Norwegian University of Life Sciences
Faculty of Biosciences
Department of Plant Science

Philosophiae Doctor (PhD)
Thesis 2022:36

A study of freezing tolerance in red clover (*Trifolium pratense* L.)

En studie av frysetoleranse i rødkløver (*Trifolium pratense* L.)

Stefano Zanotto

A study of freezing tolerance in red clover (*Trifolium pratense* L.)

En studie av frysetoleranse i rødkløver (*Trifolium pratense* L.)

Philosophiae Doctor (PhD) Thesis

Stefano Zanotto

Norwegian University of Life Sciences
Faculty of Biosciences
Department of Plant Science

Ås (2022)

Thesis number 2022:36
ISSN 1894-6402
ISBN 978-82-575-1909-4

Supervisors and Evaluation Committee

The candidate's supervisory group consists of:

Main Supervisor: Ass. Prof. Åshild Ergon, NMBU
Co-supervisor: Prof. Jorunn Elisabeth Olsen, NMBU
Co-supervisor: Prof. Odd Arne Rognli, NMBU

Members of the evaluation committee are:

Dr. Beat Boller
Address: Agroscope, Langwiesstrasse 14,
CH-8050 Zürich, Switzerland
Phone: +41 76 393 09 06
E-mail: beat.c.boller@bluewin.ch

Dr. Heathcliffe Riday
Address: USDA-ARS, U.S. Dairy Forage Research Center
1925 Linden Drive West, Madison,
WI 53706 USA
Phone: 608-890-0077
E-mail: heathcliffe.riday@usda.gov

Committee coordinator:
Prof. Sissel Torre
Address: Faculty of Biosciences, Department of Plant Sciences
P.O. Box 5003 NMBU, NO-1432 Ås, NORWAY
Phone: +47 67232828
E-mail: sissel.torre@nmbu.no

Acknowledgements

This PhD project was financed mainly by NMBU and was partly realized within the framework of the EUCLEG project (EU project No. 727312). In addition, I have used data obtained in the Arctic Clover project, which received funding from the Nordic Council of Ministers' Arctic Co-operation Programme, and data obtained in a study funded by Nansenfondet.

I would like to express my gratitude to my main supervisor Associate Prof. Åshild Ergon for giving me the opportunity to work in this PhD project and for being available to help. Thank you very much for your support throughout this PhD and especially for your guidance during the last period up to the defence. Thank you also for giving me the opportunity to participate at international projects and to introduce me to several scientists working in the plant sciences in Norway and abroad. Thank you also for the confrontations we had during these years, I personally learned a lot from the discussions on things that we disagreed on. It helped me to be more critical on my own ideas and convictions, and open to change them. At the same time, I learned to be more self-confident and argue for my own ideas and beliefs.
I also want to thank my co-supervisors Prof. Jorunn E. Olsen and Prof. Odd Arne Rognli for their useful contributions to the planning and the evaluation of my research, as well as for all your suggestions on the data analyses and especially the contributions to the writing of the manuscripts and supervision in writing this thesis.

I would also like to thank all the other co-authors for sharing your expertise and for all your assistance in writing the manuscripts. Thanks in particular to Annick Bertrand from, and Randy W. Purves for your help in planning and completing the analyses and writing of paper 2.

Additionally, I would like to thank the technical support from our technicians. A personal thank you goes to Liv Berge and especially Øyvind Jørgensen who provided me with much more than just technical assistance in these years of work as PhD. Especially at the beginning of my experience you have been an invaluable source of human relations, and through our improvised “language and culture sessions” you really helped me to integrate in the working and social environment. Also, I want to thank Tina Schmaucks for her technical help in the greenhouse.

Then, I want to express my gratitude to all the members of the Genetics and Plant Breeding section at IPV. Thanks to everyone for being welcoming and for creating a friendly working environment. Despite the special times we have been through in the last two years, I think that we manged to maintain a sense of collaboration and desire of doing things together. In particular I want to thank personally my office mates and friends Martin and Camilla. You have been much more than colleagues for me and as we known I would say that we have been more like a family. Thanks a lot for your support in solving problems at work and especially for always having a word of comfort during the hard times, and at the same time a happy smile in the positive ones. A special thank you goes to Martin for all the time and patience that you used to teach me statistical analyses in R and for helping me in all bureaucratic problems typical of PhD life.

Finally, I want to thank my family and friends. There are too many that I would like to thank personally, but there would not be enough space here. I just want to say that you all owe a big credit for what I manage to achieve during my PhD. There is only one person that I want to name here, my mother Rita. I cannot spend enough words for thanking you for all the support that you gave me in those years. Thanks for all the times you answered the phone in the middle of the night, for your patience, your forgiveness, and your love.

Ås, April 2022

Stefano

Table of Contents

1 Abbreviations and definitions

ANOVA: analysis of variance
CA: cold acclimation
CCA: canonical correlation analysis
DAD: diode array detector
FS: freezing susceptible
FT: freezing tolerance/freezing tolerant
GABA: gamma aminobutyric acid
GBS: genotyping by sequencing
GRM: genomic relationship matrix
GWAS: genome wide association studies
GxE: genotype by environment
HPLC: high-performance liquid chromatography
HRMS: high-resolution mass spectrometry
Kbp: kilo base pair
LC: liquid chromatography
LC-MS: liquid chromatography-mass spectrometry
LD: linkage disequilibrium
LIS: legume information system
MAGIC: multiparent advanced generation intercross
MFA: multi factor analysis
MLMM: multi-locus mixed model
QTL: quantitative trait locus
REML: residual maximum likelihood
PCA: principal component analysis
PPO: polyphenol oxidase
SNP: single nucleotide polymorphism
UPLC: ultra-performance liquid chromatography
WS: winter survival

2 List of papers

I. Zanotto, S., Palmé, A., Helgadóttir, Á., Daugstad, K., Isolahti, M., Öhlund, L., Marum, P., Moen, M.A., Veteläinen, M., Rognli, O.A., Ergon, Å., 2021. Trait characterization of genetic resources reveals useful variation for the improvement of cultivated Nordic red clover. Journal of Agronomy and Crop Science 207, 492–503. doi.org/10.1111/jac.12487

II. Zanotto, S., Bertrand, A., Purves, R.P., Olsen, J.E., Elessawy, F.M., Ergon, Å., 2022. Biochemical changes after cold acclimation in Nordic red clover (*Trifolium pratense* L.) accessions with contrasting levels of freezing tolerance.
Submitted manuscript

III. Zanotto, S., Ruttink, T., Pegard, M., Skøt, L., Grieder, C., Kölliker, R., Ergon, Å, 2022. A genome-wide association study of freezing tolerance in red clover (*Trifolium pratense* L.).
Manuscript

3 Abstract

The ability to tolerate freezing temperatures is one of the traits that allows successful overwintering of red clover (*Trifolium pratense* L.) at Nordic latitudes. Freezing tolerance (FT) is enhanced during a period of cold acclimation (CA) at low above freezing temperatures, which induces major changes at the transcriptional, proteomic and metabolic level in perennial plants. However, in Nordic red clover, little is known about biochemical changes induced by CA and how these are related to the level of FT. Furthermore, even if persistency has been one of the main breeding goals in red clover in the Nordic countries since systematic breeding started, there is a lack of knowledge on the morpho-physiological traits associated with improved FT as well as its genetic control in red clover. Genetic resources conserved in gene banks may be an important source for the improvement of FT and thereby winter survival of red clover. However, there are very few recent studies which characterized gene bank material of Nordic origin for traits related to winter survival and no studies specifically characterized FT in this species. A better understanding of FT from the biochemical and genetic point of view, as well as the characterization of available genetic resources for FT and other morpho-physiological traits related to it is of primary importance for further improvement of this trait through selection and breeding.

Overall, the aim of this PhD project was to conduct a comprehensive study of freezing tolerance in red clover, with three main sub-goals. The first was to assess the phenotypic variation for FT among gene bank accessions of Nordic origin, also considering its association with other morpho-physiological traits (research component 1). The second was to investigate the major biochemical changes induced by CA and if and how these differ between accessions with different levels of FT (research component 2). Lastly, we wanted to detect genomic regions involved in the control of FT in red clover, possibly identifying candidate genes for the trait (research component 3).

To achieve these goals, we characterized 48 red clover accessions of Nordic origin for several phenotypic traits in a multi-location multi-year field trial and in a controlled condition experiment and compared them with commercial cultivars. We found that traits related to plant size were strongly associated with late flowering and upright growth and differed between landraces/cultivars on the one hand and wild populations on the other. There was a large genotype by environment interaction on

winter survival, which only partially correlated with FT under controlled conditions. Several gene bank accessions had better winter survival than commercial cultivars and large variation for FT was also identified among the former, suggesting that these accessions are a precious genetic resource for future improvement of these traits. The climate at the collection site of gene bank accessions was also significantly associated with phenotypic variation. This suggests that the integration of climatic information in analyses for the characterization of gene bank material can be useful to facilitate the collection, conservation, and utilization of red clover genetic resources in the Nordic region.

