chromosome C06. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	Arabidopsis orthologue	Annotated gene function
Bo6rg119680	AT1G75520.1	SHI-related sequence 5
Bo6rg119690		
Bo6rg119700	AT1G75540.1	salt tolerance homolog2
Bo6rg119710	AT1G75550.1	glycine-rich protein
Bo6rg119720	AT1G75560.2	zinc knuckle (CCHC-type) family protein
Bo6rg119730		
Bo6rg119740	AT1G75590.1	SAUR-like auxin-responsive protein family
Bo6rg119750	AT1G75620.1	glyoxal oxidase-related protein
Bo6rg119760	AT1G75640.1	Leucine-rich receptor-like protein kinase family protein
Bo6rg119770	AT1G75660.1	5'-3' exoribonuclease 3
Bo6rg119780	AT1G75680.1	glycosyl hydrolase 9B7
Bo6rg119790	AT1G75690.1	DnaJ/Hsp40 cysteine-rich domain superfamily protein
Bo6rg119800	AT1G75700.1	HVA22-like protein G
Bo6rg119810	AT1G75710.1	C2H2-like zinc finger protein

Bo6rg119820	AT1G75720.2	Plant protein of unknown function (DUF827)
Bo6rg119830		
Bo6rg119840	AT1G75750.1	GAST1 protein homolog 1
Bo6rg119850	AT1G75760.1	ER lumen protein retaining receptor family protein
Bo6rg119860	AT5G11210.1	glutamate receptor 2.5
Bo6rg119870	AT1G75780.1	tubulin beta-1 chain
Bo6rg119880	AT1G75800.1	Pathogenesis-related thaumatin superfamily protein
Bo6rg119890		
Bo6rg119900	AT1G75820.1	Leucine-rich receptor-like protein kinase family protein
Bo6rg119910	AT1G75830.1	low-molecular-weight cysteine-rich 67
Bo6rg119920	AT1G75840.1	RAC-like GTP binding protein 5
Bo6rg119930	AT1G75850.1	VPS35 homolog B
Bo6rg119940		
Bo6rg119950	AT1G75890.2	GDSL-like Lipase/Acylhydrolase superfamily protein
Bo6rg119960		
Bo6rg119970	AT1G75900.1	GDSL-like Lipase/Acylhydrolase superfamily protein
Bo6rg119980	AT1G75950.1	S phase kinase-associated protein 1
Bo6rg119990		
Bo6rg120000	AT1G75980.1	Single hybrid motif superfamily protein
Bo6rg120010	AT1G75990.1	PAM domain (PCI/PINT associated module) protein
Bo6rg120020	AT1G76020.1	Thioredoxin superfamily protein
Bo6rg120030	AT1G76030.1	ATPase, V1 complex, subunit B protein
Bo6rg120040	AT1G76040.2	calcium-dependent protein kinase 29
Bo6rg120050		
Bo6rg120060	AT2G17090.1	Protein kinase protein with tetratricopeptide repeat domain
Bo6rg120070		
Bo6rg120080		
Bo6rg120090		
Bo6rg120100	AT1G76070.1	
Bo6rg120110	AT1G76080.1	chloroplastic drought-induced stress protein of 32 kD
Bo6rg120120	AT1G76090.1	sterol methyltransferase 3
Bo6rg120130	AT1G76100.1	plastocyanin 1
Bo6rg120140	AT1G76110.1	HMG (high mobility group) box protein with ARID/BRIGHT DNA-binding domain
Bo6rg120150	AT1G76120.1	Pseudouridine synthase family protein
Bo6rg120160	AT1G76130.1	alpha-amylase-like 2

Table B8.8: Table showing genes within 150kb of SNP marker in BnaC09g02780D associated with SPP on chromosome C09. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	Arabidopsis orthologue	Annotated gene function
Bo9g135310	AT5G58560.1	Phosphatidate cytidylyltransferase family protein
Bo9g135320	AT5G58580.1	TOXICOS EN LEVADURA 63
Bo9g135330	AT4G01975.1	transposable element gene
Bo9g135340		
Bo9g135350		
Bo9g135360		
Bo9g135370	AT5G58600.1	Plant protein of unknown function (DUF828)
Bo9g135380	AT5G58610.1	PHD finger transcription factor, putative
Bo9g135390		
Bo9g135400	AT5G58620.1	zinc finger (CCCH-type) family protein
Bo9g135410	AT5G58630.1	
Bo9g135420	AT5G58640.1	Selenoprotein, Rdx type
Bo9g135430		
Bo9g135440		
Bo9g135450	AT5G58660.1	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
Bo9g135460		
Bo9g135470	AT5G58670.1	phospholipase C1
Bo9g135480	AT5G58680.1	ARM repeat superfamily protein
Bo9g135490	AT5G58690.1	phosphatidylinositol-speciwc phospholipase C5
Bo9g135500	AT5G58700.1	phosphatidylinositol-speciwc phospholipase C4
Bo9g135510	AT5G58710.1	rotamase CYP 7

Bo9g135520		
Bo9g135530	AT5G58740.1	HSP20-like chaperones superfamily protein
Bo9g135540	AT5G58750.1	NAD(P)-binding Rossmann-fold superfamily protein
Bo9g135550		
Bo9g135560	AT5G58770.1	Undecaprenyl pyrophosphate synthetase family protein
Bo9g135570	AT5G58784.1	Undecaprenyl pyrophosphate synthetase family protein
Bo9g135580	AT5G56900.2	CwfJ-like family protein / zinc finger (CCCH-type) family protein
Bo9g135590		
Bo9g135600	AT5G58770.1	Undecaprenyl pyrophosphate synthetase family protein
Bo9g135610	AT5G58784.1	Undecaprenyl pyrophosphate synthetase family protein
Bo9g135620		
Bo9g135630		
Bo9g135640	AT3G15310.1	transposable element gene
Bo9g135650		
Bo9g135660		
Bo9g135670	AT5G59090.1	subtilase 4.12
Bo9g135680	AT5G58870.1	FTSH protease 9
Bo9g135690		
Bo9g135700	AT2G42480.1	TRAF-like family protein
Bo9g135710	AT5G58880.1	
Bo9g135720		

Table B8.9: Table showing genes within 150kb of SNP marker in BnaC01g36120D associated with SPP on chromosome C01. Empty cells indicate that no Arabidopsis orthologue was identified, or no gene function was annotated, within the pangenome.

SNP Marker	Arabidopsis orthologue	Annotated gene function
Bo1g134550		
Bo1g134560	AT3G14200.1	Chaperone DnaJ-domain superfamily protein
Bo1g134570	AT3G14190.1	
Bo1g134580		
Bo1g134590		
Bo1g134600	AT3G14120.2	
Bo1g134610	AT3G14080.2	Small nuclear ribonucleoprotein family protein
Bo1g134620	AT3G14075.2	Mono-/di-acylglycerol lipase, N-terminal;Lipase, class 3
Bo1g134630	AT3G14067.1	Subtilase family protein
Bo1g134640		
Bo1g134650	AT3G14020.1	nuclear factor Y, subunit A6
Bo1g134660	AT3G14010.4	CTC-interacting domain 4
Bo1g134670	AT3G14000.2	DZC (Disease resistance/zinc finger/chromosome condensation-like region) domain containing protein
Bo1g134680	AT3G13990.2	Kinase-related protein of unknown function (DUF1296)
Bo1g134690		
Bo1g134700		
Bo1g134710		
Bo1g134720		
Bo1g134730		
Bo1g134740	AT3G13960.1	growth-regulating factor 5

Bo1g134750		
Bo1g134760		
Bo1g134770	AT3G13920.1	eukaryotic translation initiation factor 4A1
Bo1g134780		
Bo1g134790	AT3G13890.2	myb domain protein 26
Bo1g134800	AT3G13880.1	Tetratricopeptide repeat (TPR)-like superfamily protein
Bo1g134810	AT3G13880.1	Tetratricopeptide repeat (TPR)-like superfamily protein
Bo1g134820		
Bo1g134830	AT3G13850.1	LOB domain-containing protein 22
Bo1g134840		
Bo1g134850		
Bo1g134860		
Bo1g134870		
Bo1g134880		
Bo1g134890		
Bo1g134900		
Bo1g134910		
Bo1g134920	AT3G13810.1	indeterminate(ID)-domain 11
Bo1g134930	AT3G13790.1	Glycosyl hydrolases family 32 protein
Bo1g134940	AT3G13750.1	beta galactosidase 1

Table B8.10: GWAS results showing highest p-values for each association in identified regions of interest for TGW. Left hand column indicates treatments.

Chr	A01	A03	A03	A03	A04	A06	A07	A08	C06	C08
Location	*25.9 - 26.1*	*8.7 - 8.8*	*14.8 - 14.9*	*16.2 - 16.4*	*21.5 - 21.7*	*26.5 - 26.8*	*24.1 - 24.2*	*17.8 - 18*	*36.8 - 37.7*	*25.3 - 25.9*
5°C 6wks 18°C	2.46E-02	1.05E-06	3.93E-03	2.79E-03	2.14E-07	1.50E-02	3.60E-06	8.91E-04	1.01E-04	1.30E-03
5°C 12wks 18°C	4.93E-02	3.94E-04	6.45E-03	5.39E-05	3.52E-05	1.03E-03	4.48E-04	1.28E-03	2.75E-06	5.25E-04
5°C 6wks 24°C	7.18E-02	1.72E-03	5.88E-02	3.20E-03	1.06E-04	6.01E-03	3.01E-04	1.12E-03	6.44E-05	3.68E-02
5°C 12 wks 24°C	1.04E-02	1.66E-03	2.41E-02	1.50E-02	9.22E-03	9.64E-03	4.73E-02	1.07E-02	5.40E-04	2.36E-02
10°C 6wks 18°C	1.79E-03	1.59E-05	4.46E-03	8.71E-03	2.16E-05	7.50E-03	2.05E-05	5.82E-03	8.41E-05	6.51E-04
10°C 12wks 18°C	3.34E-02	6.67E-06	9.27E-03	1.78E-03	7.95E-05	1.03E-02	1.11E-04	2.25E-02	5.07E-06	8.52E-03
10°C 6wks 24°C	2.45E-02	3.77E-02	2.50E-03	3.32E-03	3.97E-04	1.44E-02	1.86E-04	1.20E-02	2.36E-05	3.34E-03
10°C 12 wks 24°C	4.71E-03	2.99E-03	1.04E-02	1.15E-03	1.06E-05	7.47E-03	1.29E-05	3.22E-03	2.33E-04	2.01E-02
15°C 6wks 18°C	7.82E-04	8.45E-03	2.60E-03	4.78E-05	8.31E-04	6.69E-05	5.16E-06	1.84E-02	9.91E-04	4.89E-03
15°C 12wks 18°C	2.38E-02	2.03E-04	2.79E-03	7.95E-03	3.18E-04	4.72E-03	9.52E-05	5.97E-03	1.92E-04	3.47E-04
15°C 6wks 24°C	2.92E-04	1.78E-02	6.44E-02	2.00E-02	5.29E-02	3.41E-02	2.84E-02	5.44E-02	5.70E-03	3.22E-02
15°C 12 wks 24°C	2.64E-02	1.17E-02	2.45E-02	7.06E-03	7.44E-03	1.95E-03	4.58E-03	9.50E-03	3.23E-04	1.44E-02
5°C/15°C, 6wks 18°C	3.57E-03	8.44E-03	5.77E-03	3.74E-03	1.10E-03	4.96E-04	1.01E-02	6.74E-03	1.42E-03	5.97E-04
5°C/15°C, 12wks 18°C	1.06E-03	1.14E-04	8.83E-03	2.92E-04	2.49E-04	8.41E-04	6.01E-03	2.59E-05	1.90E-03	3.88E-07
5°C/15°C, 6wks 24°C	1.96E-02	3.07E-03	1.42E-02	5.14E-03	1.01E-01	3.53E-02	1.26E-01	3.52E-04	7.78E-03	3.86E-02
5°C/15°C, 12wks 24°C	3.20E-05	7.12E-02	4.51E-03	1.63E-02	2.01E-02	1.39E-02	2.94E-01	1.85E-02	1.79E-03	6.31E-03
18°C/24°C, 5°C 6wks	2.79E-03	7.71E-03	2.66E-02	7.87E-03	3.08E-03	6.40E-03	2.84E-02	9.91E-03	1.35E-03	1.87E-02
18°C/24°C, 5°C 12wks	6.09E-02	6.53E-03	4.03E-04	8.02E-04	1.19E-03	1.95E-03	4.02E-02	1.75E-03	1.11E-03	1.63E-03
18°C/24°C, 10°C 6wks	6.37E-03	7.98E-03	3.11E-06	1.02E-02	1.23E-02	6.19E-03	4.47E-03	2.05E-02	8.56E-04	7.31E-03
18°C/24°C, 10°C 12wks	5.32E-02	3.58E-02	1.08E-04	1.99E-03	1.45E-03	2.08E-02	1.97E-02	1.65E-03	8.80E-03	4.96E-03
18°C/24°C, 15°C 6wks	3.47E-03	2.63E-02	2.06E-02	2.84E-03	1.92E-02	4.92E-03	1.77E-02	5.59E-03	7.74E-03	8.99E-03

18°C/24°C, 15°C 12wks	4.22E-02	1.21E-03	1.40E-03	4.67E-03	5.00E-03	4.02E-05	3.80E-02	7.77E-04	4.58E-04	4.69E-03

Table B8.11: GWAS results showing highest p-values for each association in identified regions of interest for SPP. Left hand column indicates treatments.

Chr	A02	A05	A08	C01	C04	C07	C06	C09
Location	*34.6 - 34.7*	*1.5 - 1.6*	*15.4 - 16*	*38.6 - 39.1*	*35.1 - 35.2*	*43.6 - 43.7*	*16.7 - 16.9*	*2.7 - 2.9*
5°C 6wks 18°C	2.17E-04	5.50E-04	1.66E-03	1.80E-03	3.52E-02	8.10E-03	3.70E-03	2.89E-03
5°C 12wks 18°C	2.61E-03	7.30E-03	9.32E-05	8.97E-05	1.69E-01	4.83E-02	5.22E-03	7.35E-05
5°C 6wks 24°C	4.15E-04	1.68E-03	3.38E-04	5.65E-06	1.62E-02	8.75E-02	1.84E-03	1.74E-04
5°C 12 wks 24°C	8.15E-05	1.97E-04	2.80E-03	1.39E-05	4.95E-02	1.35E-02	3.57E-03	8.10E-05
10°C 6wks 18°C	4.81E-03	9.04E-03	4.00E-05	2.21E-04	1.32E-01	3.79E-02	3.30E-03	6.93E-06
10°C 12wks 18°C	3.88E-03	2.65E-03	1.20E-05	5.79E-04	4.65E-02	1.85E-02	3.40E-04	2.29E-04
10°C 6wks 24°C	6.55E-03	2.31E-05	3.32E-04	1.51E-05	8.33E-03	7.12E-02	5.11E-04	4.05E-06
10°C 12 wks 24°C	2.72E-05	2.28E-03	3.76E-03	1.69E-05	5.62E-02	1.02E-01	3.67E-03	7.84E-04
15°C 6wks 18°C	2.74E-02	4.15E-02	2.48E-06	4.39E-04	6.09E-02	2.02E-01	1.60E-04	1.79E-03
15°C 12wks 18°C	3.92E-04	1.04E-02	5.30E-04	9.59E-05	1.86E-02	1.46E-01	7.25E-03	6.13E-05
15°C 6wks 24°C	1.32E-02	4.76E-03	3.54E-03	1.39E-02	4.81E-02	4.59E-03	1.33E-03	3.13E-03
15°C 12 wks 24°C	3.59E-03	1.98E-03	1.25E-03	1.39E-04	7.74E-02	1.29E-01	1.18E-03	1.65E-03
5°C/15°C, 6wks 18°C	2.42E-03	2.48E-02	6.95E-04	4.95E-03	1.78E-04	2.57E-02	5.23E-03	6.76E-04
5°C/15°C, 12wks 18°C	7.00E-02	4.25E-02	3.32E-05	3.37E-04	1.01E-01	8.11E-02	7.80E-03	2.05E-02
5°C/15°C, 6wks 24°C	5.28E-02	1.76E-03	3.44E-03	1.78E-02	1.51E-01	7.84E-02	6.71E-02	2.39E-02
5°C/15°C, 12wks 24°C	2.26E-02	3.22E-02	6.09E-03	1.60E-02	6.24E-03	1.58E-04	9.86E-03	2.17E-02
18°C/24°C, 5°C 6wks	9.37E-02	1.79E-01	1.39E-03	1.86E-03	1.05E-03	1.90E-01	1.87E-02	7.21E-03

18°C/24°C, 5°C 12wks	4.08E-02	7.92E-02	1.53E-03	9.60E-03	1.33E-02	5.14E-04	1.22E-02	2.13E-03
18°C/24°C, 10°C 6wks	1.60E-02	4.52E-03	4.97E-05	2.20E-02	6.03E-02	4.81E-02	1.93E-04	9.26E-04
18°C/24°C, 10°C 12wks	6.50E-02	1.02E-01	1.36E-02	1.27E-02	3.22E-04	1.31E-02	2.01E-02	3.00E-02
18°C/24°C, 15°C 6wks	2.08E-02	2.47E-03	1.66E-04	1.10E-01	1.60E-01	6.28E-02	9.02E-03	1.65E-02
18°C/24°C, 15°C 12wks	6.41E-02	2.14E-02	8.60E-03	2.51E-02	4.04E-02	5.29E-02	4.84E-02	8.17E-02

Glossary

Abbreviation	Definition
ABA	Abscisic acid
DAP	Days after pollination
DFFS	Diversity Fixed Foundation Set
GA	Gibberellic acid
OSR	Oilseed rape
SNP	Single Nucleotide Polymorphism
SPP	Seed number per pod
TGW	Thousand grain weight
Gene Name Abbreviations	
AN3	ANGUSTIFOLIA3
AXR6	AUXIN RESISTANT 6
BRR2A	EMBRYO DEFECTIVE 1507
CAND1	CULLIN-ASSOCIATED AND NEDDYLATION DISSOCIATED 1
CLV1	CLAVATA1
FLC	FLOWERING LOCUS C
GATL8	GALACTURONOSYLTRANSFERASE-LIKE 8
GDSL1	GDSL-LIKE LIPASE-ESTERASE 1
GID1A	GA-INSENSITIVE DWARF 1A
IKU2	HAIKU2
LFY	LEAFY
LNO1	EMBRYO DEFECTIVE 1011
MAF3	MADS AFFECTING FLOWERING 3
MERI-5	MERISTEM 5
MYB26	MYB DOMAIN PROTEIN 26
MYB44	MYB DOMAIN PROTEIN 44
MYB77	MYB DOMAIN PROTEIN 77
PAPP2C	PHYTOCHROME-ASSOCIATED PROTEIN PHOSPHATASE TYPE 2C
PBP1	PYK10 BINDING PROTEIN 1
REV	REVOLUTA

SOC1	SUPPRESSOR OF OVEREXPRESSION OF CO 1
SUS2	ABNORMAL SUSPENSOR 2
TCH4	TOUCH4
TFL1	TERMINAL FLOWER 1
UFO	UNUSUAL FLORAL ORGANS
VEL1	VERNALISATION/VIN3-LIKE 1
VIN3	VERNALISATION INSENSITIVE 3
VRN1	REDUCED VERNALISATION RESPONSE 1
WOX1	WUSCHEL-RELATED HOMEOBOX 1
WOX2	WUSCHEL-RELATED HOMEOBOX 2
WOX4	WUSCHEL-RELATED HOMEOBOX 4
WOX9	WUSCHEL-RELATED HOMEOBOX 9

References

Ahmad, M., Waraich, E. A., Skalicky, M., Hussain, S., Zulfiqar, U., Anjum, M. Z., Habib Ur Rahman, M., Brestic, M., Ratnasekera, D., Lamilla-Tamayo, L., Al-Ashkar, I., & El Sabagh, A. (2021). Adaptation Strategies to Improve the Resistance of Oilseed Crops to Heat Stress Under a Changing Climate: An Overview. *Frontiers in Plant Science*, *12*, 767150. https://doi.org/10.3389/fpls.2021.767150

Akter, A., Itabashi, E., Kakizaki, T., Okazaki, K., Dennis, E. S., & Fujimoto, R. (2021). Genome Triplication Leads to Transcriptional Divergence of FLOWERING LOCUS C Genes During Vernalization in the Genus Brassica. *Frontiers in Plant Science*, *11*, 619417. https://doi.org/10.3389/fpls.2020.619417

Alexandre, C., & Hennig, L. (2007). *FLC-independent vernalization responses*.

Allender, C. J., & King, G. J. (2010). Origins of the amphiploid species Brassica napusL. investigated by chloroplast and nuclear molecular markers. *BMC Plant Biology*, *10*(1), 54. https://doi.org/10.1186/1471-2229-10-54

Alonso-Blanco, C., Blankestijn-de Vries, H., Hanhart, C. J., & Koornneef, M. (1999). Natural allelic variation at seed size loci in relation to other life history traits of Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(8), 4710–4717.

Andalo, C., Mazer, S. J., Godelle, B., & Machon, N. (1999). Parental environmental effects on life history traits in Arabidopsis thaliana (Brassicaceae). *New Phytologist*, *142*(2), 173–184. https://doi.org/10.1046/j.1469-8137.1999.00396.x

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Angelovici, R., Galili, G., Fernie, A. R., & Fait, A. (2010). Seed desiccation: A bridge between maturation and germination. *Trends in Plant Science*, *15*(4), 211–218. https://doi.org/10.1016/j.tplants.2010.01.003

Antoniou-Kourounioti, R. L., Hepworth, J., Heckmann, A., Duncan, S., Qüesta, J., Rosa, S., Säll, T., Holm, S., Dean, C., & Howard, M. (2018). Temperature Sensing Is Distributed throughout the Regulatory Network that Controls FLC Epigenetic Silencing in Vernalization. *Cell Systems*, *7*(6), 643-655.e9. https://doi.org/10.1016/j.cels.2018.10.011

Atwell, S., Huang, Y. S., Vilhjálmsson, B. J., Willems, G., Horton, M., Li, Y., Meng, D., Platt, A., Tarone, A. M., Hu, T. T., Jiang, R., Muliyati, N. W., Zhang, X., Amer, M. A., Baxter, I., Brachi, B., Chory, J., Dean, C., Debieu, M., … Nordborg, M. (2010). Genome-wide association study of 107 phenotypes in a common set of Arabidopsis thaliana inbred lines. *Nature*, *465*(7298), 627–631. https://doi.org/10.1038/nature08800

Auge, G. A., Blair, L. K., Neville, H., & Donohue, K. (2017). Maternal vernalization and vernalization-pathway genes influence progeny seed germination. *New Phytologist*, *216*(2), 388–400. https://doi.org/10.1111/nph.14520

Bac-Molenaar, J. A., Fradin, E. F., Becker, F. F. M., Rienstra, J. A., van der Schoot, J., Vreugdenhil, D., & Keurentjes, J. J. B. (2015). Genome-Wide Association Mapping of Fertility Reduction upon Heat Stress Reveals Developmental Stage-Specific QTLs in

Arabidopsis thaliana. *The Plant Cell*, *27*(7), 1857–1874. https://doi.org/10.1105/tpc.15.00248

Bai, G., Yang, D.-H., Cao, P., Yao, H., Zhang, Y., Chen, X., Xiao, B., Li, F., Wang, Z.-Y., Yang, J., & Xie, H. (2019). Genome-Wide Identification, Gene Structure and Expression Analysis of the MADS-Box Gene Family Indicate Their Function in the Development of Tobacco (Nicotiana tabacum L.). *International Journal of Molecular Sciences*, *20*(20), Article 20. https://doi.org/10.3390/ijms20205043

Balfagón, D., Sengupta, S., Gómez-Cadenas, A., Fritschi, F. B., Azad, R. K., Mittler, R., & Zandalinas, S. I. (2019). Jasmonic Acid Is Required for Plant Acclimation to a Combination of High Light and Heat Stress1 [OPEN]. *Plant Physiology*, *181*(4), 1668–1682. https://doi.org/10.1104/pp.19.00956

Barnabás, B., Jäger, K., & Fehér, A. (2008). The effect of drought and heat stress on reproductive processes in cereals. *Plant, Cell & Environment*, *31*(1), 11–38. https://doi.org/10.1111/j.1365-3040.2007.01727.x

Basnet, R. K., Duwal, A., Tiwari, D. N., Xiao, D., Monakhos, S., Bucher, J., Visser, R. G. F., Groot, S. P. C., Bonnema, G., & Maliepaard, C. (2015). Quantitative Trait Locus Analysis of Seed Germination and Seedling Vigor in Brassica rapa Reveals QTL Hotspots and Epistatic Interactions. *Frontiers in Plant Science*, *6*, 1032. https://doi.org/10.3389/fpls.2015.01032

Basunanda, P., Radoev, M., Ecke, W., Friedt, W., Becker, H. C., & Snowdon, R. J. (2010). Comparative mapping of quantitative trait loci involved in heterosis for seedling and

yield traits in oilseed rape (Brassica napus L.). *Theoretical and Applied Genetics*, *120*(2), 271–281. https://doi.org/10.1007/s00122-009-1133-z

Batista, R. A., Figueiredo, D. D., Santos-González, J., & Köhler, C. (2019). Auxin regulates endosperm cellularization in Arabidopsis. *Genes & Development*, *33*(7–8), 466–476. https://doi.org/10.1101/gad.316554.118

Batool, M., El-Badri, A. M., Hassan, M. U., Haiyun, Y., Chunyun, W., Zhenkun, Y., Jie, K., Wang, B., & Zhou, G. (2022). Drought Stress in Brassica napus: Effects, Tolerance Mechanisms, and Management Strategies. *Journal of Plant Growth Regulation*. https://doi.org/10.1007/s00344-021-10542-9

Beauzamy, L., Fourquin, C., Dubrulle, N., Boursiac, Y., Boudaoud, A., & Ingram, G. (2016). Endosperm turgor pressure decreases during early Arabidopsis seed development. *Development*, *143*(18), 3295–3299. https://doi.org/10.1242/dev.137190

Bentsink, L., Jowett, J., Hanhart, C. J., & Koornneef, M. (2006). Cloning of DOG1, a quantitative trait locus controlling seed dormancy in Arabidopsis. *Proceedings of the National Academy of Sciences*, *103*(45), 17042–17047. https://doi.org/10.1073/pnas.0607877103

Berger, F. (1999). Endosperm development. *Current Opinion in Plant Biology*, *2*(1), 28–32. https://doi.org/10.1016/S1369-5266(99)80006-5

Bergonzi, S., Albani, M. C., Themaat, E. V. L. van, Nordström, K. J. V., Wang, R., Schneeberger, K., Moerland, P. D., & Coupland, G. (2013). Mechanisms of Age-Dependent Response to Winter Temperature in Perennial Flowering of Arabis alpina. *Science*, *340*(6136), 1094–1097. https://doi.org/10.1126/science.1234116

Bernardes, W. S., & Menossi, M. (2020). Plant 3' Regulatory Regions From mRNA-Encoding Genes and Their Uses to Modulate Expression. *Frontiers in Plant Science*, *11*, 1252. https://doi.org/10.3389/fpls.2020.01252

Bethke, P. C., Libourel, I. G. L., Aoyama, N., Chung, Y.-Y., Still, D. W., & Jones, R. L. (2007). The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. *Plant Physiology*, *143*(3), 1173–1188. https://doi.org/10.1104/pp.106.093435

Bettey, M., Finch-Savage, W. E., King, G. J., & Lynn, J. R. (2000). Quantitative genetic analysis of seed vigour and pre-emergence seedling growth traits in Brassica oleracea. *New Phytologist*, *148*(2), 277–286. https://doi.org/10.1046/j.1469-8137.2000.00760.x

Bewley, J. (1997). Seed Germination and Dormancy. *The Plant Cell*, *9*(7), 1055–1066.