We then selected the five accessions with the highest (freezing tolerant group, FT) and the five with the lowest (freezing susceptible group, FS) freezing tolerance. These accessions were characterized for their contents of carbohydrates, amino acids, and phenolic compounds in the plant crowns after a short-term cold acclimation (CA) treatment. The major effects of CA were the depletion of starch reserves and the accumulation of water-soluble sugars and free amino acids. Specific phenolic compounds were also affected in their content by CA. The contents of various carbohydrates (raffinose, pinitol), amino acids (arginine, serine, alanine, valine, phenylalanine) and one phenolic compound (a pinocembrin isomer-G derivative) were more elevated in FT than FS accessions, suggesting that these compounds have a role in the freezing tolerance of red clover.

Finally, 393 red clover accessions of diverse origin were genotyped using genotyping by sequencing (GBS) on pools of individuals, and phenotyped for FT in a controlled conditions experiment. Genome wide association analyses (GWAS) for FT were conducted using both single nucleotide polymorphisms (SNPs) and within GBS loci haplotypes (HTPs) allele frequency data. The analyses detected various genomic regions significantly associated with FT. Ten independent SNP and HTP markers were found within or at a short distance from genes possibly involved in mechanisms associated with FT. The results of this study may be used for the development of molecular markers useful for the improvement of FT in red clover.

Overall, the findings of this PhD project significantly enhanced the knowledge of FT in red clover. We described the phenotypic variation available for this trait and how it is associated with other phenotypic traits. We have now a better understanding of CA at the biochemical level and how this is associated with the level of FT in Nordic red clover. Furthermore, the identification of genomic regions significantly associated with FT is an important step forward for the dissection of the genetic base of this trait and is a starting point for further studies aiming at the

development of genomic tools for the improvement of FT in red clover through breeding.

4 Norsk sammendrag

Evnen til å tåle kuldegrader er en av egenskapene som muliggjør vellykket overvintring av rødkløver (*Trifolium pratense* L.) ved nordiske breddegrader. Frosttoleranse (FT) øker i en periode med kulde-akklimatisering (CA) ved lave temperaturer over frysepunktet, noe som induserer store endringer på transkripsjonelt, proteomisk og metabolsk nivå i flerårige planter. Men lite er kjent om biokjemiske endringer som er indusert av CA og hvordan disse er relatert til nivået av FT i nordisk rødkløver. Og selv om varighet har vært et av hovedmålene i rødkløver-foredling i Norden siden systematisk planteforedling startet, er det mangel på kunnskap om de morfologiske og fysiologiske egenskapene knyttet til forbedret FT, samt genetisk kontroll av FT, i rødkløver. Genetiske ressurser bevart i genbanker kan være en viktig kilde til forbedring av FT og dermed vinteroverlevelse hos rødkløver. Det er imidlertid svært få nyere studier som har karakterisert genbankmateriale av nordisk opprinnelse med tanke på egenskaper relatert til vinteroverlevelse, og ingen studier som spesifikt har karakterisert FT i denne arten. En bedre forståelse av FT fra et biokjemisk og genetisk synspunkt, samt karakterisering av FT og tilknyttede morfologiske og fysiologiske egenskaper i tilgjengelige genetiske ressurser er av grunnleggende betydning for videre forbedring av denne egenskapen gjennom planteforedling.

Det overordnede målet for dette ph.d.-prosjektet var å gjennomføre en omfattende studie av frysetoleranse i rødkløver, med tre delmål. Det første var å evaluere den fenotypiske variasjonen i FT blant genbank-aksesjoner av nordisk opprinnelse, og også ta i betraktning assosiasjoner med andre morfologiske og fysiologiske egenskaper (forskningskomponent 1). Det andre var å undersøke de viktigste biokjemiske endringene indusert av CA og hvordan disse varierer mellom aksesjoner med ulike nivåer av FT (forskningskomponent 2). Til slutt ønsket vi å finne genomiske regioner som er involvert i kontrollen av FT i rødkløver og om mulig identifisere kandidatgener for egenskapen (forskningskomponent 3). For å oppnå disse målene karakteriserte vi 48 rødkløver-aksesjoner av nordisk opprinnelse for flere fenotypiske egenskaper i et flerårig feltforsøk på flere steder og i et forsøk under kontrollerte betingelser, og sammenlignet dem med kommersielle sorter. Vi fant at egenskaper knyttet til plantestørrelse var sterkt assosiert med sen blomstring og oppreist vekst og var ulike i landsorter/sorter på den ene siden og ville populasjoner på den andre siden. Det var en stor genotype-miljøinteraksjon på vinteroverlevelse,

og vinteroverlevelse var bare delvis korrelert med FT under kontrollerte forhold. Flere genbank-aksesjoner hadde bedre vinteroverlevelse enn kommersielle sorter, og stor variasjon for FT ble også identifisert blant de første noe som antyder at disse aksesjonene er en verdifull genetisk ressurs for fremtidig forbedring av disse egenskapene Klimaet på innsamlingsstedet for genbanktilgang var også signifikant assosiert med fenotypisk variasjon. Dette tyder på at integrering av klimatisk informasjon i analyser for karakterisering av genbankmateriale kan være nyttig for å lette innsamling, bevaring og utnyttelse av rødkløverens genetiske ressurser i Norden.

Vi valgte deretter de fem aksesjonene med høyest og de fem med lavest frosttoleranse (hhv. frosttolerant gruppe, FT, og frost-sensitiv gruppe, FS). Disse aksesjonene ble karakterisert for innholdet av karbohydrater, aminosyrer og fenolforbindelser i stengelbasis etter en kortvarig kulde-akklimatisering (CA). De viktigste effektene av CA var reduksjon i stivelsesreserver og akkumulering av vannløselige sukkerarter og frie aminosyrer. Innholdet av spesifikke fenol-forbindelser ble også påvirket av CA. Innholdet av forskjellige karbohydrater (raffinose, pinitol), aminosyrer (arginin, serin, alanin, valin, fenylalanin) og en fenolforbindelse (et pinocembrin-G-derivat) var mer forhøyet i FT enn FS-aksesjoner, noe som tyder på at disse forbindelsene har en rolle i frysetoleransen til rødkløver. Til slutt ble 393 rødkløver-aksesjoner av forskjellig opprinnelse genotypet ved hjelp av «genotyping by sequencing» (GBS) på aksesjons-nivå, ved å slå sammen bladprøver fra mange individer per aksesjon før DNA-ekstraksjon. Aksesjonenes FT ble også målt i et kontrollert eksperiment. Genomiske assosiasjonsanalyser (GWAS) for FT ble utført ved bruk av allelfrekvensdata for både polymorfismer i enkeltnukleotider (SNPs) og for polymorfe haplotyper innen GBS loci (HTPs). Analysene detekterte ulike genomiske regioner som er signifikant assosiert med FT. Ti uavhengige SNP- og HTP-markører ble funnet i eller i nærheten av gener som muligens er involvert i mekanismer assosiert med plantens evne til å tolerere lave frysetemperaturer. Resultatene av denne studien har implikasjoner for utviklingen av molekylære markører som er nyttige for forbedring av FT i rødkløver.

Samlet sett har funnene fra denne studien forbedret kunnskapen om FT i rødkløver betydelig. Vi har beskrevet den fenotypiske variasjonen som er tilgjengelig for denne egenskapen og hvordan dette er assosiert med andre fenotypiske egenskaper. Vi har nå en bedre forståelse av CA på biokjemisk nivå og hvordan dette henger sammen med nivået av FT i nordisk rødkløver. Videre er identifiseringen av genomiske regioner som er signifikant assosiert med FT et viktig skritt fremover for disseksjon av det genetiske grunnlaget til denne egenskapen og er et utgangspunkt

for videre studier som tar sikte på utvikling av genomiske verktøy for forbedring av FT i rødkløver gjennom planteforedling.

5 Synopsis

5.1 Introduction

5.1.1 Red clover

5.1.1.1 Origin, history and diffusion

Red clover (*Trifolium pratense* L.) is a natural diploid (2n=2x=14) although autotetraploid varieties have been produced following chromosome doubling (Annicchiarico et al., 2015). The species, which probably originated near the Mediterranean basin (Boller et al., 2010), remained undomesticated until the second millennium (Kjærgaard, 2003). The first description of cultivated red clover dates back to the end of the 13th century in Moorish Spain, from which the crop progressively spread to other parts of Europe and reached the Nordic countries in the 18th century (Kjærgaard, 2003).

In the United States the species was introduced in the 17th century by European colonists in the east and later it spread to the west and further all over the US (Taylor and Quesenberry, 1996). Red clover cultivation started in Russia in the central and western regions in the 18th century, likely by adapting cultivars introduced from western Europe (Semerikov et al., 2002). Red clover, thanks to the symbiosis with the bacteria *Rhizobium leguminosarum* biovar *trifolii* resulting in atmospheric nitrogen fixation, played a central role in European and North American agriculture as the chief provider of nitrogen for cereals, needed to feed an expanding population (Kjærgaard, 2003). The decline of red clover cultivation started at the beginning of the 20th century when the process of extracting nitrogen through ammonia synthesis was discovered. By the second half of the century the industrial production of nitrogen expanded dramatically, leading to the disappearance of red clover from many areas (Taylor, 2008). Recently an increased interest has developed in the cultivation of forage legumes because of the energy and environmental costs related to the use of synthetic nitrogen. Current estimates of the area cultivated with red clover amounts to four million hectares worldwide (Isobe et al., 2013).