Biro, J. (2006). Amino acid size, charge, hydropathy indices and matrices for protein structure analysis. *Theoretical Biology and Medical Modelling*, *3*(1), 15. https://doi.org/10.1186/1742-4682-3-15

Boisnard-Lorig, C., Colon-Carmona, A., Bauch, M., Hodge, S., Doerner, P., Bancharel, E., Dumas, C., Haseloff, J., & Berger, F. (2001). Dynamic analyses of the expression of the HISTONE::YFP fusion protein in arabidopsis show that syncytial endosperm is divided in mitotic domains. *The Plant Cell*, *13*(3), 495–509. https://doi.org/10.1105/tpc.13.3.495

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borg, M., & Berger, F. (2015). Chromatin remodelling during male gametophyte development. *The Plant Journal*, *83*(1), 177–188. https://doi.org/10.1111/tpj.12856

Borg, M., Jacob, Y., Susaki, D., LeBlanc, C., Buendía, D., Axelsson, E., Kawashima, T., Voigt, P., Boavida, L., Becker, J., Higashiyama, T., Martienssen, R., & Berger, F. (2020). Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nature Cell Biology*, *22*(6), Article 6. https://doi.org/10.1038/s41556-020-0515-y

Bouttier, C., & Morgan, D. G. (1992). Ovule Development and Determination of Seed Number Per Pod in Oilseed Rape (Brassica napus L.). *Journal of Experimental Botany*, *43*(5), 709–714. https://doi.org/10.1093/jxb/43.5.709

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics*, *23*(19), 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Bradford, K. J., TARQUIS, A. M., & DURÁN, J. M. (1993). A Population-based Threshold Model Describing the Relationship Between Germination Rates and Seed Deterioration. *Journal of Experimental Botany*, *44*(7), 1225–1234. https://doi.org/10.1093/jxb/44.7.1225

Braud, C., Zheng, W., & Xiao, W. (2012). LONO1 encoding a nucleoporin is required for embryogenesis and seed viability in Arabidopsis. *Plant Physiology*, *160*(2), 823–836. https://doi.org/10.1104/pp.112.202192

Brown, J. K. M., Beeby, R., & Penfield, S. (2019). Yield instability of winter oilseed rape modulated by early winter temperature. *Scientific Reports*, *9*(1), 6953. https://doi.org/10.1038/s41598-019-43461-7

Brunel-Muguet, S., D'Hooghe, P., Bataillé, M.-P., Larré, C., Kim, T.-H., Trouverie, J., Avice, J.-C., Etienne, P., & Dürr, C. (2015). Heat stress during seed filling interferes with sulfur restriction on grain composition and seed germination in oilseed rape (Brassica napus L.). *Frontiers in Plant Science*, *6*. https://www.frontiersin.org/articles/10.3389/fpls.2015.00213

Bušić, A., Kundas, S., Morzak, G., Belskaya, H., Marđetko, N., Ivančić Šantek, M., Komes, D., Novak, S., & Šantek, B. (2018). Recent Trends in Biodiesel and Biogas Production. *Food Technology and Biotechnology*, *56*(2), 152–173. https://doi.org/10.17113/ftb.56.02.18.5547

Cai, D., Xiao, Y., Yang, W., Ye, W., Wang, B., Younas, M., Wu, J., & Liu, K. (2014). Association mapping of six yield-related traits in rapeseed (Brassica napus L.). *Theoretical and Applied Genetics*, *127*(1), 85–96. https://doi.org/10.1007/s00122-013-2203-9

Cai, G., Yang, Q., Yi, B., Fan, C., Edwards, D., Batley, J., & Zhou, Y. (2014). A Complex Recombination Pattern in the Genome of Allotetraploid Brassica napus as Revealed by a

High-Density Genetic Map. *PLoS ONE*, *9*(10), e109910. https://doi.org/10.1371/journal.pone.0109910

Calarco, J. P., Borges, F., Donoghue, M. T. A., Van Ex, F., Jullien, P. E., Lopes, T., Gardner, R., Berger, F., Feijó, J. A., Becker, J. D., & Martienssen, R. A. (2012). Reprogramming of DNA Methylation in Pollen Guides Epigenetic Inheritance via Small RNA. *Cell*, *151*(1), 194–205. https://doi.org/10.1016/j.cell.2012.09.001

Calderwood, A., Lloyd, A., Hepworth, J., Tudor, E. H., Jones, D. M., Woodhouse, S., Bilham, L., Chinoy, C., Williams, K., Corke, F., Doonan, J. H., Ostergaard, L., Irwin, J. A., Wells, R., & Morris, R. J. (2021). Total FLC transcript dynamics from divergent paralogue expression explains flowering diversity in Brassica napus. *New Phytologist*, *229*(6), 3534–3548. https://doi.org/10.1111/nph.17131

Čanak, P., Jeromela, A. M., Vujošević, B., Kiprovski, B., Mitrović, B., Alberghini, B., Facciolla, E., Monti, A., & Zanetti, F. (2020). Is Drought Stress Tolerance Affected by Biotypes and Seed Size in the Emerging Oilseed Crop Camelina? *Agronomy*, *10*(12), Article 12. https://doi.org/10.3390/agronomy10121856

Cao, J., Li, G., Qu, D., Li, X., & Wang, Y. (2020). Into the Seed: Auxin Controls Seed Development and Grain Yield. *International Journal of Molecular Sciences*, *21*(5), 1662. https://doi.org/10.3390/ijms21051662

Cenci, A., Concepción-Hernández, M., Guignon, V., Angenon, G., & Rouard, M. (2022). Genome-Wide Classification and Phylogenetic Analyses of the GDSL-Type Esterase/Lipase (GELP) Family in Flowering Plants. *International Journal of Molecular Sciences*, *23*(20), Article 20. https://doi.org/10.3390/ijms232012114

Chalhoub, B. A., Denœud, F., Liu, S., Parkin, I. A. P., Tang, H., Wang, X., Chiquet, J., Belcram, H., Tong, C., Samans, B., Corréa, M., Da Silva, C. S. R., Just, J., Falentin, C., Koh, C., Le Clainche, I., Bernard, M., Bento, P., Noël, B., … Wincker, P. (2014). Early allopolyploid evolution in the post-neolithic Brassica napus oilseed genome. *Science*, *345*(6199), 950–953. https://doi.org/10.1126/science.1253435

Challinor, A. J., Watson, J., Lobell, D. B., Howden, S. M., Smith, D. R., & Chhetri, N. (2014). A meta-analysis of crop yield under climate change and adaptation. *Nature Climate Change*, *4*(4), Article 4. https://doi.org/10.1038/nclimate2153

Chaudhury, A. M., Koltunow, A., Payne, T., Luo, M., Tucker, M. R., Dennis, E. S., & Peacock, W. J. (2001). Control of Early Seed Development. *Annual Review of Cell and Developmental Biology*, *17*(1), 677–699. https://doi.org/10.1146/annurev.cellbio.17.1.677

Chen, M., Du, X., Zhu, Y., Wang, Z., Hua, S., Li, Z., Guo, W., Zhang, G., Peng, J., & Jiang, L. (2012). Seed Fatty Acid Reducer acts downstream of gibberellin signalling pathway to lower seed fatty acid storage in Arabidopsis. *Plant, Cell & Environment*, *35*(12), 2155–2169. https://doi.org/10.1111/j.1365-3040.2012.02546.x

Chen, M., MacGregor, D. R., Dave, A., Florance, H., Moore, K., Paszkiewicz, K., Smirnoff, N., Graham, I. A., & Penfield, S. (2014). Maternal temperature history activates Flowering Locus T in fruits to control progeny dormancy according to time of year. *Proceedings of the National Academy of Sciences*, *111*(52), 18787–18792. https://doi.org/10.1073/pnas.1412274111

Chen, M., & Penfield, S. (2018). Feedback regulation of COOLAIR expression controls seed dormancy and flowering time. *Science*, *360*(6392), 1014–1017. https://doi.org/10.1126/science.aar7361

Chen, X., Yoong, F.-Y., O'Neill, C. M., & Penfield, S. (2021). Temperature during seed maturation controls seed vigour through ABA breakdown in the endosperm and causes a passive effect on DOG1 mRNA levels during entry into quiescence. *New Phytologist*, *n/a*(n/a). https://doi.org/10.1111/nph.17646

Chen, Y., Liu, L., Shen, Y., Liu, S., Huang, J., Long, Q., Wu, W., Yang, C., Chen, H., Guo, X., Cheng, Z., Jiang, L., & Wan, J. (2015). Loss of Function of the Cytochrome P450 Gene CYP78B5 Causes Giant Embryos in Rice. *Plant Molecular Biology Reporter*, *33*(1), 69–83. https://doi.org/10.1007/s11105-014-0731-3

Chen, Y., Lun, A. T., McCarthy, D. J., Ritchie, M. E., Phipson, B., Hu, Y., Zhou, X., Robinson, M. D., & Smyth, G. K. (2022). *edgeR: Empirical Analysis of Digital Gene Expression Data in R* (3.40.0). Bioconductor version: Release (3.16). https://doi.org/10.18129/B9.bioc.edgeR

Cheng, Y., Dai, X., & Zhao, Y. (2004). AtCAND1, A HEAT-Repeat Protein That Participates in Auxin Signaling in Arabidopsis. *Plant Physiology*, *135*(2), 1020–1026. https://doi.org/10.1104/pp.104.044495

Cheng, Y., Zhang, Y., Liu, C., Ai, P., & Liu, J. (2018). Identification of genes regulating ovary differentiation after pollination in hazel by comparative transcriptome analysis. *BMC Plant Biology*, *18*(1), 84. https://doi.org/10.1186/s12870-018-1296-3

Cheng, Z. J., Zhao, X. Y., Shao, X. X., Wang, F., Zhou, C., Liu, Y. G., Zhang, Y., & Zhang, X. S. (2014). Abscisic Acid Regulates Early Seed Development in Arabidopsis by ABI5-Mediated Transcription of SHORT HYPOCOTYL UNDER BLUE1[C][W][OPEN]. *The Plant Cell*, *26*(3), 1053–1068. https://doi.org/10.1105/tpc.113.121566

Chepyshko, H., Lai, C.-P., Huang, L.-M., Liu, J.-H., & Shaw, J.-F. (2012). Multifunctionality and diversity of GDSL esterase/lipase gene family in rice (Oryza sativa L. japonica) genome: New insights from bioinformatics analysis. *BMC Genomics*, *13*(1), 309. https://doi.org/10.1186/1471-2164-13-309

Chiang, G. C. K., Bartsch, M., Barua, D., Nakabayashi, K., Debieu, M., Kronholm, I., Koornneef, M., Soppe, W. J. J., Donohue, K., & De MEAUX, J. (2011). DOG1 expression is predicted by the seed-maturation environment and contributes to geographical variation in germination in Arabidopsis thaliana. *Molecular Ecology*, *20*(16), 3336–3349. https://doi.org/10.1111/j.1365-294X.2011.05181.x

Chiang, G. C. K., Barua, D., Kramer, E. M., Amasino, R. M., & Donohue, K. (2009). Major flowering time gene, FLOWERING LOCUS C, regulates seed germination in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, *106*(28), 11661–11666. https://doi.org/10.1073/pnas.0901367106

Chiu, R. S., Pan, S., Zhao, R., & Gazzarrini, S. (2016). ABA-dependent inhibition of the ubiquitin proteasome system during germination at high temperature in Arabidopsis. *The Plant Journal*, *88*(5), 749–761. https://doi.org/10.1111/tpj.13293

Choi, J., Hyun, Y., Kang, M.-J., In Yun, H., Yun, J.-Y., Lister, C., Dean, C., Amasino, R. M., Noh, B., Noh, Y.-S., & Choi, Y. (2009). Resetting and regulation of Flowering Locus C

expression during Arabidopsis reproductive development. *The Plant Journal: For Cell and Molecular Biology*, *57*(5), 918–931. https://doi.org/10.1111/j.1365-313X.2008.03776.x

Choi, K., Kim, J., Hwang, H.-J., Kim, S., Park, C., Kim, S. Y., & Lee, I. (2011). The FRIGIDA Complex Activates Transcription of FLC, a Strong Flowering Repressor in Arabidopsis, by Recruiting Chromatin Modification Factors. *The Plant Cell*, *23*(1), 289–303. https://doi.org/10.1105/tpc.110.075911

Clauß, K., von Roepenack-Lahaye, E., Böttcher, C., Roth, M. R., Welti, R., Erban, A., Kopka, J., Scheel, D., Milkowski, C., & Strack, D. (2011). Overexpression of Sinapine Esterase BnSCE3 in Oilseed Rape Seeds Triggers Global Changes in Seed Metabolism. *Plant Physiology*, *155*(3), 1127–1145. https://doi.org/10.1104/pp.110.169821

Creff, A., Ali, O., Bayle, V., Ingram, G., & Landrein, B. (2021). *Endosperm turgor pressure both promotes and restricts seed growth and size* [Preprint]. Plant Biology. https://doi.org/10.1101/2021.03.22.436392

Creff, A., Brocard, L., & Ingram, G. (2015). A mechanically sensitive cell layer regulates the physical properties of the Arabidopsis seed coat. *Nature Communications*, *6*(1), Article 1. https://doi.org/10.1038/ncomms7382

Crevillén, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., Cao, X., & Dean, C. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. *Nature*, *515*(7528), 587–590. https://doi.org/10.1038/nature13722

Csorba, T., Questa, J. I., Sun, Q., & Dean, C. (2014). Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proceedings of*

the National Academy of Sciences of the United States of America, *111*(45), 16160–16165. https://doi.org/10.1073/pnas.1419030111

Daie, J., & Campbell, W. F. (1981). Response of Tomato Plants to Stressful Temperatures: INCREASE IN ABSCISIC ACID CONCENTRATIONS. *Plant Physiology*, *67*(1), 26–29. https://doi.org/10.1104/pp.67.1.26

Dawson, J., Sözen, E., Vizir, I., Waeyenberge, S. V., Wilson, Z. A., & Mulligan, B. J. (1999). Characterization and genetic mapping of a mutation (ms35) which prevents anther dehiscence in Arabidopsis thaliana by affecting secondary wall thickening in the endothecium. *The New Phytologist*, *144*(2), 213–222. https://doi.org/10.1046/j.1469-8137.1999.00507.x

De Lucia, F., Crevillen, P., Jones, A. M. E., Greb, T., & Dean, C. (2008). A PHD-Polycomb Repressive Complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proceedings of the National Academy of Sciences*, *105*(44), 16831–16836. https://doi.org/10.1073/pnas.0808687105

DEFRA. (2021). *Agriculture in the United Kingdom 2020*. https://www.gov.uk/government/statistics/agriculture-in-the-united-kingdom-2020

del Olmo, I., López, J. A., Vázquez, J., Raynaud, C., Piñeiro, M., & Jarillo, J. A. (2016). Arabidopsis DNA polymerase ϵ recruits components of Polycomb repressor complex to mediate epigenetic gene silencing. *Nucleic Acids Research*, *44*(12), 5597–5614. https://doi.org/10.1093/nar/gkw156

Deryng, D., Conway, D., Ramankutty, N., Price, J., & Warren, R. (2014). Global crop yield response to extreme heat stress under multiple climate change futures. *Environmental Research Letters*, *9*(3), 034011. https://doi.org/10.1088/1748-9326/9/3/034011

Dewar, A. M., Ferguson, A., Pell, J. K., Nicholls, C., & Watts, J. (2016). *A review of pest management in cereals and oilseed rape in the UK* (No. 86; AHDB Research Review). https://ahdb.org.uk/a-review-of-pest-management-in-cereals-and-oilseed-rape-in-the-uk

Ding, G., Zhao, Z., Liao, Y., Hu, Y., Shi, L., Long, Y., & Xu, F. (2012). Quantitative trait loci for seed yield and yield-related traits, and their responses to reduced phosphorus supply in Brassica napus. *Annals of Botany*, *109*(4), 747–759. https://doi.org/10.1093/aob/mcr323

Ding, L.-N., Guo, X.-J., Li, M., Fu, Z.-L., Yan, S.-Z., Zhu, K.-M., Wang, Z., & Tan, X.-L. (2019). Improving seed germination and oil contents by regulating the GDSL transcriptional level in Brassica napus. *Plant Cell Reports*, *38*(2), 243–253. https://doi.org/10.1007/s00299-018-2365-7

Ding, L.-N., Li, M., Wang, W.-J., Cao, J., Wang, Z., Zhu, K.-M., Yang, Y.-H., Li, Y.-L., & Tan, X.-L. (2019). Advances in plant GDSL lipases: From sequences to functional mechanisms. *Acta Physiologiae Plantarum*, *41*(9), 151. https://doi.org/10.1007/s11738-019-2944-4

Dong, H., Tan, C., Li, Y., He, Y., Wei, S., Cui, Y., Chen, Y., Wei, D., Fu, Y., He, Y., Wan, H., Liu, Z., Xiong, Q., Lu, K., Li, J., & Qian, W. (2018). Genome-Wide Association Study Reveals Both Overlapping and Independent Genetic Loci to Control Seed Weight and

Silique Length in Brassica napus. *Frontiers in Plant Science*, *9*. https://www.frontiersin.org/articles/10.3389/fpls.2018.00921

Dong, X., Yi, H., Han, C.-T., Nou, I.-S., & Hur, Y. (2016). GDSL esterase/lipase genes in Brassica rapa L.: Genome-wide identification and expression analysis. *Molecular Genetics and Genomics: MGG*, *291*(2), 531–542. https://doi.org/10.1007/s00438-015-1123-6

Donohue, K. (2009). Completing the cycle: Maternal effects as the missing link in plant life histories. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1520), 1059–1074. https://doi.org/10.1098/rstb.2008.0291

Dubas, E., Moravčíková, J., Libantová, J., Matušíková, I., Benková, E., Żur, I., & Krzewska, M. (2014). The influence of heat stress on auxin distribution in transgenic B. napus microspores and microspore-derived embryos. *Protoplasma*, *251*(5), 1077–1087. https://doi.org/10.1007/s00709-014-0616-1

Duncan, S., Holm, S., Questa, J., Irwin, J., Grant, A., & Dean, C. (2015). Seasonal shift in timing of vernalization as an adaptation to extreme winter. *ELife*, *4*, e06620. https://doi.org/10.7554/eLife.06620

Egli, D. B., TeKrony, D. M., Heitholt, J. J., & Rupe, J. (2005). Air Temperature During Seed Filling and Soybean Seed Germination and Vigor. *Crop Science*, *45*(4), 1329–1335. https://doi.org/10.2135/cropsci2004.0029

Elliott, R. C., Betzner, A. S., Huttner, E., Oakes, M. P., Tucker, W. Q., Gerentes, D., Perez, P., & Smyth, D. R. (1996). AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with

pleiotropic roles in ovule development and floral organ growth. *The Plant Cell*, *8*(2), 155–168. https://doi.org/10.1105/tpc.8.2.155

Fan, C., Cai, G., Qin, J., Li, Q., Yang, M., Wu, J., Fu, T., Liu, K., & Zhou, Y. (2010). Mapping of quantitative trait loci and development of allele-specific markers for seed weight in Brassica napus. *Theoretical and Applied Genetics*, *121*(7), 1289–1301. https://doi.org/10.1007/s00122-010-1388-4

Fang, W., Wang, Z., Cui, R., Li, J., & Li, Y. (2012). Maternal control of seed size by EOD3/CYP78A6 in Arabidopsis thaliana. *The Plant Journal*, *70*(6), 929–939. https://doi.org/10.1111/j.1365-313X.2012.04907.x

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M. F., Kater, M. M., & Colombo, L. (2003). MADS-Box Protein Complexes Control Carpel and Ovule Development in Arabidopsis. *The Plant Cell*, *15*(11), 2603–2611. https://doi.org/10.1105/tpc.015123

Fenner, M. (1991). The effects of the parent environment on seed germinability. *Seed Science Research*, *1*(2), 75–84. https://doi.org/10.1017/S0960258500000696

Ferreira, M. E., Satagopan, J., Yandell, B. S., Williams, P. H., & Osborn, T. C. (1995). Mapping loci controlling vernalization requirement and flowering time in Brassica napus. *Theoretical and Applied Genetics*, *90*(5), 727–732. https://doi.org/10.1007/BF00222140

Marcos Filho, J. (2015). Seed vigor testing: an overview of the past, present and future perspective. *Scientia agricola*, *72*, 363-374. https://doi.org/10.1590/0103-9016-2015-0007

Finch-Savage, W. E., & Bassel, G. W. (2016). Seed vigour and crop establishment: Extending performance beyond adaptation. *Journal of Experimental Botany*, *67*(3), 567–591. https://doi.org/10.1093/jxb/erv490

Finkelstein, R. R., Gampala, S. S. L., & Rock, C. D. (2002). Abscisic Acid Signaling in Seeds and Seedlings. *The Plant Cell*, *14*(Suppl), s15. https://doi.org/10.1105/tpc.010441

Finnegan, E. J., Bond, D. M., Buzas, D. M., Goodrich, J., Helliwell, C. A., Tamada, Y., Yun, J.-Y., Amasino, R. M., & Dennis, E. S. (2011). Polycomb proteins regulate the quantitative induction of VERNALIZATION INSENSITIVE 3 in response to low temperatures. *The Plant Journal*, *65*(3), 382–391. https://doi.org/10.1111/j.1365-313X.2010.04428.x

Finnegan, E. J., Peacock, W. J., & Dennis, E. S. (2000). DNA methylation, a key regulator of plant development and other processes. *Current Opinion in Genetics & Development*, *10*(2), 217–223. https://doi.org/10.1016/S0959-437X(00)00061-7

Fletcher, J. C., Brand, U., Running, M. P., Simon, R., & Meyerowitz, E. M. (1999). Signaling of Cell Fate Decisions by CLAVATA3 in Arabidopsis Shoot Meristems. *Science*, *283*(5409), 1911–1914. https://doi.org/10.1126/science.283.5409.1911

Floris, M., Mahgoub, H., Lanet, E., Robaglia, C., & Menand, B. (2009). Post-transcriptional Regulation of Gene Expression in Plants during Abiotic Stress. *International Journal of Molecular Sciences*, *10*(7), 3168–3185. https://doi.org/10.3390/ijms10073168

Folsom, J. J., Begcy, K., Hao, X., Wang, D., & Walia, H. (2014). Rice Fertilization-Independent Endosperm1 Regulates Seed Size under Heat Stress by Controlling Early

Endosperm Development1[W]. *Plant Physiology*. https://doi.org/10.1104/pp.113.232413

Futschik, M. E. (2022). *Mfuzz: Soft clustering of time series gene expression data* (2.58.0). Bioconductor version: Release (3.16). https://doi.org/10.18129/B9.bioc.Mfuzz

Futschik, M. E., & Carlisle, B. (2005). Noise-robust soft clustering of gene expression time-course data. *Journal of Bioinformatics and Computational Biology*, *03*(04), 965–988. https://doi.org/10.1142/S0219720005001375

Gajardo, H. A., Wittkop, B., Soto-Cerda, B., Higgins, E. E., Parkin, I. A. P., Snowdon, R. J., Federico, M. L., & Iniguez-Luy, F. L. (2015). Association mapping of seed quality traits in Brassica napus L. using GWAS and candidate QTL approaches. *Molecular Breeding*, *35*(6), 143. https://doi.org/10.1007/s11032-015-0340-3

Galili, T., Benjamini, Y., Simpson, G., package), G. J. (imported code from his dendroextras, marcog), M. G. (a k a:, Renaudie (https://github.com/plannapus), J., Infastructure, T. R. C. T. (Thanks for the, examples), and code in the, Hornik, K., Ligges, U., Spiess, A.-N., Horvath, S., Langfelder, P., skullkey, d3dendrogram), M. V. D. L. (https://github com/markvanderloo, author), A. de V. (ggdendro, author), Z. G. (circlize, Cath (https://github.com/CathG), Ma (https://github.com/JohnMCMa), J., … Schupp (https://github.com/pschupp), P. (2022). *dendextend: Extending "dendrogram" Functionality in R* (1.16.0). https://CRAN.R-project.org/package=dendextend

Gallego-Giraldo, C., Hu, J., Urbez, C., Gomez, M. D., Sun, T., & Perez-Amador, M. A. (2014). Role of the gibberellin receptors GID1 during fruit-set in Arabidopsis. *The Plant Journal*, *79*(6), 1020–1032. https://doi.org/10.1111/tpj.12603

Garcia, D., Gerald, J. N. F., & Berger, F. (2005). Maternal Control of Integument Cell Elongation and Zygotic Control of Endosperm Growth Are Coordinated to Determine Seed Size in Arabidopsis. *The Plant Cell*, *17*(1), 52–60. https://doi.org/10.1105/tpc.104.027136

Garcia, D., Saingery, V., Chambrier, P., Mayer, U., Jürgens, G., & Berger, F. (2003). Arabidopsis haiku Mutants Reveal New Controls of Seed Size by Endosperm. *Plant Physiology*, *131*(4), 1661–1670. https://doi.org/10.1104/pp.102.018762

Ge, S. X., Jung, D., & Yao, R. (2020). ShinyGO: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics*, *36*(8), 2628–2629. https://doi.org/10.1093/bioinformatics/btz931

Gendall, A. R., Levy, Y. Y., Wilson, A., & Dean, C. (2001). The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis. *Cell*, *107*(4), 525–535. https://doi.org/10.1016/S0092-8674(01)00573-6

Graeber, K., Nakabayashi, K., Miatton, E., Leubner-Metzger, G., & Soppe, W. J. J. (2012). Molecular mechanisms of seed dormancy. *Plant, Cell & Environment*, *35*(10), 1769–1786. https://doi.org/10.1111/j.1365-3040.2012.02542.x

Gray, D., Wurr, D. C. E., Ward, J. A., & Fellows, J. R. (1988). Influence of post-flowering temperature on seed development, and subsequent performance of crisp lettuce. *Annals of Applied Biology*, *113*(2), 391–402. https://doi.org/10.1111/j.1744-7348.1988.tb03315.x

Gray, W. M., Östin, A., Sandberg, G., Romano, C. P., & Estelle, M. (1998). High temperature promotes auxin-mediated hypocotyl elongation in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(12), 7197–7202.

Griffiths, J., Murase, K., Rieu, I., Zentella, R., Zhang, Z.-L., Powers, S. J., Gong, F., Phillips, A. L., Hedden, P., Sun, T., & Thomas, S. G. (2006). Genetic Characterization and Functional Analysis of the GID1 Gibberellin Receptors in Arabidopsis. *The Plant Cell*, *18*(12), 3399–3414. https://doi.org/10.1105/tpc.106.047415

Grossniklaus, U., & Paro, R. (2014). Transcriptional silencing by polycomb-group proteins. *Cold Spring Harbor Perspectives in Biology*, *6*(11), a019331. https://doi.org/10.1101/cshperspect.a019331

Gu, W., Wang, X., Zhai, C., Xie, X., & Zhou, T. (2012). Selection on Synonymous Sites for Increased Accessibility around miRNA Binding Sites in Plants. *Molecular Biology and Evolution*, *29*(10), 3037–3044. https://doi.org/10.1093/molbev/mss109

Gu, W., Zhou, T., Ma, J., Sun, X., & Lu, Z. (2003). Folding type specific secondary structure propensities of synonymous codons. *IEEE Transactions on NanoBioscience*, *2*(3), 150–157. https://doi.org/10.1109/TNB.2003.817024

Habekotte´, B. (1997). A model of the phenological development of winter oilseed rape (Brassica napus L.). *Field Crops Research*, *54*(2), 127–136. https://doi.org/10.1016/S0378-4290(97)00043-9

Hackenberg, D., Asare-Bediako, E., Baker, A., Walley, P., Jenner, C., Greer, S., Bramham, L., Batley, J., Edwards, D., Delourme, R., Barker, G., Teakle, G., & Walsh, J. (2020). Identification and QTL mapping of resistance to Turnip yellows virus (TuYV) in oilseed

rape, Brassica napus. *Theoretical and Applied Genetics*, *133*(2), 383–393. https://doi.org/10.1007/s00122-019-03469-z

Harper, A. L., Trick, M., Higgins, J., Fraser, F., Clissold, L., Wells, R., Hattori, C., Werner, P., & Bancroft, I. (2012). Associative transcriptomics of traits in the polyploid crop species Brassica napus. *Nature Biotechnology*, *30*(8), 798–802. https://doi.org/10.1038/nbt.2302

Hashimoto, K., Miyashima, S., Sato-Nara, K., Yamada, T., & Nakajima, K. (2018). Functionally Diversified Members of the MIR165/6 Gene Family Regulate Ovule Morphogenesis in Arabidopsis thaliana. *Plant and Cell Physiology*, *59*(5), 1017–1026. https://doi.org/10.1093/pcp/pcy042

Haughn, G., & Chaudhury, A. (2005). Genetic analysis of seed coat development in Arabidopsis. *Trends in Plant Science*, *10*(10), 472–477. https://doi.org/10.1016/j.tplants.2005.08.005

He, D., Wang, E., Wang, J., & Lilley, J. M. (2017). Genotype×environment×management interactions of canola across China: A simulation study. *Agricultural and Forest Meteorology*, *247*, 424–433. https://doi.org/10.1016/j.agrformet.2017.08.027

He, H., de Souza Vidigal, D., Snoek, L. B., Schnabel, S., Nijveen, H., Hilhorst, H., & Bentsink, L. (2014). Interaction between parental environment and genotype affects plant and seed performance in Arabidopsis. *Journal of Experimental Botany*, *65*(22), 6603–6615. https://doi.org/10.1093/jxb/eru378

He, H., Willems, L. A. J., Batushansky, A., Fait, A., Hanson, J., Nijveen, H., Hilhorst, H. W. M., & Bentsink, L. (2016). Effects of Parental Temperature and Nitrate on Seed

Performance are Reflected by Partly Overlapping Genetic and Metabolic Pathways. *Plant and Cell Physiology*, *57*(3), 473–487. https://doi.org/10.1093/pcp/pcv207

He, Y., Doyle, M. R., & Amasino, R. M. (2004). PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. *Genes & Development*, *18*(22), 2774–2784. https://doi.org/10.1101/gad.1244504

He, Y., Wu, D., Wei, D., Fu, Y., Cui, Y., Dong, H., Tan, C., & Qian, W. (2017). GWAS, QTL mapping and gene expression analyses in Brassica napus reveal genetic control of branching morphogenesis. *Scientific Reports*, *7*(1), Article 1. https://doi.org/10.1038/s41598-017-15976-4

He, Z., Cheng, F., Li, Y., Wang, X., Parkin, I. A. P., Chalhoub, B., Liu, S., & Bancroft, I. (2015). Construction of Brassica A and C genome-based ordered pan-transcriptomes for use in rapeseed genomic research. *Data in Brief*, *4*, 357–362. https://doi.org/10.1016/j.dib.2015.06.016

Hepworth, J., Antoniou-Kourounioti, R. L., Berggren, K., Selga, C., Tudor, E. H., Yates, B., Cox, D., Collier Harris, B. R., Irwin, J. A., Howard, M., Säll, T., Holm, S., & Dean, C. (2020). Natural variation in autumn expression is the major adaptive determinant distinguishing Arabidopsis FLC haplotypes. *ELife*, *9*, e57671. https://doi.org/10.7554/eLife.57671

Hepworth, J., Antoniou-Kourounioti, R. L., Bloomer, R. H., Selga, C., Berggren, K., Cox, D., Collier Harris, B. R., Irwin, J. A., Holm, S., Säll, T., Howard, M., & Dean, C. (2018).

Absence of warmth permits epigenetic memory of winter in Arabidopsis. *Nature Communications*, *9*(1), Article 1. https://doi.org/10.1038/s41467-018-03065-7

Hobbie, L., McGovern, M., Hurwitz, L. R., Pierro, A., Liu, N. Y., Bandyopadhyay, A., & Estelle, M. (2000). The axr6 mutants of Arabidopsis thaliana define a gene involved in auxin response and early development. *Development*, *127*(1), 23–32.