5.1.1.2 Agronomic aspects

Red clover is the major forage legume cultivated in Northern Europe (Annicchiarico et al., 2015) where it is mainly grown in mixtures with grasses, of which timothy (*Phleum pratense* L.), meadow fescue (*Festuca pratensis* Huds.), tall fescue (*Festuca arundinacea* Schreb.) and perennial ryegrass (*Lolium perenne* L.) are the most common, depending on the climatic conditions (Boller, 2010). The forage produced by these mixtures is mainly used to produce hay and particularly silage which is fed to animals throughout most of that part of the year when fresh feed from pastures is not available. Inclusion of clover in forage mixtures reduces the need for concentrates (such as soybean), favoring the self-sufficiency for the production of livestock feed (reviewed by Rognli et al., 2021). Red clover forage is characterized by high protein content (Frame et al., 1997), which contributes to high quality in terms of crude protein and unsaturated fatty acids contents of animal end products such as milk and meat (Lee et al., 2009; Taylor and Quesenberry, 1996). Red clover also increases the palatability of the mixtures (Helgadóttir et al., 2018), while the relatively high content of polyphenol oxidase reduces the degradability of protein making protein digestion more efficient and reducing the risk of bloating in ruminants as well as decreasing protein degradation during ensiling (Taylor, 2008).

Red clover is able to fix atmospheric nitrogen thanks to the symbiosis with the soil bacteria *Rhizobium leguminosarum* biovar *trifolii*. Estimates of the nitrogen fixation capability of red clover are up to 400 kg $ha^{-1}year^{-1}$ and this is likely to be positively correlated with dry matter yield (Boller, 2010). Atmospheric nitrogen fixation reduces the requirement of nitrogen fertilizer in the cropping system when red clover, or other legumes, are used in the crop rotation (Jensen et al., 2012; Lüscher et al., 2014; Reckling et al., 2016). From this perspective the cultivation of red clover will help to increase the sustainability of the Nordic agriculture and at the same time to enhance the self-sufficiency in terms of protein production. In fact, the Nordic countries, like the rest of Europe, are strongly dependent on the import of protein based products for food and feed (Visser et al., 2014; Voisin et al., 2014).

5.1.1.3 Red clover in the Norwegian climate

Norway is located in western Scandinavia (57-71°N), the agricultural area covers about 3.3% (1.0 million hectares) of the country surface and 2/3 of this is used for temporary and permanent grasslands (Steinshamn et al., 2016). Grassland

cultivation is practiced all over the country, therefore the species are affected by a wide range of climatic conditions both in term of temperature and precipitation. The growing season, defined as the period of the year with daily mean temperatures above 5 °C, ranges between 100 and 225 days (Steinshamn et al., 2016). Furthermore, another important feature of the growing season in Norway is the high variation in photoperiod during the growing season, typical of regions located at high latitudes. In particular, the combinations of temperature and photoperiod generally result in a relatively high level of daily radiation relative to the mean temperature in the spring and early summer. This situation is particularly evident at the northernmost part of the country, which on the other hand experience low global radiation in the autumn. These dynamics are shown in Figure 1, adapted from Steinshamn et al. (2016), where they give a comprehensive description of the climatic conditions where the Norwegian agriculture find its context.

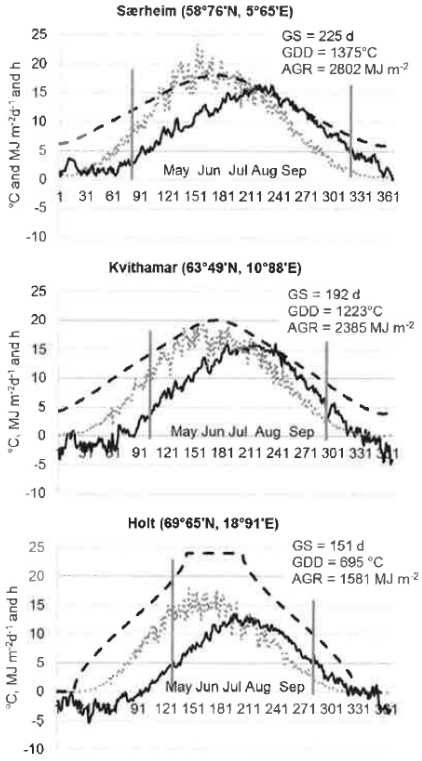


Figure 1. Daily mean temperature (°C, black line), daily global radiation (Mj m^{-2} d^{-1}, grey dotted line) and day length (h, black dashed line) at locations spanning the latitudinal range of most of the Norwegian agriculture (Julian days). The temperature and radiation data are averaged across the years 1999-2011. Days between the vertical lines indicate the growing season (GS) with daily mean temperature > 5°C, GDD is growing degree days (accumulated degree sum above 5°C) and AGR is accumulated global radiation during GS. (Adapted from Steinshamn et al. 2016).

Norway, like the rest of the Nordic region, is being significantly affected by climate change. This is at the same time raising important challenges and opportunities for grassland production. At northern latitudes the observed and projected climate warming is faster than the global average, and the temperature increase is highest during late autumn, winter and spring (Ergon et al., 2018). Annual precipitation is also predicted to increase, especially during the winter (reviewed by Ergon et al., 2018). Warmer temperatures in autumn and spring may lead to a longer growing season in the north, resulting in higher dry matter yield. However, the combination of several situations such as lack of snow cover during the winter, waterlogging and ice cover may challenge plant survival during the winter, resulting in an overall lower production of perennial species (Rapacz et al., 2014). Furthermore, higher

temperatures in autumn will delay acclimation of perennials to a time when less light is available, which may affect the capability of plants to fully cold acclimate (Åshild Ergon, 2017; Dalmannsdottir et al., 2017). On the other hand, unstable winter temperatures, lack of snow cover and earlier springs can cause plants to de-acclimate in periods where frost events may still happen (Rapacz et al., 2014).

Considering the unique combination of annual cycles of photoperiod and temperature and the new challenges imposed by the climate change, the availability of cultivars of grassland species particularly adapted to the Nordic climate is of primary importance. In a recent review, Rognli et al., (2021) suggested that breeding for climate change in northern Europe, requires increasing genetic diversity by introgressing exotic material in species that lack sufficient variation for key traits such as red clover. This is particularly important for traits conferring adaptability, e.g., winter hardiness, freezing and ice-encasement tolerance, timing of growth cessation and utilization of longer growing seasons, and resistance to fungal diseases.

5.1.2 Red clover breeding

5.1.2.1 Breeding goals and methods

Across all the regions where red clover was cultivated, a common practice was to re-sow seed harvested from a restricted area, often within a single farm. This led to the development of locally adapted populations, as a result of semi-conscious selection by the farmers, who exploited the natural selection by producing seeds from the surviving plants in the field. This process gave birth to several landraces, which

were named after the farm or region where they were generated. Farmers were particularly careful to preserve these locally adapted populations avoiding the contamination with seed of different origin. Among the forage species, red clover received early attention by plant breeders, because of its recognized potential to contribute to soil fertility. However, it was only after the second world war that high performing red clover cultivars started to be released and adopted in substitution of locally adapted landraces (Boller 2010). The major breeding goals of red clover are relatively common across its main areas of cultivation, with focus on improving dry matter yield, persistency, disease resistance, seed yield and quality, for example improving the PPO enzyme (polyphenol oxidase) because of its positive effect on N utilization (Abberton and Marshall, 2005; Annicchiarico et al., 2015; Taylor, 2008). However, other goals have been targeted in specific areas depending on the agricultural system where the species is adopted. For example specific morphological characteristics has been prioritized for the improvement of tolerance to grazing or the competition with companion grasses (Abberton and Marshall, 2005).

Red clover is an allogamous self-incompatible species (gametophytic self-incompatibility). Most adopted breeding methods are based on recurrent mass and maternal line selection (Boller et al., 2010). New varieties are created by combining superior progenies of elite parent plants and are normally obtained by open pollination or pair crosses. Progenies are chosen after testing in rows or plots and are then allowed to intermate and form the new synthetic variety (Boller, 2010). Details on how this strategy is adopted in the current Norwegian breeding programs are given in the next section. A synthetic variety can be defined as an heterogenous populations composed of heterozygous genotypes; these varieties can be generated by any number of parents. However, a distinction is made between narrow-based synthetics with fewer than ten parents and broad based synthetics having more than 50 (Annicchiarico et al., 2015). Because of difficulties related to conduct progeny testing of the parent plants, this method may result in a low efficiency of selecting for characters with low heritability. Better selection for these characters may be achieved through the polycross method, however, the length of the time required, and the expenses related to it reduces the adoption of this method (Taylor, 2008). Other breeding techniques such as strain building, backcrossing and hybrid varieties (Riday and Krohn, 2010) have been successfully used for the production of red clover varieties, although they are less common than the traditional methods based on mass selection (Taylor, 2008, Annicchiarico, 2015). Production of interspecific hybrids for the improvement of specific traits where variation was not available within the species were also reported (Abberton and Marshall, 2005).

5.1.2.2 Red clover breeding in Norway

Red clover was introduced in Norway during the 18th and 19th century from the Netherlands and Great Britain (Wexelsen, 1937). However, it was soon clear that the development of adapted material was necessary to improve the use of red clover in Norway, in particular because of the poor winter hardiness of red clover grown from imported seeds. The earliest research on red clover conducted in Norway, in the first half of the 20th century, showed that several Norwegian varieties (local types or landraces) were significantly better than foreign material (Vestad, 1990) in terms of yield and persistency. These local varieties were developed within farms by farmers who set aside some of the seed from their pastures and repeatedly grew the same for several years. The first Norwegian red clover variety was registered in 1953 under the name Molstad, a winter hardy landrace which had been grown already for 80-90 years in the Molstad farm near Brandbu in south-eastern Norway (Vestad, 1990) before its registration in the catalog of cultivated varieties. The work for developing the first Norwegian tetraploid red clover started in 1947, and the first variety (Tripo) was registered in 1964; tetraploid red clover had higher yield and better resistance to clover rot (*Sclerotinia Trifolium*) but at the same time had lower fertility and seed yield than diploid red clover (Vestad, 1990). Since the start of the first Norwegian breeding program for red clover, resistance to diseases and nematodes, winter hardiness, persistency, yield and quality have been the main breeding goals (Vestad, 1990). Current red clover breeding in Norway is conducted at Graminor AS and is based mostly on progeny testing and phenotypic selection of full-sib families, sometimes combined with phenotypic mass selection, followed by creation of synthetic populations. Population improvement through recurrent phenotypic selection is also used. On average the production of a new cultivar requires about 20 years. In a first phase (year 1-8), genotypes from already adopted cultivars are crossed (about 40 initial crosses). In this phase selection is done at family level after progeny testing. After that the best families (3-6) are crossed together to create a synthetic population which is further tested at several locations and years before a new variety is released (year 9-20). The main focus of the breeding program at Graminor AS is on forage yield, winter hardiness, persistence (as tolerance to biotic and abiotic stresses) and seed yield (Amdahl, 2016).

5.1.2.3 Genetic diversity and genetic resources

Until the middle of the 20th century a great diversity of landraces was available worldwide, however, a significant reduction in the cultivation of landraces began around the seventies as a consequence of the development of high performing cultivars by modern breeding programs (Boller, 2010). Fortunately, a large number of these landraces are still preserved at different gene bank institutions, which also conserve wild populations. Wild populations have probably either naturalized from cultivated populations or are the result of hybridization between wild and cultivated populations. Compared to landraces wild populations are less productive in favorable cropping environments but tend to have greater general persistency and may be higher-yielding even in the short term under severe stress conditions, such as drought or severe winter conditions (Annicchiarico and Pagnotta, 2012; Zanotto et al., 2021).

Recently the conservation of landraces and wild populations has been carried out also *in situ,* a method that includes the cultivation of one or more of these landraces "on farm" for some years before seed harvest. In this way these genetic resources can evolve simultaneously as being actively used (Daugstad, 2015).

The genetic resources found within landraces and wild populations are precious as starting points for subsequent breeding programs as well as for the introduction of specific traits of interest in cultivated materials. However, these accessions are often underutilized in plant breeding, both in the Nordic countries and in the rest of Europe (Brozynska et al., 2016; Rognli et al., 2013). The characterization of this germplasm is of primary importance for the improvement of its utilization in breeding programs and for efficient germplasm conservation. Various studies focused on the characterization of red clover genetic resources at the molecular (Greene et al., 2004; Kölliker et al., 2003; Pagnotta et al., 2011; Ulloa et al., 2003) and phenotypic levels (Annicchiarico and Pagnotta, 2012; Dias et al., 2008; Greene et al., 2004; Kouamé and Quesenberry, 1993; Pagnotta et al., 2011; Rosso and Pagano, 2005; Solberg et al., 2017; Tucak et al., 2013; Zanotto et al., 2021). However, a combination of the two is likely the most effective strategy for the characterization of agronomically relevant traits, since the characterization based on selectively neutral marker data tends not to coincide with the phenotypic variation for those traits (Dias et al., 2008; Riday, 2010.; Rognli et al., 2013). The use of molecular markers found within gene sequences or promoters (gene-targeted markers, GTMs) (Poczai et al., 2013) were suggested as being more effective for describing genetic variation for specific traits of interest than selectively neutral markers (van Tienderen et al., 2002).

5.1.2.4 Red clover genomic resources

Red clover, as well as other forage legumes, have lagged behind other crops in terms of available genomic data. The first linkage map for red clover was based on 158 RFLP loci (Isobe et al., 2003). Afterwards, an updated map with 1,399 markers consisting of mainly SSRs and RFLPs was published by Sato et al. in 2005, while an integrated consensus linkage map including 1804 markers distributed over seven linkage groups with a total length of 836 cM was published in 2009 by Isobe et al. (2009). The first draft genome assembly for red clover was published in 2014 by Ištvánek et al. Soon after, an updated draft genome for red clover amounting to 309 Mb, of which 164 Mb are placed on chromosomes, was published by De Vega et al. (2015). Following this work, a coding sequence assembly based on a cultivated red clover population was developed and is available at the Legume Information System (LIS). These achievements were a fundamental step forward and have been an invaluable resource for conducting genomic studies in red clover. These tools were recently used to discover new genes (Ištvánek et al., 2017) as well as for population structure and genetic diversity studies. Recently, a pipeline to identify single nucleotide polymorphisms (SNPs) based GTMs have been successfully developed for red clover (Li et al., 2019). The same study identified three mutations associated with several biological processes, one of this likely involved in molecular mechanisms of self-incompatibility in red clover. In another work, Osterman et al. (2021) used SNPs found within GTMs for a study of genetic diversity and population structure in a panel of Nordic red clover varieties. Similarly, Jones et al. (2020) used SNPs generated by genotyping by sequencing (GBS) to determine the genetic variation and population structure in natural populations of red clover from Europe and Asia, as well as varieties or synthetic populations. The development of modern genomic tools such as GTMs are of primary importance for a species like red clover characterized by high genetic diversity, heterozygosity and genome complexity (Annicchiarico, 2015). Molecular markers may support red clover breeders and managers of gene banks institutions in their decisions on how to maintain or increase genetic diversity in their germplasm collections.

5.1.2.5 Genomic tools for plant breeding

Among the various technologies made available after the last advances in genomics, genotyping by sequencing (GBS) and genome wide association studies

(GWAS) have been widely adopted in plant sciences. GBS has emerged as a popular method to identify genetic variation in species both with and without a reference genome (Elshire et al., 2011). GBS is used for the identification of a high number of markers in the form of single nucleotide polymorphisms (SNPs) which have several applications, from the study of genetic diversity and dissection of complex traits through association studies (reviewed by Jones et al., 2020). Recently, the genotyping of pools of DNA samples from several instead of single individuals, was applied and revealed as a time and cost effective method when dealing with a large number of samples (Byrne et al., 2013). This technique may be particularly useful for outbreeding species with large within population genetic variation, whose characterization at population level is more relevant than at individual level (Ashraf et al., 2014). Furthermore, breeding of these species is mostly based on testing differences between populations (half-sib and full-sib families) which are often the unit of selection.

Genome-wide association study (GWAS) makes it possible to simultaneously screen a very large number of accessions for genetic variation underlying diverse complex traits (Keep et al., 2020; Raggi et al., 2019; Zhao et al., 2011). This technique was successfully used in forage legumes to shed light on the genetic control of quantitative traits and to detect several SNPs within genes related to these traits (Biazzi et al., 2017; Inostroza et al., 2018). However, the power of GWAS to detect true significant associations may be limited by the level of relatedness of accessions (Zhao et al., 2011), leading to the detection of false positives. False positive associations may arise because of two main reasons connected to relatedness of accessions as reviewed by Vilhjálmsson and Nordborg, (2013). The first is the presence of population structure, which is: the study population is a mixture of populations that differ with respect to allele frequencies but also with respect to the traits of interest, controlled by a multitude of different loci. The second is the genetic background, which causes the estimate of the effect of a particular locus to be confounded by the other causal loci in the genome. The use of GWAS models that incorporates population structure and kinship among individuals, through mixed modelling, allows for an effective control of false positives (Liu et al., 2016). Among these models, the Multiple Loci Linear Mixed Model (MLMM), incorporates multiple markers simultaneously as covariates in a stepwise fashion to partially remove the confounding between testing markers and kinship (Segura et al., 2012).