Hodgson, A. S. (1978). Rapeseed adaptation in Northern New South Wales. II.* Predicting plant development of Brassica campestris L. and Brassica napus L. and its implications for planting time, designed to avoid water deficit and frost. *Australian Journal of Agricultural Research*, *29*(4), 711–726. https://doi.org/10.1071/ar9780711

Holdsworth, M. J., Bentsink, L., & Soppe, W. J. J. (2008). Molecular networks regulating Arabidopsis seed maturation, after-ripening, dormancy and germination. *New Phytologist*, *179*(1), 33–54. https://doi.org/10.1111/j.1469-8137.2008.02437.x

Hong, J. K., Kim, S.-Y., Kim, J. S., Kim, J. A., Park, B.-S., & Lee, Y.-H. (2011). Promoters of three Brassica rapa FLOWERING LOCUS C differentially regulate gene expression during growth and development in Arabidopsis. *Genes & Genomics*, *33*(1), 75–82. https://doi.org/10.1007/s13258-010-0117-3

Horiguchi, G., Nakayama, H., Ishikawa, N., Kubo, M., Demura, T., Fukuda, H., & Tsukaya, H. (2011). ANGUSTIFOLIA3 Plays Roles in Adaxial/Abaxial Patterning and Growth in Leaf Morphogenesis. *Plant and Cell Physiology*, *52*(1), 112–124. https://doi.org/10.1093/pcp/pcq178

Hou, J., Long, Y., Raman, H., Zou, X., Wang, J., Dai, S., Xiao, Q., Li, C., Fan, L., Liu, B., & Meng, J. (2012). A Tourist-like MITE insertion in the upstream region of the

BnFLC.A10 gene is associated with vernalization requirement in rapeseed (Brassica napus L.). *BMC Plant Biology*, *12*(1), 238. https://doi.org/10.1186/1471-2229-12-238

Huang, R., Liu, Z., Xing, M., Yang, Y., Wu, X., Liu, H., & Liang, W. (2019). Heat stress suppresses Brassica napus seed oil accumulation by inhibition of photosynthesis and BnWRI1 pathway. *Plant & Cell Physiology*. https://doi.org/10.1093/pcp/pcz052

Imaizumi, T., Auge, G., & Donohue, K. (2017). Photoperiod throughout the maternal life cycle, not photoperiod during seed imbibition, influences germination in Arabidopsis thaliana. *American Journal of Botany*, *104*(4), 516–526. https://doi.org/10.3732/ajb.1600389

Ingouff, M., Rademacher, S., Holec, S., Šoljić, L., Xin, N., Readshaw, A., Foo, S. H., Lahouze, B., Sprunck, S., & Berger, F. (2010). Zygotic Resetting of the HISTONE 3 Variant Repertoire Participates in Epigenetic Reprogramming in Arabidopsis. *Current Biology*, *20*(23), 2137–2143. https://doi.org/10.1016/j.cub.2010.11.012

Ingram, G. C. (2010). Family life at close quarters: Communication and constraint in angiosperm seed development. *Protoplasma*, *247*(3), 195–214. https://doi.org/10.1007/s00709-010-0184-y

Jaime, R., Alcántara, J. M., Manzaneda, A. J., & Rey, P. J. (2018). Climate change decreases suitable areas for rapeseed cultivation in Europe but provides new opportunities for white mustard as an alternative oilseed for biofuel production. *PLOS ONE*, *13*(11), e0207124. https://doi.org/10.1371/journal.pone.0207124

Jofuku, K. D., den Boer, B. G., Van Montagu, M., & Okamuro, J. K. (1994). Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *The Plant Cell*, *6*(9), 1211–1225.

Jofuku, K. D., Omidyar, P. K., Gee, Z., & Okamuro, J. K. (2005). Control of seed mass and seed yield by the floral homeotic gene APETALA2. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(8), 3117–3122. https://doi.org/10.1073/pnas.0409893102

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., & Dean, C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science (New York, N.Y.)*, *290*(5490), 344–347. https://doi.org/10.1126/science.290.5490.344

Johnson, C. S., Kolevski, B., & Smyth, D. R. (2002). TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *The Plant Cell*, *14*(6), 1359–1375.

Jones, D. M., Wells, R., Pullen, N., Trick, M., Irwin, J. A., & Morris, R. J. (2018). Spatio-temporal expression dynamics differ between homologues of flowering time genes in the allopolyploid Brassica napus. *The Plant Journal*, *96*(1), 103–118. https://doi.org/10.1111/tpj.14020

Kanno, Y., Jikumaru, Y., Hanada, A., Nambara, E., Abrams, S. R., Kamiya, Y., & Seo, M. (2010). Comprehensive Hormone Profiling in Developing Arabidopsis Seeds: Examination of the Site of ABA Biosynthesis, ABA Transport and Hormone Interactions. *Plant and Cell Physiology*, *51*(12), 1988–2001. https://doi.org/10.1093/pcp/pcq158

Kaufmann, K., Muiño, J. M., Jauregui, R., Airoldi, C. A., Smaczniak, C., Krajewski, P., & Angenent, G. C. (2009). Target Genes of the MADS Transcription Factor SEPALLATA3: Integration of Developmental and Hormonal Pathways in the Arabidopsis Flower. *PLOS Biology*, *7*(4), e1000090. https://doi.org/10.1371/journal.pbio.1000090

Kegode, G. O., & Pearce, R. B. (1998). Influence of environment during maternal plant growth on dormancy of shattercane (Sorghum bicolor) and giant foxtail (Setaria faberi) seed. *Weed Science*, *46*(3), 322–329. https://doi.org/10.1017/S0043174500089499

Kelley, D. R., Skinner, D. J., & Gasser, C. S. (2009). Roles of polarity determinants in ovule development. *The Plant Journal*, *57*(6), 1054–1064. https://doi.org/10.1111/j.1365-313X.2008.03752.x

Kendall, S. L., Hellwege, A., Marriot, P., Whalley, C., Graham, I. A., & Penfield, S. (2011). Induction of Dormancy in Arabidopsis Summer Annuals Requires Parallel Regulation of DOG1 and Hormone Metabolism by Low Temperature and CBF Transcription Factors. *The Plant Cell*, *23*(7), 2568–2580. https://doi.org/10.1105/tpc.111.087643

Khan, M. N., Zhang, J., Luo, T., Liu, J., Ni, F., Rizwan, M., Fahad, S., & Hu, L. (2019). Morpho-physiological and biochemical responses of tolerant and sensitive rapeseed cultivars to drought stress during early seedling growth stage. *Acta Physiologiae Plantarum*. https://doi.org/10.1007/s11738-019-2812-2

Khan, S. U., Yangmiao, J., Liu, S., Zhang, K., Khan, M. H. U., Zhai, Y., Olalekan, A., Fan, C., & Zhou, Y. (2019). Genome-wide association studies in the genetic dissection of ovule number, seed number, and seed weight in Brassica napus L. *Industrial Crops and Products*, *142*, 111877. https://doi.org/10.1016/j.indcrop.2019.111877

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), Article 4. https://doi.org/10.1038/nmeth.3317

Kim, J. H., & Kende, H. (2004). A transcriptional coactivator, AtGIF1, is involved in regulating leaf growth and morphology in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(36), 13374–13379. https://doi.org/10.1073/pnas.0405450101

Kim, Y.-C., Nakajima, M., Nakayama, A., & Yamaguchi, I. (2005). Contribution of gibberellins to the formation of Arabidopsis seed coat through starch degradation. *Plant & Cell Physiology*. https://doi.org/10.1093/PCP/PCI141

Kipp-Sinanis, E. (2008). Heat stress effects on growth and development in three ecotypes of varying latitude of Arabidopsis. *Applied Ecology and Environmental Research*, *6*. https://doi.org/10.15666/aeer/0604_001014

Klucher, K. M., Chow, H., Reiser, L., & Fischer, R. L. (1996). The AINTEGUMENTA gene of Arabidopsis required for ovule and female gametophyte development is related to the floral homeotic gene APETALA2. *The Plant Cell*, *8*(2), 137–153. https://doi.org/10.1105/tpc.8.2.137

Kochanek, J., Buckley, Y. M., Probert, R. J., Adkins, S. W., & Steadman, K. J. (2010). Pre-zygotic parental environment modulates seed longevity. *Austral Ecology*, *35*(7), 837–848. https://doi.org/10.1111/j.1442-9993.2010.02118.x

Koornneef, M., Alonso-Blanco, C., Peeters, A. J. M., & Soppe, W. (1998). Genetic Control of Flowering Time in Arabidopsis. *Annual Review of Plant Physiology and Plant Molecular Biology*, *49*(1), 345–370. https://doi.org/10.1146/annurev.arplant.49.1.345

Koornneef, M., Reuling, G., & Karssen, C. M. (1984). The isolation and characterization of abscisic acid-insensitive mutants of Arabidopsis thaliana. *Physiologia Plantarum*, *61*(3), 377–383. https://doi.org/10.1111/j.1399-3054.1984.tb06343.x

Körber, N., Bus, A., Li, J., Parkin, I. A. P., Wittkop, B., Snowdon, R. J., & Stich, B. (2016). Agronomic and Seed Quality Traits Dissected by Genome-Wide Association Mapping in Brassica napus. *Frontiers in Plant Science*, *7*. https://www.frontiersin.org/articles/10.3389/fpls.2016.00386

Korte, A., & Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods*, *9*(1), 29. https://doi.org/10.1186/1746-4811-9-29

Kourani, M., Mohareb, F., Rezwan, F. I., Anastasiadi, M., & Hammond, J. P. (2022). Genetic and physiological responses to heat stress in Brassica napus. Frontiers in Plant Science, 13. https://doi.org/10.3389/fpls.2022.832147

Kroj, T., Savino, G., Valon, C., Giraudat, J., & Parcy, F. (2003). Regulation of storage protein gene expression in Arabidopsis. *Development (Cambridge, England)*, *130*(24), 6065–6073. https://doi.org/10.1242/dev.00814

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lamichaney, A., Parihar, A. K., Hazra, K. K., Dixit, G. P., Katiyar, P. K., Singh, D., Singh, A. K., Kumar, N., & Singh, N. P. (2021). Untangling the Influence of Heat Stress on Crop Phenology, Seed Set, Seed Weight, and Germination in Field Pea (Pisum sativum L.). *Frontiers in Plant Science*, *12*. https://www.frontiersin.org/articles/10.3389/fpls.2021.635868

Lancashire, P. D., Bleiholder, H., Boom, T. V. D., Langelüddeke, P., Stauss, R., Weber, E., & Witzenberger, A. (1991). A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology*, *119*(3), 561–601. https://doi.org/10.1111/j.1744-7348.1991.tb04895.x

Larkindale, J., & Knight, M. R. (2002). Protection against heat stress-induced oxidative damage in Arabidopsis involves calcium, abscisic acid, ethylene, and salicylic acid. *Plant Physiology*, *128*(2), 682–695. https://doi.org/10.1104/pp.010320

Latunde-Dada, A. O., Teakle, G. R., Evans, N., Hornby, P. E., Pink, D. a. C., Fitt, B. D. L., King, G. J., & Crute, I. R. (2004). *Biodiversity and genetic resources for Brassica napus and its major fungal pathogens within the UK Defra Oilseed RapE Genetic Improvement Network (OREGIN)*. http://uhra.herts.ac.uk/handle/2299/15837

Le, B. H., Cheng, C., Bui, A. Q., Wagmaister, J. A., Henry, K. F., Pelletier, J., Kwong, L., Belmonte, M., Kirkbride, R., Horvath, S., Drews, G. N., Fischer, R. L., Okamuro, J. K., Harada, J. J., & Goldberg, R. B. (2010). Global analysis of gene activity during Arabidopsis seed development and identification of seed-specific transcription factors. *Proceedings of the National Academy of Sciences*, *107*(18), 8063–8070. https://doi.org/10.1073/pnas.1003530107

Lee, B. H., Wynn, A. N., Franks, R. G., Hwang, Y., Lim, J., & Kim, J. H. (2014). The Arabidopsis thaliana GRF-INTERACTING FACTOR gene family plays an essential role in control of male and female reproductive development. *Developmental Biology*, *386*(1), 12–24. https://doi.org/10.1016/j.ydbio.2013.12.009

Lee, I., Wolfe, D. S., Nilsson, O., & Weigel, D. (1997). A LEAFY co-regulator encoded by UNUSUAL FLORAL ORGANS. *Current Biology*, *7*(2), 95–104. https://doi.org/10.1016/S0960-9822(06)00053-4

Leijten, W., Koes, R., Roobeek, I., & Frugis, G. (2018). Translating Flowering Time from Arabidopsis thaliana to Brassicaceae and Asteraceae Crop Species. *Plants*, *7*(4), 111. https://doi.org/10.3390/plants7040111

Lenth, R. V., Buerkner, P., Giné-Vázquez, I., Herve, M., Jung, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2022). *Package "emmeans."* https://cran.r-project.org/web/packages/emmeans/index.html

Li, F., Chen, B., Xu, K., Wu, J., Song, W., Bancroft, I., Harper, A. L., Trick, M., Liu, S., Gao, G., Wang, N., Yan, G., Qiao, J., Li, J., Li, H., Xiao, X., Zhang, T., & Wu, X. (2014). Genome-Wide Association Study Dissects the Genetic Architecture of Seed Weight and Seed Quality in Rapeseed (Brassica napus L.). *DNA Research*, *21*(4), 355–367. https://doi.org/10.1093/dnares/dsu002

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, N., & Li, Y. (2015). Maternal control of seed size in plants. *Journal of Experimental Botany*, *66*(4), 1087–1097. https://doi.org/10.1093/jxb/eru549

Li, N., & Li, Y. (2016). Signaling pathways of seed size control in plants. *Current Opinion in Plant Biology*, *33*, 23–32. https://doi.org/10.1016/j.pbi.2016.05.008

Li, N., Peng, W., Shi, J., Wang, X., Liu, G., & Wang, H. (2015). The Natural Variation of Seed Weight Is Mainly Controlled by Maternal Genotype in Rapeseed (Brassica napus L.). *PLOS ONE*, *10*(4), e0125360. https://doi.org/10.1371/journal.pone.0125360

Li, Y., Zheng, L., Corke, F., Smith, C., & Bevan, M. W. (2008). Control of final seed and organ size by the DA1 gene family in Arabidopsis thaliana. *Genes & Development*, *22*(10), 1331–1336. https://doi.org/10.1101/gad.463608

Li, Z., Wu, B., J., Lu, & G. Y., Chen, Y., Zou, C. S., & Zhang, X. K. (2012). Differences in physiological responses of Brassica napus genotypes under water stress during seedling stage. *Chin. J. Oil Crops Sci*, *34*, 033–039.