5.1.3 Persistency and freezing tolerance

5.1.3.1 Persistency in the Nordic climate

Improvement of field persistency is a long standing challenge in red clover, and one of the most important breeding goals across the different environments where the species is cultivated (Annicchiarico et al., 2015; Annicchiarico and Pagnotta, 2012; Ergon et al., 2019; Herrmann et al., 2008; Klimenko et al., 2010; Taylor, 2008; Zanotto et al., 2021). Persistency is a complex trait which is affected by many biotic and abiotic factors (Ergon et al., 2019; Herrmann et al., 2008), making its phenotypic evaluation particularly complex. The improvement of persistency of red clover plants will also result in higher dry matter and protein yield of red clover/grass swards in the long term (Marshall et al., 2017). In the Nordic region the main reason for limited persistency is the poor winter survival (Abberton and Marshall, 2005; Helgadóttir et al., 2014), which is also a complicated trait affected by several stresses (Bélanger et al., 2006). Among these, low freezing temperatures are one of the main causes of mortality at locations with continental climate characterized by harsh winters.

The improvement of freezing tolerance (FT) in the north is particularly important in the context of the greater climatic variation linked to the current and projected climate change. The increased climatic variation would indeed imply greater frequency of early or late winter frosts on non-hardened or dehardened plants, as well as lack of snow cover resulting in plants being more vulnerable to low freezing temperatures (Ergon et al., 2018; Rapacz et al., 2014). With this background, the development of red clover cultivars with improved freezing tolerance, as well as wide adaptation to tackle various other stresses occurring during the winter is desired to enhance winter survival and thus persistency at Nordic latitudes (Helgadóttir et al., 2000; Zanotto et al., 2021).

5.1.3.2 Morpho-physiological and biochemical aspects

Since persistency is such a complex trait affected by many biotic and abiotic stresses, it is not easy to identify a plant ideotype of persistent red clover from the morpho-physiological point of view. Several studies underlined the importance of biotic stress resistance as major targets for the improvement of field persistency in red clover (Boller et al., 2010; Taylor, 2008). However, I am not dealing with the biotic stress component of persistency in this thesis. Different studies found that smaller

plant size is associated with better persistency (Annicchiarico and Pagnotta, 2012; Helgadóttir et al., 2000). In the study by Annicchiarico and Pagnotta (2012) it was found that smaller plant size was a feature of wild populations and this was associated to a higher level of persistency than that of land races (Annicchiarico and Pagnotta, 2012). Swiss Mattenklee ecotypes, characterized by early flowering, have been a source of persistence in breeding programs (Kölliker et al., 2003), suggesting that this trait may be important for improved persistency. FT is the main limitation to field persistency in those regions where harsh winter temperatures are the overwhelming cause of mortality such as the Nordic region (Helgadóttir et al., 2014) and Canada (Bertrand et al., 2016). FT is achieved during a period of cold acclimation (CA) at low above freezing temperatures. CA in winter annual and perennial herbaceous plants adapted to temperate climate is associated with huge changes at the transcriptional, proteomic and metabolomic levels (Fürtauer et al., 2019; John et al., 2016; Theocharis et al., 2012). Metabolic changes during CA typically include an increase in the content (on a dry matter basis) of soluble sugars, sugar alcohols, organic acids, amino acids and amino acid derivatives, and other substrates for secondary metabolites. Soluble carbohydrates have multiple functions in FT, such as stabilization of membranes and proteins, lowering of the freezing point within the cell and quenching or scavenging of reactive oxygen species (Pommerrenig et al., 2018). Among the amino acids, proline, the glycine derivative glycine betaine and arginine are accumulated in response to CA in a wide range of species. Proline and glycine betaine function as osmolytes. In addition, proline protect photosystem II under stress and stabilize membranes, while glycine betaine protects against oxidative stress and acts as a signaling molecule (Szabados and Savouré, 2010). Arginine is a precursor in the synthesis of polyamines, which play roles in abiotic and biotic stress resistance by stabilizing membranes, acting as scavengers of reactive oxygen species, and by participating in complex signaling networks that regulate stress responses (Alcázar et al., 2011; Menéndez et al., 2019; Pál et al., 2015). In an analysis of biochemical changes induced by CA in Canadian red clover recurrently selected for improved FT, Bertrand et al. (2020) underlined the importance of the accumulation of soluble sugars and of specific amino acids in plant crowns for the acquisition of superior FT. Similar results were found also in alfalfa (Castonguay et al., 2011), however, in both species the responses differed between genetic backgrounds.

Secondary metabolites in the class of phenolic compounds, including flavonoids, play important roles in the defense towards abiotic and biotic stress and have been shown to act as antioxidants (Agati et al., 2012). Also, flavonoids were shown to contribute to improved freezing tolerance in *Arabidopsis thaliana* (Schulz et al.,

2016). Red clover accumulates several phenolic compounds, in particular isoflavones, which are involved in ozone stress responses in both roots and leaves (Saviranta et al., 2010; Saviranta et al., 2010).

CA also induces major changes in protein abundance; proteins involved in carbohydrate and amino acid metabolism, energy production were upregulated in response to cold acclimation and showed different responses in different genetic backgrounds (Bertrand et al., 2016). In addition, the identification of up- and downregulated proteins involved in the biosynthesis of phenolic compounds after CA (Bertrand et al., 2016), such as the isoflavone-reductase family proteins, suggests that phenolic compounds may play a role in the acquisition of higher freezing tolerance in red clover.

5.1.3.3 Genetic aspects

Various studies have investigated plant persistency in red clover at the genomic level, however, none of these specifically targeted FT as its main component. The first of these studies dates back to 2008 (Herrmann et al., 2008). Considering the difficulty of phenotyping such a complex trait, the authors found that a weighted average of vigor scores assessed for two winter and three growing seasons was the optimal measurement for the trait. They identified a QTL explaining 12.2% of the total phenotypic variation for this trait on LG3. The trait had low heritability, confirming the strong effect of the environment, but it was positively correlated with stem length and with seed yield, although less strongly. Another study (Klimenko et al., 2010) found several QTLs related to winter hardiness, the trait was measured both in controlled environment and in the field, as well as to resistance to diseases affecting red clover during the winter. Most of these QTLs were mapped on LG3 and LG6. Recently Ergon et al. (2019) found various chromosomal loci associated to persistency which were under selection in red clover populations grown for two and a half years comparing to the original populations. These loci were identified by their significant shifts in allele frequency between the originally sown population and the survivor populations. The identified loci under selection were spread across all seven red clover chromosomes, however, one of the loci with the highest allele frequency change was found on chromosome 3. Very little is known about the genetic control of freezing tolerance in red clover however, the trait was demonstrated as being improvable through selection and breeding (Bertrand et al., 2016). In a study on white clover Inostroza et al., (2018) was able to identify a total of 53 loci associated

with FT through GWAS, proving the efficacy of this method for the investigation of the genetics of complex traits. In this context, a comprehensive study of FT in red clover which deals with the morpho-physiological, biochemical and genetic aspects is of primary importance for providing plant breeders with the necessary knowledge and tools for the improvement of this trait in Nordic red clover

5.2 Background and objectives

Red clover (*Trifolium pratense* L.) is the main forage legume cultivated in the Nordic countries. Thanks to its high protein content and nitrogen fixation ability, it plays a central role to enhance protein self-sufficiency and to reduce the use of nitrogen fertilizers, contributing to higher sustainability of the Nordic agriculture. However, cultivation of red clover in the Nordic countries is challenged by poor persistency in the field, mainly because of low winter survival. Among the various causes of plant mortality during the winter, low freezing temperatures are one of the most severe and freezing tolerance (FT) is one of the least investigated traits in red clover. Acquisition of FT in red clover and other perennial species is achieved after a period of cold acclimation (CA) at low above freezing temperatures which induces important changes at the biochemical as well as molecular levels, but none of these mechanisms were studied in red clover of Nordic origin. Furthermore, the achievement of CA may be challenged by the effects of current and future climate change. Finally, very little is known about the genetic control of FT in red clover. Therefore, a better understanding of the CA mechanism, the genetics of FT and an evaluation of genetic resources for the latter and related traits are all priorities for breeding activities aiming at the improvement of FT in Nordic red clover. With this background this PhD project was designed with the aim to investigate the various aspects of FT in red clover at the morpho-physiological, biochemical, and genetic levels, the following objectives have been targeted:

1. To evaluate red clover gene bank accessions of Nordic origin from the collection conserved at NordGen and to compare them with currently cultivated representative cultivars in the Nordic countries. The focus was to better understand the association among various morphological and physiological traits, with particular attention to those associated with FT and winter survival. We also wanted to investigate the relationship between phenotypic traits and climatic variables at the sites of origin of the gene bank

accessions. We proposed that this information may be used to improve collection, conservation, and utilization of red clover genetic resources in the Nordic region.

2. To identify cold-induced biochemical changes associated with freezing tolerance in Nordic red clover. We hypothesize that the content of specific carbohydrates, amino acids and phenolic compounds in plant crowns are significantly affected by short-term CA and that accessions with different levels of freezing tolerance differ in response of these traits to CA.

3. To investigate the genetic control of FT in red clover through Genome Wide Association Studies (GWAS) considering a wide panel of red clover accessions mainly of European origin genotyped at accession level with genotyping by sequencing (GBS). Our hypothesis was that this technique could successfully identify genomic regions associated with FT in red clover and to detect candidate genes which may be involved in biochemical mechanisms related to improved FT in red clover.

5.3 Materials and Methods

A detailed description of the materials and methods used for the completion of the various parts of this PhD project is given in each of the manuscripts. Here I provide an overview on the plant material and analytical approach which were used.

5.3.1 Plant material and experimental locations

For the characterization of red clover accessions of Nordic origin in field and controlled condition experiments (research component 1) a set of 48 accessions of red clover were selected from the NordGen collection, representing the geographical distribution of the species. This included 19 landraces, 2 semi-wild and 27 wild populations. In addition, the commercial cultivars Lea, Linus and Gandalf (Graminor, Norway), SW Ares and SW Yngve (Lantmännen, Sweden) and Saija (Boreal, Finland) were included in the experiment under controlled conditions, while Lea, SW Yngve and Saija were included in the field trials. Among the accessions used in the first research component, the five with the worst and the five with the best freezing tolerance (FS, freezing susceptible and FT, freezing tolerant) were used for the

characterization of the biochemical changes induced by cold acclimation in plant crowns (research component 2). For the genome wide association analyses on freezing tolerance (research component 3) a wide panel of 393 accessions was used. The accessions were wild populations, landraces, cultivars and breeding material, mostly of European origin.

The field experiments to assess the phenotypic variation among gene bank accessions (research component 1) were established at Jokioinen, Finland (60°48'N, 23°29'E, 115 m a.s.l.), Korpa, Iceland (64°09'N, 21°84'W, 35 m a.s.l.), Løken, Norway (61°12'N, 9°06'E, 545 m a.s.l.) and Lännäs, Sweden (63°08'N, 17°43'E, 25 m a.s.l.). Experiments under controlled conditions (research component 1,2 and 3) were conducted in a greenhouse located at Ås, Norway (59°40' N, 10°47' E).

5.3.2 Cold acclimation (CA) and estimates of freezing tolerance (FT)

Greenhouse grown six weeks old seedlings were cold acclimated at 3–4 °C, 12 hr photoperiod and 110 μmol m^2s^{-1} PAR, for two and three weeks in research component 1 and 3 and in research component 2, respectively. For determination of freezing tolerance (LT50; temperature required to kill 50% of the plants), cold acclimated plants were exposed to four different test temperatures (slightly variable for the four research components). Plants of each accession were placed in a programmed freezing chamber initially set at 2°C. The temperature was first lowered from 2°C to −3°C at 1 °C h^{-1} and kept at this level for 12 hr to ensure even freezing, after which the temperature was lowered again by 1 °C h^{-1} down to the test temperature. When the temperature reached the test temperature, it was kept there for one hour before the temperature was raised, again by 1 °C h^{-1}, up to 2 °C. After thawing, the pots were returned to the greenhouse and kept there for three weeks, before survival of each plant was recorded. The experiment was repeated 3 or 4 times, and each accession was represented by 12 or 18 plants per population and test temperature, depending on the research component. The data from the repeated experiments were pooled before LT50 was calculated.

5.3.3 Variation and association among phenotypic traits

The phenotypic traits measured in the field and controlled experiments in research component 1 were: autumn status, growth type, leaf area, stem length,

flowering date, flowering score, yield score, winter survival, spring status, number of leaves, petiole length, days to elongation and freezing tolerance (FT). Trait variation was investigated with boxplots. The association among phenotypic traits was investigated with phenotypic correlation on average data for field and controlled conditions experiments, and genotypic correlations using a mixed model using replicated data for each accession, for the field traits only. Analyses of variance (ANOVA) was performed on field and controlled condition experiments to assess the difference between accession types (cultivars, landraces, and wild types). Variance component analysis on the field traits was conducted using a linear mixed model with the residual maximum likelihood (REML) procedure. The statistical significance of the variance components was further assessed using the likelihood ratio test. Furthermore, the variance components were used to obtain an estimate of entry-mean broad sense heritability (h_b^2) for the field traits. A canonical correlation analysis (CCA) was performed to test the association of each of the phenotypic traits (population means) recorded, with ten variables describing the climate at the collection sites of the gene bank accessions.

5.3.4 Biochemical analyses on red clover crowns before and after cold acclimation

Cold acclimated (as described in section 5.3.2) and non-acclimated plant crowns, a 2-cm transition zone between shoots and roots, of freezing tolerant (FT) and freezing susceptible (FS) red clover accessions were used to prepare freeze-dried and powdered samples for determination of carbohydrate, amino acids and polyphenols content and composition. Plant crown was chosen because it is one of the overwintering tissues in perennial clovers and constitutes the meristematic tissue from which new leaves and branches may grow in the spring. Soluble sugars were separated and quantified on a high-performance liquid chromatography (HPLC) (Waters) analytical system while amino acids were separated and quantified using an ultra-performance liquid chromatography (UPLC) (Waters ACQUITY) analytical system. Both systems were controlled by the Empower II software (Waters).

Polyphenols were separated and determined using liquid chromatography (LC) and high-resolution mass spectrometry (HRMS). The LC system was a Dionex 3000 LC equipped with a Diode Array Detector (DAD) that was placed in-line between the column outlet and the inlet to the ion source of the high resolution Thermo Fisher Q-Exactive mass spectrometer (Thermo Fisher). The contents of carbohydrates and amino acids were subjected to analyses of variance (ANOVA), performed with

treatment (CA vs NA), accession group (FT vs FS) and the interaction between them as factors in the analyses. Also, a Multi Factor Analysis (MFA) was performed to summarize the grouped contribution of the amino acid and carbohydrate contents to the biochemical changes induced by the acclimation treatment in red clover crowns. For the analyses of the polyphenols data, a customized untargeted workflow was developed by adapting an existing workflow in the Compound Discoverer 3.1 software (Thermo Fisher) to process LC-MS raw data and determination of the compounds. Separate analyses were carried out for comparing accession group (FT vs FS) and acclimation treatment (NA vs CA) effects on the content of the compounds detected.

5.3.5 Genome wide association analyses for freezing tolerance

Genotyping of accessions was done at accession level using genotyping by sequencing (GBS, Elshire et al., 2011) of DNA extracted from pools of singles leaves from 200 individuals per accession. Haplotype variants within GBS loci were called and their relative frequencies estimated with the SMAP package (Schaumont et al., 2022). Both data sets were imputed for missing values with the simple mean method. The population structure and kinship among accessions were investigated with a principal component analysis (PCA) and a genomic relationship matrix (GRM), respectively. Linkage disequilibrium (LD) was determined as the squared partial correlation between the allele frequencies of pairs of SNPs (r^2) (Lin et al., 2012; Mangin et al., 2012). SNP- and haplotype-based genome-wide association studies (GWAS) for FT were performed with the Multi Locus Mixed Model (MLMM), which incorporates multiple markers identified as significantly associated with the phenotypic trait as covariates simultaneously (Segura et al. 2012). This method allows for the inclusion of kinship, accounted for by the GRM. Furthermore, regression models between FT and the allele frequencies of the significant markers (SNPs and haplotypes) were performed to estimate the allele effect on the phenotype (slope) and the phenotypic variance explained (R^2) by it. The variance explained (R^2) by all the significant markers simultaneously were also calculated with linear regressions of these markers and FT.

Candidate genes were identified within regions of +/- 0.5 Kbp flanking the significant SNPs and haplotypes identified by the GWAS in the red clover genome Tp2.0 (De Vega et al., 2015) with the gbrowse function available within the Legume Information System (LIS, legumeinfo.org).

5.4 Results and Discussion

5.4.1 Genetic variation for FT and traits associated to winter survival.

The evaluations of freezing tolerance (FT) that we conducted in research components 1 and 3 revealed that a large variation is available for this trait in red clover. Our results point out that the geographic origin of accessions is linked to their level of FT. Among the 393 accessions evaluated in research component 3 those from Scandinavian and eastern European countries tended to have better FT than the other groups, suggesting that the former adapted to the harsher winters typical of their areas of origin and cultivation. This is supported by the results of the field experiments conducted in research component 1, were we found that freezing tolerance was significantly correlated to winter survival at a continental location affected by low winter temperatures. The lack of significant correlations between FT and winter survival at other locations confirms that winter survival is a complex trait affected by high GxE. At these locations other stresses than freezing temperatures may be predominant, and resistance mechanisms towards the various stress factors are likely multiple and, to some extent, genetically independent (Bélanger et al., 2006; Rapacz et al., 2014). FT was negatively correlated with stem and petiole length measured in the field and controlled conditions, respectively. These results suggest that smaller plants may be able to tolerate lower freezing temperatures. The evaluation of Nordic red clover for various morpho-physiological traits done in research component 1 revealed that an erect growth habit and late flowering were associated with high values for several growth and size-related traits. Furthermore, for this material we found that the phenotypic variation was associated with two main axes of climatic variation at the collection site of accessions. The correlation analysis indicates that traits related to bigger plant size, particularly petiole length, increased with increasing summer temperature and decreasing summer precipitation, suggesting that warmer and dryer summers favors the evolution of bigger plant types, and on the other hand colder and wetter summers favors smaller plant types.