Lim, S., Park, J., Lee, N., Jeong, J., Toh, S., Watanabe, A., Kim, J., Kang, H., Kim, D. H., Kawakami, N., & Choi, G. (2013). ABA-INSENSITIVE3, ABA-INSENSITIVE5, and DELLAs Interact to Activate the Expression of SOMNUS and Other High-Temperature-Inducible Genes in Imbibed Seeds in Arabidopsis. *The Plant Cell*, *25*(12), 4863–4878. https://doi.org/10.1105/tpc.113.118604

Lin, S.-I., Wang, J.-G., Poon, S.-Y., Su, C.-L., Wang, S.-S., & Chiou, T.-J. (2005). Differential regulation of FLOWERING LOCUS C expression by vernalization in cabbage and Arabidopsis. *Plant Physiology*, *137*(3), 1037–1048. https://doi.org/10.1104/pp.104.058974

Ling, H., Zhao, J., Zuo, K., Qiu, C., Yao, H., Qin, J., Sun, X., & Tang, K. (2006). Isolation and expression analysis of a GDSL-like lipase gene from Brassica napus L. *Journal of Biochemistry and Molecular Biology*, *39*(3), 297–303. https://doi.org/10.5483/bmbrep.2006.39.3.297

Liu, C., Liu, Z., Li, C., Zhang, Y., & Feng, H. (2016). Comparative transcriptome analysis of fertile and sterile buds from a genetically male sterile line of Chinese cabbage. *In Vitro Cellular & Developmental Biology - Plant*, *52*(2), 130–139. https://doi.org/10.1007/s11627-016-9754-9

Liu, T., Li, Y., Zhang, C., Qian, Y., Wang, Z., & Hou, X. (2012). Overexpression of FLOWERING LOCUSC, Isolated from Non-Heading Chinese Cabbage (Brassica campestris ssp. Chinensis Makino), Influences Fertility in Arabidopsis. *Plant Molecular Biology Reporter*, *30*(6), 1444–1449. https://doi.org/10.1007/s11105-012-0469-8

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate Trends and Global Crop Production Since 1980. *Science*, *333*(6042), 616–620. https://doi.org/10.1126/science.1204531

Lotan, T., Ohto, M., Yee, K. M., West, M. A. L., Lo, R., Kwong, R. W., Yamagishi, K., Fischer, R. L., Goldberg, R. B., & Harada, J. J. (1998). Arabidopsis LEAFY COTYLEDON1 Is Sufficient to Induce Embryo Development in Vegetative Cells. *Cell*, *93*(7), 1195–1205. https://doi.org/10.1016/S0092-8674(00)81463-4

Love, M., Ahlmann-Eltze, C., Forbes, K., Anders, S., Huber, W., FP7, R. E., NHGRI, N., & CZI. (2022). *DESeq2: Differential gene expression analysis based on the negative*

binomial distribution (1.38.1). Bioconductor version: Release (3.16). https://doi.org/10.18129/B9.bioc.DESeq2

Luo, M., Dennis, E. S., Berger, F., Peacock, W. J., & Chaudhury, A. (2005). MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(48), 17531–17536. https://doi.org/10.1073/pnas.0508418102

Luo, X., Ou, Y., Li, R., & He, Y. (2020). Maternal transmission of the epigenetic 'memory of winter cold' in Arabidopsis. *Nature Plants*, *6*(10), Article 10. https://doi.org/10.1038/s41477-020-00774-0

Lur, H. S., & Setter, T. L. (1993). Role of Auxin in Maize Endosperm Development (Timing of Nuclear DNA Endoreduplication, Zein Expression, and Cytokinin). *Plant Physiology*, *103*(1), 273–280. https://doi.org/10.1104/pp.103.1.273

MacGregor, D. R., Kendall, S. L., Florance, H., Fedi, F., Moore, K., Paszkiewicz, K., Smirnoff, N., & Penfield, S. (2015). Seed production temperature regulation of primary dormancy occurs through control of seed coat phenylpropanoid metabolism. *New Phytologist*, *205*(2), 642–652. https://doi.org/10.1111/nph.13090

Mácová, K., Prabhullachandran, U., Spyroglou, I., Štefková, M., Pěnčík, A., Endlová, L., Novák, O., & Robert, H. S. (2021). *Effects of long-term high-temperature stress on reproductive growth and seed development in development in Brassica napus* (p. 2021.03.11.434971). bioRxiv. https://doi.org/10.1101/2021.03.11.434971

Magno Massuia de Almeida, L., Avice, J.-C., Morvan-Bertrand, A., Wagner, M.-H., González-Centeno, M. R., Teissedre, P.-L., Bessoule, J.-J., Le Guédard, M., Kim, T. H., Mollier, A., & Brunel-Muguet, S. (2021). High temperature patterns at the onset of seed maturation determine seed yield and quality in oilseed rape (Brassica napus L.) in relation to sulphur nutrition. *Environmental and Experimental Botany*, *185*, 104400. https://doi.org/10.1016/j.envexpbot.2021.104400

Mahrez, W., Shin, J., Muñoz-Viana, R., Figueiredo, D. D., Trejo-Arellano, M. S., Exner, V., Siretskiy, A., Gruissem, W., Köhler, C., & Hennig, L. (2016). BRR2a Affects Flowering Time via FLC Splicing. *PLoS Genetics*, *12*(4), e1005924. https://doi.org/10.1371/journal.pgen.1005924

Marquardt, S., Raitskin, O., Wu, Z., Liu, F., Sun, Q., & Dean, C. (2014). Functional consequences of splicing of the antisense transcript COOLAIR on FLC transcription. *Molecular Cell*, *54*(1), 156–165. https://doi.org/10.1016/j.molcel.2014.03.026

Mateos, J. L., Madrigal, P., Tsuda, K., Rawat, V., Richter, R., Romera-Branchat, M., Fornara, F., Schneeberger, K., Krajewski, P., & Coupland, G. (2015). Combinatorial activities of SHORT VEGETATIVE PHASE and FLOWERING LOCUS C define distinct modes of flowering regulation in Arabidopsis. *Genome Biology*, *16*(1), 31. https://doi.org/10.1186/s13059-015-0597-1

McAdam, E. L., Meitzel, T., Quittenden, L. J., Davidson, S. E., Dalmais, M., Bendahmane, A. I., Thompson, R., Smith, J. J., Nichols, D. S., Urquhart, S., Gélinas-Marion, A., Aubert, G., & Ross, J. J. (2017). Evidence that auxin is required for normal seed size and starch synthesis in pea. *New Phytologist*, *216*(1), 193–204. https://doi.org/10.1111/nph.14690

Meinke, D. W. (2020). Genome-wide identification of EMBRYO-DEFECTIVE (EMB) genes required for growth and development in Arabidopsis. *New Phytologist*, *226*(2), 306–325. https://doi.org/10.1111/nph.16071

Mendiburu, F. de. (2021). *agricolae: Statistical Procedures for Agricultural Research* (1.3-5). https://CRAN.R-project.org/package=agricolae

Meng, L.-S., Wang, Y.-B., Loake, G. J., & Jiang, J.-H. (2016). Seed Embryo Development Is Regulated via an AN3-MINI3 Gene Cascade. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01645

Meng, L.-S., Xu, M.-K., Li, D., Zhou, M.-M., & Jiang, J.-H. (2017). Soluble Sugar Accumulation Can Influence Seed Size via AN3-YDA Gene Cascade. *Journal of Agricultural and Food Chemistry*, *65*(20), 4121–4132. https://doi.org/10.1021/acs.jafc.7b00228

Meng, L.-S., & Yao, S.-Q. (2015). Transcription co-activator Arabidopsis ANGUSTIFOLIA3 (AN3) regulates water-use efficiency and drought tolerance by modulating stomatal density and improving root architecture by the transrepression of YODA (YDA). *Plant Biotechnology Journal*. https://doi.org/10.1111/pbi.12324

Michaels, S. D., & Amasino, R. M. (1999). FLOWERING LOCUS C Encodes a Novel MADS Domain Protein That Acts as a Repressor of Flowering. *The Plant Cell*, *11*(5), 949–956. https://doi.org/10.1105/tpc.11.5.949

Michaels, S. D., Ditta, G., Gustafson-Brown, C., Pelaz, S., Yanofsky, M., & Amasino, R. M. (2003). AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated

by vernalization. *The Plant Journal: For Cell and Molecular Biology*, *33*(5), 867–874. https://doi.org/10.1046/j.1365-313x.2003.01671.x

Miller, C., Wells, R., McKenzie, N., Trick, M., Ball, J., Fatihi, A., Dubreucq, B., Chardot, T., Lepiniec, L., & Bevan, M. W. (2019). Variation in Expression of the HECT E3 Ligase UPL3 Modulates LEC2 Levels, Seed Size, and Crop Yields in Brassica napus. *The Plant Cell*, *31*(10), 2370–2385. https://doi.org/10.1105/tpc.18.00577

Mittler, R. (2006). Abiotic stress, the field environment and stress combination. *Trends in Plant Science*, *11*(1), 15–19. https://doi.org/10.1016/j.tplants.2005.11.002

Morice, C. P., Kennedy, J. J., Rayner, N. A., Winn, J. P., Hogan, E., Killick, R. E., Dunn, R. J. H., Osborn, T. J., Jones, P. D., & Simpson, I. R. (2021). An Updated Assessment of Near-Surface Temperature Change From 1850: The HadCRUT5 Data Set. *Journal of Geophysical Research: Atmospheres*, *126*(3), e2019JD032361. https://doi.org/10.1029/2019JD032361

Moriondo, M., Giannakopoulos, C., & Bindi, M. (2011). Climate change impact assessment: The role of climate extremes in crop yield simulation. *Climatic Change*, *104*(3), 679–701. https://doi.org/10.1007/s10584-010-9871-0

Morrison, M. J., & Stewart, D. W. (2002). Heat Stress during Flowering in Summer Brassica. *Crop Science*, *42*(3), 797–803. https://doi.org/10.2135/cropsci2002.7970

Nagaharu, U., & Nagaharu, N. (1935). Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. *Jpn J Bot*, *7*(7), 389–452.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. https://doi.org/10.1111/j.2041-210x.2012.00261.x

Nambara, E., Hayama, R., Tsuchiya, Y., Nishimura, M., Kawaide, H., Kamiya, Y., & Naito, S. (2000). The Role of ABI3 and FUS3 Loci in Arabidopsis thaliana on Phase Transition from Late Embryo Development to Germination. *Developmental Biology*, *220*(2), 412–423. https://doi.org/10.1006/dbio.2000.9632

Nelson, M. N., Nesi, N., Barrero, J. M., Fletcher, A. L., Greaves, I. K., Hughes, T., Laperche, A., Snowdon, R. J., Rebetzke, G. J., & Kirkegaard, J. A. (2022). Strategies to improve field establishment of canola: A review. *Advances in Agronomy*.

Ni, D. A., Wang, L. J., Ding, C. H., & Xu, Z. H. (2001). Auxin distribution and transport during embryogenesis and seed germination of Arabidopsis. *Cell Research*, *11*(4), Article 4. https://doi.org/10.1038/sj.cr.7290096

Nicholls, C. (2016). *A review of AHDB impact assessments following the neonicotinoid seed treatment restrictions in winter oilseed rape* (No. 84; AHDB Research Review, p. 31).

Niu, D., & He, Y. (2019). LEAFY COTYLEDONs: Old genes with new roles beyond seed development. *F1000Research*, *8*, F1000 Faculty Rev-2144. https://doi.org/10.12688/f1000research.21180.1

Ohto, M., Floyd, S. K., Fischer, R. L., Goldberg, R. B., & Harada, J. J. (2009). Effects of APETALA2 on embryo, endosperm, and seed coat development determine seed size in Arabidopsis. *Sexual Plant Reproduction*, *22*(4), 277–289. https://doi.org/10.1007/s00497-009-0116-1

Olesen, J. E., Trnka, M., Kersebaum, K. C., Skjelvåg, A. O., Seguin, B., Peltonen-Sainio, P., Rossi, F., Kozyra, J., & Micale, F. (2011). Impacts and adaptation of European crop production systems to climate change. *European Journal of Agronomy*, *34*(2), 96–112. https://doi.org/10.1016/j.eja.2010.11.003

Olsen, O.-A. (2004). Nuclear endosperm development in cereals and Arabidopsis thaliana. *The Plant Cell*, *16*(Suppl), S214–S227. https://doi.org/10.1105/tpc.017111

O'Neill, C. M., Lu, X., Calderwood, A., Tudor, E. H., Robinson, P., Wells, R., Morris, R., & Penfield, S. (2019). Vernalization and Floral Transition in Autumn Drive Winter Annual Life History in Oilseed Rape. *Current Biology*, *29*(24), 4300-4306.e2. https://doi.org/10.1016/j.cub.2019.10.051

OONO, Y., OOURA, C., & UCHIMIYA, H. (2002). Expression Pattern of Aux/IAA Genes in the iaa3/shy2-1D Mutant of Arabidopsis thaliana (L.). *Annals of Botany*, *89*(1), 77–82. https://doi.org/10.1093/aob/mcf007

Osborn, T. C., Chris Pires, J., Birchler, J. A., Auger, D. L., Jeffery Chen, Z., Lee, H.-S., Comai, L., Madlung, A., Doerge, R. W., Colot, V., & Martienssen, R. A. (2003). Understanding mechanisms of novel gene expression in polyploids. *Trends in Genetics*, *19*(3), 141–147. https://doi.org/10.1016/S0168-9525(03)00015-5

Østergaard, L., & King, G. J. (2008). Standardized gene nomenclature for the Brassica genus. *Plant Methods*, *4*, 10. https://doi.org/10.1186/1746-4811-4-10

Ottoline Leyser, H. M., & Furner, I. J. (1992). Characterisation of three shoot apical meristem mutants of Arabidopsis thaliana. *Development*, *116*(2), 397–403. https://doi.org/10.1242/dev.116.2.397

Pal, L., Sandhu, S. K., Bhatia, D., & Sethi, S. (2021). Genome-wide association study for candidate genes controlling seed yield and its components in rapeseed (Brassica napus subsp. Napus). *Physiology and Molecular Biology of Plants*, *27*(9), 1933–1951. https://doi.org/10.1007/s12298-021-01060-9

Penfield, S. (2008). Temperature perception and signal transduction in plants. *New Phytologist*, *179*(3), 615–628. https://doi.org/10.1111/j.1469-8137.2008.02478.x

Penfield, S. (2017). Seed dormancy and germination. *Current Biology*, *27*(17), R874–R878. https://doi.org/10.1016/j.cub.2017.05.050

Penfield, S., & MacGregor, D. R. (2017). Effects of environmental variation during seed production on seed dormancy and germination. *Journal of Experimental Botany*, *68*(4), 819–825. https://doi.org/10.1093/jxb/erw436

Penfield, S., Warner, S., & Wilkinson, L. (2021). Molecular responses to chilling in a warming climate and their impacts on plant reproductive development and yield. *Journal of Experimental Botany*, *72*(21), 7374–7383. https://doi.org/10.1093/jxb/erab375

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X., Centeno, G. S., Khush, G. S., & Cassman, K. G. (2004). Rice yields decline with higher night temperature from global warming. *Proceedings of the National Academy of Sciences*, *101*(27), 9971–9975. https://doi.org/10.1073/pnas.0403720101

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T.-C., Mendell, J. T., & Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, *33*(3), Article 3. https://doi.org/10.1038/nbt.3122

Pien, S., Fleury, D., Mylne, J. S., Crevillen, P., Inzé, D., Avramova, Z., Dean, C., & Grossniklaus, U. (2008). ARABIDOPSIS TRITHORAX1 Dynamically Regulates FLOWERING LOCUS C Activation via Histone 3 Lysine 4 Trimethylation. *The Plant Cell*, *20*(3), 580–588. https://doi.org/10.1105/tpc.108.058172