For almost all traits we found significant differences between accession types, with wild populations on the one hand and cultivars and landraces on the other hand. Wild populations tended to be smaller with later flowering and lower yield and had significantly better winter survival at a continental location affected by low winter

temperatures compared to landraces. Even if we didn't find significant differences in freezing tolerance between accession type, several wild types and landraces had clearly better freezing tolerance than that of cultivars. All these aspects hint to the fact that the survival of small and compact plants may be favored compared to other plant types at locations affected by severe winter temperatures. The tendency of wild accessions of being more prostrate and smaller in size, but more persistent, is something which was already reported for Italian red clover (Annicchiarico and Pagnotta, 2012) as well as for Nordic white clover populations (Aasmo Finne et al., 2000). We may here see a difference between 'competitors', allocating resources to growth and 'stress tolerators' which allocates resources to storage and maintenance, as described by Grime (1971). The relevance of this theory for our results is supported by the negative correlations that we found between plant size in the autumn and freezing tolerance and winter survival (at two of three locations), which suggest that earlier growth cessation in the autumn favors better freezing tolerance and winter survival. However, these results are apparently not in agreement with the results of the biochemical analyses conducted in research component 2, where we found that freezing susceptible accessions had higher starch content (a reserve carbohydrate) compared to freezing tolerant accessions (more details in the next section). This suggests that the accumulation of reserves in the form of starch in the crowns is not a primary determinant for higher freezing tolerance in itself, but more freezing tolerant accessions are better able to use starch as a building block for the production of other carbohydrates (such as raffinose and pinitol) which have better cryoprotective function than starch. Therefore, the accumulation of reserves may still be a valid explanation for the better freezing tolerance showed by small plants in the autumn. In this perspective, gene bank accessions with such characteristics may indeed be used as a source of useful phenotypic variation for traits related to winter survival and thus persistency in the field. However, our results showed that better FT and persistency is achieved at the expense of other characteristics related to larger size and favorable for higher yield. We found a positive correlation of an erect growth type with bigger plant size and yield, which is in agreement with several studies in red clover of various origin (Annicchiarico and Pagnotta, 2012; Greene et al., 2004; Herrmann et al., 2008; Pagnotta et al., 2011; Rosso and Pagano, 2005; Solberg et al., 2017; Tucak et al., 2013). The negative association between yield and persistency was previously reported in red clover (Annicchiarico and Pagnotta, 2012; Helgadóttir et al., 2000). However, we found that some landraces exceeded cultivars for yield, suggesting that this trait can be indeed improved in cultivated material exploiting the genetic variation conserved in gene bank accessions. It is possible that among these,

genes with positive effect on yield without detrimental effect on persistency are available. Selection for these genes may allow the improvement of the two traits simultaneously.

5.4.2 Biochemical changes induced by cold acclimation (CA) in crowns of red clover with Nordic origin with different level of freezing tolerance (FT)

It is well known that cold acclimation induces major changes at the biochemical, proteomic and metabolomic level in perennial species, however, an investigation of these mechanisms in Nordic red clover is still lacking. Considering the important variation for FT and several other morphological and physiological traits among Nordic gene bank materials found in research component 1, we aimed to understand what characteristics at the biochemical level distinguish accessions with contrasting level of FT before and after cold acclimation.

Cold acclimation provoked marked increases in the content of total soluble sugars which was counterbalanced by the decrease in starch content. Starch is accumulated as carbohydrate reserve in red clover as well as in other legumes, and its depletion in favor of soluble carbohydrates is in line with the findings of Bertrand et al. (2020) in Canadian red clover acclimated in semi-natural conditions. Considering the individual soluble sugars that we analyzed, only glucose was not affected by the acclimation treatment. Pinitol, raffinose fructose and sucrose all increased in content with CA, and sucrose was the one that accounted for the largest amount. The accumulation of sucrose may be due to the reduced catabolism of this sugar in the crowns, as suggested by Bertrand et al (2016), which found a strong reduction of the sucrose-cleaving enzyme sucrose synthase in response to cold acclimation. Considering differences between FT and FS accessions, only sucrose content did not differ between the two groups. Starch, fructose, glucose and the total sum of soluble sugars had higher content in the FS group, while pinitol and raffinose had higher content in the FT group. The largest differences were seen for fructose, which had 80% higher concentration in the FS group, and raffinose, which had a 50% higher concentration in the FT group. Also, these were among the carbohydrates the only two that showed treatment by accession group interaction. Fructose had higher content in the FS group after CA, while the opposite was found for raffinose. In addition, the content of pinitol after CA, was higher in FT than in the FS group (interaction effect was close to significance), suggesting that this carbohydrate, together with raffinose has a role in the freezing tolerance of red clover. Pinitol was

proposed by Bertrand et al. (2020) to have an osmoprotective function in red clover and thus to enhance its freezing tolerance. For raffinose, various studies in alfalfa and red clover (Bertrand et al., 2020, 2017; Castonguay et al., 1995) reported a direct or indirect function of this sugar in the freezing tolerance of these species. Red clover has a limited capacity in accumulating raffinose compared to alfalfa, a reason that could partly explain its lower capacity to reach high levels of freezing tolerance than the latter species.

The amino acid content was also strongly affected by the CA treatment. All the analyzed amino acids but GABA and tyrosine increased with CA. Arginine, asparagine and proline were those with the highest abundance, with proline and asparagine accounted for 40% of all the free amino acids after CA. These three amino acids were all previously suggested to be involved in the cold stress response in red clover and alfalfa (Bertrand et al., 2020; Castonguay et al., 2011). Arginine is an important storage compound in plants, and its abundance in the taproots of alfalfa was proposed as playing a role as N reserves favoring overwintering and spring regrowth (Dhont et al., 2003). Furthermore, arginine is a precursor of polyamines which can be catabolized into GABA, which are involved in stress response in various legumes species (Li et al., 2018; Mendez et al., 2019; Yong et al., 2017; Zhou et al., 2021). However, we found a decreased content of GABA after cold acclimation, which is not consistent with a previous study in Canadian red clover (Bertrand et al., 2020), that found maximum accumulation of GABA after cold acclimation. The different acclimation conditions and the different genetic background of the materials that were used are the most probable causes of the differences that we found here compared to that study. Arginine was also among the amino acids that that showed significantly different accumulation between FT and FS accessions, with higher content in the latter. The roles of arginine in freezing tolerance of red clover are already discussed above. Serine, valine, tyrosine, alanine and phenylalanine are other amino acids which had higher content in the FT than in the FS group and that could therefore be involved in their difference in freezing tolerance.

Cold acclimation also induced the accumulation of six phenolic compounds whose identity could be assigned (Daidzein, Biochanin A isomer, L-phenylalanine, L-tryptophan, Genistein, N-(phenylacetyl) aspartic acid) while an equal number of identified compounds had higher content in non-acclimated crowns (Formononetin derivative, Dideoxyclovamide, Tetrahydro-carboline-3-carboxylic acid, Irilone derivative, Deoxyclovamide (caffeoyl tyrosine), Irilone isomer derivative). However, we found a general trend in the accumulation of a higher number of compounds (detected by the analyses but whose identity could not be assigned) after CA,

suggesting that the treatment indeed shifts the phenolic compound metabolism towards the accumulation of a larger number of compounds, in agreement with the findings by Schulz et al., (2015) in A. thaliana. Also, four phenolic compounds had different intensities (as detected my MS analyses) between FT and FS accessions. However, only one of them (pinocembrin-G derivative) was found with higher intensity in FT while the other three were more intense in FS accessions (Dihydrokaempferol, Biochanin A-G-M-M, Phloretin isomer).

The majority of the phenolic compounds that we identified in this study belongs to the isoflavones, a class of phenolic compounds particularly abundant in legumes (Mazur et al., 1998). There is a lack of knowledge on the role of specific phenolic compounds in abiotic stress response. However, different studies in red clover and *A. thaliana* (Saviranta et al., 2010 a; Saviranta et al., 2010 b, Schulz et al., 2016), suggested that a combined contribution of a set of different phenolic compounds is more likely than the specific action of a particular compound in the response to ozone and freezing stress respectively. Overall, our results support the initial hypothesis that phenolic compounds are involved in the response to cold stress in red clover, but further research is needed to elucidate what is the mode of action of these metabolites and if this differs among different compounds.

5.4.3 Genomic regions associated with FT revealed by SNPs and haplotypes-based genome wide association analyses (GWAS)

To complement the knowledge at the morphological and biochemical levels acquired in research component 1 and 2, we conducted a genome wide association study (GWAS) of FT, to investigate the genetic control of the trait. The variation in FT was high, and partially associated with the origin of the material (see section 5.4.1). At the genomic level accessions appeared to be quite highly structured (revealed by principal component analysis, PCA), which was also related with the geographic origin of the plant material. Also, we found a very rapid decay in linkage disequilibrium (LD), which is typical in outcrossing species, and is in line with previous studies in red clover (Jones et al., 2020) and perennial ryegrass (Keep et al., 2020). The LD was close to the background level already at 1 Kbp distance between pairs of markers, while the average distance between two independent GBS loci was 35 kbp. Therefore, the marker density that we found was likely not sufficient for

having a major part of the QTLs (quantitative trait loci) associated with FT to be in linkage with at least one of these GBS-loci.