Piskurewicz, U., Iwasaki, M., Susaki, D., Megies, C., Kinoshita, T., & Lopez-Molina, L. (2016). Dormancy-specific imprinting underlies maternal inheritance of seed dormancy in Arabidopsis thaliana. *ELife*, *5*. https://doi.org/10.7554/eLife.19573

Polowick, P., & Sawhney, V. (1988). *High Temperature Induced Male and Female Sterility in Canola (Brassica napus L.)*. https://doi.org/10.1093/OXFORDJOURNALS.AOB.A087639

Porter, J. R., & Gawith, M. (1999). Temperatures and the growth and development of wheat: A review. *European Journal of Agronomy*, *10*(1), 23–36. https://doi.org/10.1016/S1161-0301(98)00047-1

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1093/genetics/155.2.945

Qiu, D., Morgan, C., Shi, J., Long, Y., Liu, J., Li, R., Zhuang, X., Wang, Y., Tan, X., Dietrich, E., Weihmann, T., Everett, C., Vanstraelen, S., Beckett, P., Fraser, F., Trick, M., Barnes, S., Wilmer, J., Schmidt, R., … Bancroft, I. (2006). A comparative linkage map of oilseed rape and its use for QTL analysis of seed oil and erucic acid content. *Theoretical and Applied Genetics*, *114*(1), 67–80. https://doi.org/10.1007/s00122-006-0411-2

Quint, M., Ito, H., Zhang, W., & Gray, W. M. (2005). Characterization of a novel temperature-sensitive allele of the CUL1/AXR6 subunit of SCF ubiquitin-ligases. *The Plant Journal: For Cell and Molecular Biology*, *43*(3), 371–383. https://doi.org/10.1111/j.1365-313X.2005.02449.x

Rahaman, M., Mamidi, S., & Rahman, M. (2018). Genome-wide association study of heat stress-tolerance traits in spring-type Brassica napus L. under controlled conditions. *The Crop Journal*, *6*(2), 115–125. https://doi.org/10.1016/j.cj.2017.08.003

Raissig, M. T., Bemer, M., Baroux, C., & Grossniklaus, U. (2013). Genomic Imprinting in the Arabidopsis Embryo Is Partly Regulated by PRC2. *PLOS Genetics*, *9*(12), e1003862. https://doi.org/10.1371/journal.pgen.1003862

Rashid, M., Hampton, J. G., Rolston, M. P., Khan, K. M., & Saville, D. J. (2018). Heat stress during seed development affects forage brassica (Brassica napus L.) seed quality. *Journal of Agronomy and Crop Science*, *204*(2), 147–154. https://doi.org/10.1111/jac.12251

Ratcliffe, O. J., Kumimoto, R. W., Wong, B. J., & Riechmann, J. L. (2003). Analysis of the Arabidopsis MADS AFFECTING FLOWERING Gene Family: MAF2 Prevents Vernalization by Short Periods of Cold [W]. *The Plant Cell*, *15*(5), 1159–1169. https://doi.org/10.1105/tpc.009506

Raz, V., Bergervoet, J. H., & Koornneef, M. (2001). Sequential steps for developmental arrest in Arabidopsis seeds. *Development (Cambridge, England)*, *128*(2), 243–252.

Raza, A., Charagh, S., Zahid, Z., Mubarik, M. S., Javed, R., Siddiqui, M. H., & Hasanuzzaman, M. (2021). Jasmonic acid: A key frontier in conferring abiotic stress

tolerance in plants. *Plant Cell Reports*, *40*(8), 1513–1541. https://doi.org/10.1007/s00299-020-02614-z

Reed, R. C., Bradford, K. J., & Khanday, I. (2022). Seed germination and vigor: Ensuring crop sustainability in a changing climate. *Heredity*, *128*(6), Article 6. https://doi.org/10.1038/s41437-022-00497-2

Roach, D. A., & Wulff, R. D. (2008). *Maternal Effects in Plants*. *Annual Review of Ecology and Systematics, Vol. 18 (1987), pp. 209–235*.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative Genomics Viewer. *Nature Biotechnology*, *29*(1), 24–26. https://doi.org/10.1038/nbt.1754

Rosenzweig, C., Elliott, J., Deryng, D., Ruane, A. C., Müller, C., Arneth, A., Boote, K. J., Folberth, C., Glotter, M., Khabarov, N., Neumann, K., Piontek, F., Pugh, T. A. M., Schmid, E., Stehfest, E., Yang, H., & Jones, J. W. (2014). Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proceedings of the National Academy of Sciences*, *111*(9), 3268–3273. https://doi.org/10.1073/pnas.1222463110

Roy, J., Shaikh, T. M., del Río Mendoza, L., Hosain, S., Chapara, V., & Rahman, M. (2021). Genome-wide association mapping and genomic prediction for adult stage sclerotinia stem rot resistance in Brassica napus (L) under field environments. *Scientific Reports*, *11*(1), Article 1. https://doi.org/10.1038/s41598-021-01272-9

Sanhewe, A. J., Ellis, R. H., Hong, T. D., Wheeler, T. R., Batts, G. R., Hadley, P., & Morison, J. I. L. (1996). The effect of temperature and CO2 on seed quality development in wheat

(Triticum aestivum L.). *Journal of Experimental Botany*, *47*(5), 631–637. https://doi.org/10.1093/jxb/47.5.631

Santos-Mendoza, M., Dubreucq, B., Baud, S., Parcy, F., Caboche, M., & Lepiniec, L. (2008). Deciphering gene regulatory networks that control seed development and maturation in Arabidopsis. *The Plant Journal*, *54*(4), 608–620. https://doi.org/10.1111/j.1365-313X.2008.03461.x

Saunders, R., & Deane, C. M. (2010). Synonymous codon usage influences the local protein structure observed. *Nucleic Acids Research*, *38*(19), 10.

Schiessl, S. (2020). Regulation and Subfunctionalization of Flowering Time Genes in the Allotetraploid Oil Crop Brassica napus. *Frontiers in Plant Science*, *11*. https://www.frontiersin.org/articles/10.3389/fpls.2020.605155

Schiessl, S. V., Quezada-Martinez, D., Tebartz, E., Snowdon, R. J., & Qian, L. (2019). The vernalisation regulator FLOWERING LOCUS C is differentially expressed in biennial and annual Brassica napus. *Scientific Reports*, *9*(1), Article 1. https://doi.org/10.1038/s41598-019-51212-x

Schönrock, N., Bouveret, R., Leroy, O., Borghi, L., Köhler, C., Gruissem, W., & Hennig, L. (2006). Polycomb-group proteins repressthe floral activator AGL19 in the FLC-independent vernalization pathway. *Genes & Development*, *20*(12), 1667–1678. https://doi.org/10.1101/gad.377206

Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., & Scott, R. J. (2006). The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division,

and the size of seeds and other organs. *Development (Cambridge, England)*, *133*(2), 251–261. https://doi.org/10.1242/dev.02194

Schwechheimer, C., & Mergner, J. (2014). The NEDD8 modification pathway in plants. *Frontiers in Plant Science*, *5*. https://www.frontiersin.org/articles/10.3389/fpls.2014.00103

Scott, R. J., Spielman, M., Bailey, J., & Dickinson, H. G. (1998). Parent-of-origin effects on seed development in Arabidopsis thaliana. *Development (Cambridge, England)*, *125*(17), 3329–3341. https://doi.org/10.1242/dev.125.17.3329

Semenov, M. A. (2007). Development of high-resolution UKCIP02-based climate change scenarios in the UK. *Agricultural and Forest Meteorology*, *144*(1), 127–138. https://doi.org/10.1016/j.agrformet.2007.02.003

Semenov, M. A. (2008). Impacts of climate change on wheat in England and Wales. *Journal of The Royal Society Interface*. https://doi.org/10.1098/rsif.2008.0285

Shah, D., & Shaikh, A. R. (2016). Interaction of arginine, lysine, and guanidine with surface residues of lysozyme: Implication to protein stability. *Journal of Biomolecular Structure & Dynamics*, *34*(1), 104–114. https://doi.org/10.1080/07391102.2015.1013158

Shah, S., Weinholdt, C., Jedrusik, N., Molina, C., Zou, J., Große, I., Schiessl, S., Jung, C., & Emrani, N. (2018). Whole-transcriptome analysis reveals genetic factors underlying flowering time regulation in rapeseed (Brassica napus L.). *Plant, Cell & Environment*, *41*(8), 1935–1947. https://doi.org/10.1111/pce.13353

Sheldon, C. C., Hills, M. J., Lister, C., Dean, C., Dennis, E. S., & Peacock, W. J. (2008). Resetting of FLOWERING LOCUS C expression after epigenetic repression by vernalization. *Proceedings of the National Academy of Sciences*, *105*(6), 2214–2219. https://doi.org/10.1073/pnas.0711453105

Sheldon, C. C., Rouse, D. T., Finnegan, E. J., Peacock, W. J., & Dennis, E. S. (2000). The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences*, *97*(7), 3753–3758. https://doi.org/10.1073/pnas.97.7.3753

Shen, G., Sun, W., Chen, Z., Shi, L., Hong, J., & Shi, J. (2022). Plant GDSL Esterases/Lipases: Evolutionary, Physiological and Molecular Functions in Plant Development. *Plants*, *11*(4), Article 4. https://doi.org/10.3390/plants11040468

Shen, W.-H., Parmentier, Y., Hellmann, H., Lechner, E., Dong, A., Masson, J., Granier, F., Lepiniec, L., Estelle, M., & Genschik, P. (2002). Null Mutation of AtCUL1 Causes Arrest in Early Embryogenesis in Arabidopsis. *Molecular Biology of the Cell*, *13*(6), 1916–1928. https://doi.org/10.1091/mbc.E02-02-0077

Shi, J., Li, R., Qiu, D., Jiang, C., Long, Y., Morgan, C., Bancroft, I., Zhao, J., & Meng, J. (2009). Unraveling the Complex Trait of Crop Yield With Quantitative Trait Loci Mapping in Brassica napus. *Genetics*, *182*(3), 851–861. https://doi.org/10.1534/genetics.109.101642

Shi, L., Song, J., Guo, C., Wang, B., Guan, Z., Yang, P., Chen, X., Zhang, Q., King, G. J., Wang, J., & Liu, K. (2019). A CACTA-like transposable element in the upstream region

of BnaA9.CYP78A9 acts as an enhancer to increase silique length and seed weight in rapeseed. *The Plant Journal*, *98*(3), 524–539. https://doi.org/10.1111/tpj.14236

ShiBo, F., KaiYan, T., & SanXue, R. (2010). Winter wheat yields decline with spring higher night temperature by controlled experiments. *Scientia Agricultura Sinica*, *43*(15), 3251–3258.

Sieber, P., Gheyselinck, J., Groß-Hardt, R., Laux, T., Grossniklaus, U., & Schneitz, K. (2004). Pattern formation during early ovule development in Arabidopsis thaliana. *Developmental Biology*. https://doi.org/10.1016/J.YDBIO.2004.05.037

Singh, A. K., Yadav, B. S., Dhanapal, S., Berliner, M., Finkelshtein, A., & Chamovitz, D. A. (2019). CSN5A Subunit of COP9 Signalosome Temporally Buffers Response to Heat in Arabidopsis. *Biomolecules*, *9*(12), Article 12. https://doi.org/10.3390/biom9120805

SINNIAH, U. R., ELLIS, R. H., & JOHN, P. (1998). Irrigation and Seed Quality Development in Rapid-cycling Brassica: Seed Germination and Longevity. *Annals of Botany*, *82*(3), 309–314. https://doi.org/10.1006/anbo.1998.0748

Song, G., Olatunji, D., Montes, C., Clark, N. M., Pu, Y., Kelley, D. R., & Walley, J. W. (2021). *Quantitative proteomics reveals extensive lysine ubiquitination in the Arabidopsis root proteome and uncovers novel transcription factor stability states* [Preprint]. Plant Biology. https://doi.org/10.1101/2021.01.07.425780

Springthorpe, V., & Penfield, S. (2015). Flowering time and seed dormancy control use external coincidence to generate life history strategy. *ELife*, *4*. https://doi.org/10.7554/eLife.05557

Steiner-Lange, S., Unte, U. S., Eckstein, L., Yang, C., Wilson, Z. A., Schmelzer, E., Dekker, K., & Saedler, H. (2003). Disruption of Arabidopsis thaliana MYB26 results in male sterility due to non-dehiscent anthers. *The Plant Journal*, *34*(4), 519–528. https://doi.org/10.1046/j.1365-313X.2003.01745.x

Steuernagel, B., Woodhouse, S., He, Z., Hepworth, J., Tidy, A., Siles-Suarez, L., Havlikova, L., Calderwood, A., Woolfenden, H., Chinoy, C., Bilham, L., Corke, F., Graham, N., Doonan, J., Teakle, G., Eastmond, P., Morris, R., Wilson, Z., Kurup, S., … Ostergaard, L. (2021). *BRAVO target sequence capture V3* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4473283

Stevens, M., McGrann, G., & Clark, B. (2008). *Turnip yellows virus (syn Beet western yellows virus): An emerging threat to European oilseed rape production?*

Stone, S. L., Braybrook, S. A., Paula, S. L., Kwong, L. W., Meuser, J., Pelletier, J., Hsieh, T.-F., Fischer, R. L., Goldberg, R. B., & Harada, J. J. (2008). Arabidopsis LEAFY COTYLEDON2 induces maturation traits and auxin activity: Implications for somatic embryogenesis. *Proceedings of the National Academy of Sciences*, *105*(8), 3151–3156. https://doi.org/10.1073/pnas.0712364105

Su, H.-G., Zhang, X.-H., Wang, T.-T., Wei, W.-L., Wang, Y.-X., Chen, J., Zhou, Y.-B., Chen, M., Ma, Y.-Z., Xu, Z.-S., & Min, D.-H. (2020). Genome-Wide Identification, Evolution, and Expression of GDSL-Type Esterase/Lipase Gene Family in Soybean. *Frontiers in Plant Science*, *11*. https://www.frontiersin.org/articles/10.3389/fpls.2020.00726

Sun, J., Qi, L., Li, Y., Chu, J., & Li, C. (2012). PIF4–Mediated Activation of YUCCA8 Expression Integrates Temperature into the Auxin Pathway in Regulating Arabidopsis

Hypocotyl Growth. *PLOS Genetics*, *8*(3), e1002594. https://doi.org/10.1371/journal.pgen.1002594

Sun, S., Yu-qing, H., He-xuan, W., Jing-bin, J., Huan-huan, Y., & Xiang-yang, X. (2022). Genome-wide identification and expression analysis of GDSL esterase/lipase genes in tomato. *Journal of Integrative Agriculture*, *21*(2), 389–406. https://doi.org/10.1016/S2095-3119(20)63461-X

Sundaresan, V. (2005). Control of seed size in plants. *Proceedings of the National Academy of Sciences*, *102*(50), 17887–17888. https://doi.org/10.1073/pnas.0509021102

Sung, S., & Amasino, R. M. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature*, *427*(6970), Article 6970. https://doi.org/10.1038/nature02195

Swiezewski, S., Liu, F., Magusin, A., & Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an *Arabidopsis* Polycomb target. *Nature*, *462*(7274), 799–802. https://doi.org/10.1038/nature08618

Tadege, M., Sheldon, C. C., Helliwell, C. A., Stoutjesdijk, P., Dennis, E. S., & Peacock, W. J. (2001). Control of flowering time by FLC orthologues in Brassica napus. *The Plant Journal: For Cell and Molecular Biology*, *28*(5), 545–553. https://doi.org/10.1046/j.1365-313x.2001.01182.x

Talbert, P. B., Adler, H. T., Parks, D. W., & Comai, L. (1995). The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of Arabidopsis thaliana. *Development*, *121*(9), 2723–2735.

Tao, Z., Hu, H., Luo, X., Jia, B., Du, J., & He, Y. (2019). Embryonic resetting of the parental vernalized state by two B3 domain transcription factors in Arabidopsis. *Nature Plants*, *5*(4), 424–435. https://doi.org/10.1038/s41477-019-0402-3

Tao, Z., Shen, L., Gu, X., Wang, Y., Yu, H., & He, Y. (2017). Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. *Nature*, *551*(7678), 124–128. https://doi.org/10.1038/nature24300

Team, R. C. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org

Thomashow, M. F. (1999). PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms. *Annual Review of Plant Physiology and Plant Molecular Biology*, *50*(1), 571–599. https://doi.org/10.1146/annurev.arplant.50.1.571

Tian, Q., & Reed, J. W. (1999). Control of auxin-regulated root development by the Arabidopsis thaliana SHY2/IAA3 gene. *Development*, *126*(4), 711–721. https://doi.org/10.1242/dev.126.4.711

To, A., Valon, C., Savino, G., Guilleminot, J., Devic, M., Giraudat, J., & Parcy, F. (2006). A Network of Local and Redundant Gene Regulation Governs Arabidopsis Seed Maturation. *The Plant Cell*, *18*(7), 1642–1651. https://doi.org/10.1105/tpc.105.039925

Tommey, A. M., & Evans, E. J. (1991). Temperature and daylength control of flower initiation in winter oilseed rape (Brassica napus L.). *Annals of Applied Biology*, *118*(1), 201–208. https://doi.org/10.1111/j.1744-7348.1991.tb06098.x

Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, *4*(1), Article 1. https://doi.org/10.1038/nclimate2067

Voegele, A., Linkies, A., Müller, K., & Leubner-Metzger, G. (2011). Members of the gibberellin receptor gene family GID1 (GIBBERELLIN INSENSITIVE DWARF1) play distinct roles during Lepidium sativum and Arabidopsis thaliana seed germination. *Journal of Experimental Botany*, *62*(14), 5131–5147. https://doi.org/10.1093/jxb/err214

Walker-Simmons, M. (1988). Enhancement of ABA responsiveness in wheat embryos by high temperature*†. *Plant, Cell & Environment*, *11*(8), 769–775. https://doi.org/10.1111/j.1365-3040.1988.tb01161.x

Wan, J., Griffiths, R., Ying, J., McCourt, P., & Huang, Y. (2009). Development of Drought-Tolerant Canola (Brassica napus L.) through Genetic Modulation of ABA-mediated Stomatal Responses. *Crop Science*, *49*(5), 1539–1554. https://doi.org/10.2135/cropsci2008.09.0568

Wang, A., Garcia, D., Zhang, H., Feng, K., Chaudhury, A., Berger, F., Peacock, W. J., Dennis, E. S., & Luo, M. (2010). The VQ motif protein IKU1 regulates endosperm growth and seed size in Arabidopsis. *The Plant Journal*, *63*(4), 670–679. https://doi.org/10.1111/j.1365-313X.2010.04271.x

Wei, L., Jian, H., Lu, K., Filardo, F., Yin, N., Liu, L., Qu, C., Li, W., Du, H., & Li, J. (2016). Genome-wide association analysis and differential expression analysis of resistance to

Sclerotinia stem rot in Brassica napus. *Plant Biotechnology Journal*, *14*(6), 1368–1380. https://doi.org/10.1111/pbi.12501

Wei, Z., Yuan, Q., Lin, H., Li, X., Zhang, C., Gao, H., Zhang, B., He, H., Liu, T., Jie, Z., Gao, X., Shi, S., Wang, B., Gao, Z., Kong, L., Qian, Q., & Shang, L. (2021). Linkage analysis, GWAS, transcriptome analysis to identify candidate genes for rice seedlings in response to high temperature stress. *BMC Plant Biology*, *21*(1), 85. https://doi.org/10.1186/s12870-021-02857-2

Whittaker, C., & Dean, C. (2017). The FLC Locus: A Platform for Discoveries in Epigenetics and Adaptation. *Annual Review of Cell and Developmental Biology*, *33*, 555–575. https://doi.org/10.1146/annurev-cellbio-100616-060546

Whittle, C. A., Otto, S. P., Johnston, M. O., & Krochko, J. E. (2009). Adaptive epigenetic memory of ancestral temperature regime in Arabidopsis thalianaThis paper is one of a selection of papers published in a Special Issue from the National Research Council of Canada – Plant Biotechnology Institute. *Botany*, *87*(6), 650–657. https://doi.org/10.1139/B09-030

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. https://ggplot2.tidyverse.org.

Wickham, H., François, R., Henry, L., Müller, K., & RStudio. (2022). *dplyr: A Grammar of Data Manipulation* (1.0.10). https://CRAN.R-project.org/package=dplyr

Wilkinson, M. D., & Haughn, G. W. (1995). UNUSUAL FLORAL ORGANS Controls Meristem Identity and Organ Primordia Fate in Arabidopsis. *The Plant Cell*, *7*(9), 1485–1499. https://doi.org/10.1105/tpc.7.9.1485

Willer, C. J., Li, Y., & Abecasis, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics*, *26*(17), 2190–2191. https://doi.org/10.1093/bioinformatics/btq340

Wójcikowska, B., Jaskóła, K., Gąsiorek, P., Meus, M., Nowak, K., & Gaj, M. D. (2013). LEAFY COTYLEDON2 (LEC2) promotes embryogenic induction in somatic tissues of Arabidopsis, via YUCCA-mediated auxin biosynthesis. *Planta*, *238*(3), 425–440. https://doi.org/10.1007/s00425-013-1892-2

Wollenberg, A. C., & Amasino, R. M. (2012). Natural variation in the temperature range permissive for vernalization in accessions of Arabidopsis thaliana. *Plant, Cell & Environment*, *35*(12), 2181–2191. https://doi.org/10.1111/j.1365-3040.2012.02548.x

Woodhouse, S., He, Z., Woolfenden, H., Steuernagel, B., Haerty, W., Bancroft, I., Irwin, J. A., Morris, R. J., & Wells, R. (2021). Validation of a novel associative transcriptomics pipeline in Brassica oleracea: Identifying candidates for vernalisation response. *BMC Genomics*, *22*, 539. https://doi.org/10.1186/s12864-021-07805-w

Wu, D., Liang, Z., Yan, T., Xu, Y., Xuan, L., Tang, J., Zhou, G., Lohwasser, U., Hua, S., Wang, H., Chen, X., Wang, Q., Zhu, L., Maodzeka, A., Hussain, N., Li, Z., Li, X., Shamsi, I. H., Jilani, G., … Jiang, L. (2019). Whole-Genome Resequencing of a Worldwide Collection of Rapeseed Accessions Reveals the Genetic Basis of Ecotype Divergence. *Molecular Plant*, *12*(1), 30–43. https://doi.org/10.1016/j.molp.2018.11.007

Wu, J., Zhao, Q., Liu, S., Shahid, M., Lan, L., Cai, G., Zhang, C., Fan, C., Wang, Y., & Zhou, Y. (2016). Genome-wide Association Study Identifies New Loci for Resistance to

Sclerotinia Stem Rot in Brassica napus. *Frontiers in Plant Science*, *7*. https://www.frontiersin.org/articles/10.3389/fpls.2016.01418

Xi, X., Hu, Z., Nie, X., Meng, M., Xu, H., & Li, J. (2021). Cross Inhibition of MPK10 and WRKY10 Participating in the Growth of Endosperm in Arabidopsis thaliana. *Frontiers in Plant Science*, *12*, 521. https://doi.org/10.3389/fpls.2021.640346

Xiao, Y.-G., Sun, Q.-B., Kang, X.-J., Chen, C.-B., & Ni, M. (2016). SHORT HYPOCOTYL UNDER BLUE1 or HAIKU2 mixepression alters canola and Arabidopsis seed development. *New Phytologist*, *209*(2), 636–649. https://doi.org/10.1111/nph.13632

Yan, X., Zhang, L., Chen, B., Xiong, Z., Chen, C., Wang, L., ... & Wei, W. (2012). Functional identification and characterization of the Brassica Napus transcription factor gene BnAP2, the ortholog of Arabidopsis thaliana APETALA2. *PLoS One*, *7*(3). https://doi.org/10.1371/journal.pone.0033890

Yang, C., Xu, Z., Song, J., Conner, K., Vizcay Barrena, G., & Wilson, Z. A. (2007). Arabidopsis MYB26/MALE STERILE35 Regulates Secondary Thickening in the Endothecium and Is Essential for Anther Dehiscence. *The Plant Cell*, *19*(2), 534–548. https://doi.org/10.1105/tpc.106.046391

Yang, C., Xue-kun, Z., Chong-shun, Z. O. U., Yong, C., Pu-ying, Z., & Gui-ying, L. I. (2007). Effects of drought simulated by PEG-6000 on germination and seedling growth of rapeseed (Brassica napus L.). *CHINESE JOURNAL OF OIL CROP SCIENCES*, *29*(4), 425.

Yang, P., Shu, C., Chen, L., Xu, J., Wu, J., & Liu, K. (2012). Identification of a major QTL for silique length and seed weight in oilseed rape (Brassicanapus L.). *Theoretical and Applied Genetics*, *125*(2), 285–296. https://doi.org/10.1007/s00122-012-1833-7

Yang, X., Chen, F., Lin, X., Liu, Z., Zhang, H., Zhao, J., Li, K., Ye, Q., Li, Y., Lv, S., Yang, P., Wu, W., Li, Z., Lal, R., & Tang, H. (2015). Potential benefits of climate change for crop productivity in China. *Agricultural and Forest Meteorology*, *208*, 76–84. https://doi.org/10.1016/j.agrformet.2015.04.024

Yang, X., Wang, K., Bu, Y., Niu, F., Ge, L., Zhang, L., & Song, X. (2022). Genome-wide analysis of GELP gene family in wheat and validation of TaGELP073 involved in anther and pollen development. *Environmental and Experimental Botany*, *200*, 104914. https://doi.org/10.1016/j.envexpbot.2022.104914

Yant, L., Mathieu, J., Dinh, T. T., Ott, F., Lanz, C., Wollmann, H., Chen, X., & Schmid, M. (2010). Orchestration of the Floral Transition and Floral Development in Arabidopsis by the Bifunctional Transcription Factor APETALA2. *The Plant Cell*, *22*(7), 2156–2170. https://doi.org/10.1105/tpc.110.075606

Yin, S., Wan, M., Guo, C., Wang, B., Li, H., Li, G., Tian, Y., Ge, X., King, G. J., Liu, K., Li, Z., & Wang, J. (2020). Transposon insertions within alleles of BnaFLC.A10 and BnaFLC.A2 are associated with seasonal crop type in rapeseed. *Journal of Experimental Botany*, *71*(16), 4729–4741. https://doi.org/10.1093/jxb/eraa237

Young, L. W., Wilen, R. W., & Bonham-Smith, P. C. (2004). High temperature stress of Brassica napus during flowering reduces micro- and megagametophyte fertility, induces

fruit abortion, and disrupts seed production. *Journal of Experimental Botany*, *55*(396), 485–495. https://doi.org/10.1093/jxb/erh038

Yun, H., Hyun, Y., Kang, M.-J., Noh, Y.-S., Noh, B., & Choi, Y. (2011). Identification of regulators required for the reactivation of FLOWERING LOCUS C during Arabidopsis reproduction. *Planta*, *234*(6), 1237–1250. https://doi.org/10.1007/s00425-011-1484-y

Zhang, F., Wang, H., Kalve, S., Wolabu, T. W., Nakashima, J., Golz, J., & Tadege, M. (2019). Control of leaf blade outgrowth and floral organ development by LEUNIG, ANGUSTIFOLIA3 and WOX transcriptional regulators. *The New Phytologist*. https://doi.org/10.1111/nph.15921

Zhang, L., Li, S., Chen, L., & Yang, G. (2012). Identification and mapping of a major dominant quantitative trait locus controlling seeds per silique as a single Mendelian factor in Brassica napus L. *Theoretical and Applied Genetics*, *125*(4), 695–705. https://doi.org/10.1007/s00122-012-1861-3

Zhang, X., Lu, G., Long, W., Zou, X., Li, F., & Nishio, T. (2014). Recent progress in drought and salt tolerance studies in Brassica crops. *Breeding Science*, *64*(1), 60–73. https://doi.org/10.1270/jsbbs.64.60

Zhang, Y., Zhang, B., Yan, D., Dong, W., Yang, W., Li, Q., Zeng, L., Wang, J., Wang, L., Hicks, L. M., & He, Z. (2011). Two Arabidopsis cytochrome P450 monooxygenases, CYP714A1 and CYP714A2, function redundantly in plant development through gibberellin deactivation. *The Plant Journal*, *67*(2), 342–353. https://doi.org/10.1111/j.1365-313X.2011.04596.x

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J.-L., Elliott, J., Ewert, F., Janssens, I. A., Li, T., Lin, E., Liu, Q., Martre, P., Müller, C., … Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences*, *114*(35), 9326–9331. https://doi.org/10.1073/pnas.1701762114

Zhao, Y., Zhu, P., Hepworth, J., Bloomer, R., Antoniou-Kourounioti, R. L., Doughty, J., Heckmann, A., Xu, C., Yang, H., & Dean, C. (2021). Natural temperature fluctuations promote COOLAIR regulation of FLC. *Genes & Development*, *35*(11–12), 888–898. https://doi.org/10.1101/gad.348362.121

Zheng, J., Yang, X., Harrell, J. M., Ryzhikov, S., Shim, E.-H., Lykke-Andersen, K., Wei, N., Sun, H., Kobayashi, R., & Zhang, H. (2002). CAND1 Binds to Unneddylated CUL1 and Regulates the Formation of SCF Ubiquitin E3 Ligase Complex. *Molecular Cell*, *10*(6), 1519–1526. https://doi.org/10.1016/S1097-2765(02)00784-0

Zinn, K. E., Tunc-Ozdemir, M., & Harper, J. F. (2010). Temperature stress and plant sexual reproduction: Uncovering the weakest links. *Journal of Experimental Botany*, *61*(7), 1959–1968. https://doi.org/10.1093/jxb/erq053

Zou, X., Suppanz, I., Raman, H., Hou, J., Wang, J., Long, Y., Jung, C., & Meng, J. (2012). Comparative Analysis of FLC Homologues in Brassicaceae Provides Insight into Their Role in the Evolution of Oilseed Rape. *PLoS ONE*, *7*(9), e45751. https://doi.org/10.1371/journal.pone.0045751