Our GWAS analyses found a total of thirteen loci significantly associated with freezing tolerance, while one more locus was just below the threshold of significance. Out of these fourteen loci, nine were mapped on chromosomes while five more were located on unmapped scaffold regions. This is to our knowledge the first study that used GBS-generated allele frequencies data from population pools to conduct SNP- and short read haplotype-based GWAS in red clover. Our results confirms that GBS on pools of individuals is an effective and economical technique for the generation of genotypic data in highly diverse outbreeding species, to be used in genomic studies. The fact that the two analyses (SNP and short read haplotype-based GWAS) found different independent loci and only two of these were shared across the analyses, confirm that haplotypes are able to reveal associations that are not detectable by single SNPs and vice versa (Bekele et al., 2018). Thus, the different marker types can complement each other, revealing a larger total number of significant associations (Hamblin and Jannink, 2011; Lorenz et al., 2010). The multi-locus-mixed-model (MLMM) approach that we used for both SNPs and haplotype-based GWAS analyses allowed us to account for the kinship among the accessions. This was particularly important to reduce the confounding effect due to the structure present in this material. The inclusion of kinship effect in the GWAS model ensures that the associations that are identified are truly significant, even if it is likely to induce an underestimation of the number of significant loci and the total phenotypic variance explained (Keep et al., 2020). The magnitude of the kinship effect, in terms of reduction of the phenotypic variance explained, was variable among the different loci that we detected. In addition to the inclusion of kinship correction in the GWAS models, the exclusion of SNPs and haplotypes with low frequency (MAF <0.5) is another probable cause for the relatively low number of significant loci that we identified, since rare alleles significantly associated with FT may be excluded with this stringent filtering. On the other hand, all these precautions ensure that the associations that we found are indeed true with low risk of being false positives.

Interestingly, the loci which were mostly affected by the kinship effect where also those whose allele frequency showed the strongest shift in at least one geographic group of accessions compared to the others. In other words, some of the alleles significantly associated with FT were frequent only among a particular group of accessions, something to keep in mind in local breeding programs aiming at the improvement of FT, for which it may be useful to introduce material from other countries/areas as source of useful alleles.

Overall, we found a total of 10 independent GBS-loci significantly associated with variation if FT that were mapped within gene sequences of the red clover genome, and two more loci were found very close to a gene (< 0.5kb). Markers very close or within the sequences of genes associated to phenotypic traits are very useful for the characterization of germplasm for these traits as well as for breeding purposes (Li et al., 2019; Osterman et al., 2021). Such markers for FT could potentially be developed from the loci that we identified here, after the necessary steps for the verification of the robustness of these associations and their validation across a diversified panel of accessions. In future, this kind of markers could be used to complement the phenotypic characterization with genomic information in studies such as those conducted in research component 1.

Among the gene sequences tagged by significant SNPs and HTPs some are coding for proteins that may be involved in biochemical mechanisms related to freezing tolerance. In particular, one gene coded for a membrane transporter D1 n protein in red clover, however, its corresponding sequence in the A. *thaliana* and M. *truncatula* genomes codes for an inositol transporter 1 (INT1) protein. A protein which facilitates myo-inositol import from the vacuole into the cytoplasm (Strobl et al., 2011). Myo-inositol is a cyclic sugar alcohol that is accumulated in different plants in response to abiotic stress (Sengupta et al., 2015), and which has a role in transitional metabolism and various signaling pathways in plants (Zhou et al., 2021). Among the other genes that we identified, two coded for proteins involved in oxidative stress response, one for a transducing/WD40 repeat-like superfamily protein (a vast class of proteins involved in several functions), two coded for proteins involved in DNA processing and two more for proteins with a hydrolytic function. Interestingly, the best hit of the two gene sequences in the *M*. *truncatula* genome is coding for an enzyme, caffeoyl shikimate esterase, involved in the shikimate biosynthetic pathway. In research component 2 we found that secondary metabolites of the flavonoids class were involved in the response to cold stress in red clover and the identification of a gene involved in the biosynthetic pathway of secondary metabolites cautiously suggest a possible link between our findings at the biochemical and metabolic levels.

5.5 Identified gaps for future study

5.5.1 Integration of phenotypic genotypic and environmental data for characterization of genetic resources

This study showed that the characterization of gene bank accessions is a precious activity to evaluate the variation for phenotypic traits and the associations among them. These activities are particularly useful for the identification of unexploited genetic resources that can be used to improve cultivated materials for traits of interest. In future it would be useful to create a framework which allows the integration of phenotypic, environmental and genomic data simultaneously in order to obtain a better characterization of accessions. This, as suggested by Blanco-Pastor et al., (2020) may also help us to understand the adaptive genomic diversity and phenotypic responses to environmental selection pressures. Also, in the design of similar studies in future, the choice of entries should be done in order to include accessions whose origin covers as much as possible the climatic variation of the target area including an evenly distributed number of accession types. This will help to reduce biases due to unbalanced distribution of accession type in respect to the total climatic variation of the area and further to identify a larger number of phenotype-climate gradients. This last issue was indeed a limitation in our study.

5.5.2 Deepen the understanding of CA and FT through gene expression analyses and investigation the effect of climate change on CA

This is to our knowledge the first study which analyzed the effect of CA in red clover of Nordic origin at the biochemical level. Through the investigation of these changes in accessions with a contrasting level of FT, we could also indicate which of the analyzed metabolites are associated to the capacity of red clover to tolerate low freezing temperatures, and if this is interacting with the CA process.

In future, a deeper study of the biochemical pathways of these metabolites (raffinose, pinitol, arginine, serine, alanine, valine, phenylalanine in particular), in association with gene expression analyses and additional/alternative physiological measurements for the assessment of freezing stress response, should be carried out to further dissect the mechanisms responsible for CA and the acquisition of FT in red

clover. Also, a study of CA under different conditions in terms of light and temperature would be useful to understand how these processes may be affected by the climate change. This knowledge has relevance for breeding of red clover with improved FT adapted to the future climatic conditions affecting the Nordic agriculture.

5.5.3 Development of molecular markers for FT

The results of the GWAS provided a significant advance in our knowledge of the genetic base of FT in red clover. In future we suggest to further investigate the genomic regions identified by our analysis to confirm/verify the associations that we identified. It may be worth considering a different method for phenotyping FT, for example using LD50 (lethal duration time for 50% kill) instead of LT50 (Waalen et al., 2011). Also, a different CA protocol which allows plant to reach full CA may be a useful strategy for improving the phenotyping of this trait. More accurate phenotyping and fine mapping of genomic regions identified by our GWAS in a subset of selected accessions with contrasting level of FT may lead to the identification of major QTLs for this trait.

Other strategies such as the use of multiparent advanced generation intercross (MAGIC) populations (Cavanagh et al., 2008) are worth considering for precise QTLs mapping. Another strategy could be to use the genomic information that we obtained from our GWAS analysis to develop a genomic selection (GS) model for FT in red clover. GS uses a high number of genetic markers distributed all over the genome to estimate the breeding value and/or to predict the phenotype of genotypes without the need for phenotyping thanks to the prior phenotypic and genotypic evaluation of a training population (Meuwissen et al., 2001). In this perspective, the markers that we found in the GWAS could be included in a GS model for red clover targeting the improvement of FT and several other traits of interest.

Overall, the development of strong and reliable molecular tools is a priority to speed up the improvement of FT and other complex traits through selection and breeding. Among all these proposed approaches, molecular assisted selection strategies such as GS is likely the most promising for the improvement of FT simultaneously to other traits of interest through breeding.

5.6 Conclusions

- Considerable variation was identified among gene bank accessions for FT and other morphological and physiological traits useful for the improvement of cultivated material.
- Traits, such as petiole length, which are fast and easy to measure in controlled conditions and significantly correlated with many other traits can be used as proxy for traits that are longer and more expensive to measure.
- The high GxE that we found, especially for winter survival, suggests that the breeding of red clover cultivars with improved persistency to the Nordic region requires testing at several locations over several years and that breeding strategies targeting broad instead of specific adaptation are likely more successful for this region.
- Phenotypic variation was associated with two main axes of climatic variation (1, variation in summer temperature and precipitations and 2, variation in annual and winter temperatures) at the collection site of accessions. This information can be useful for the identification of plant materials with certain characteristics based on climatic information for breeding and conservation purposes.
- The major changes in accumulation of soluble sugars and free amino acids in response to CA that we found confirm the importance of these metabolites in the CA process in Nordic red clover.
- The identification of compounds such as raffinose, pinitol, arginine and others that differ between FT and FS accessions, suggests that these compounds may have a role in the acquisition of superior FT in Nordic red clover.
- 31 major phenolic compounds mostly in the class of isoflavones were identified in crowns of red clover of Nordic origin.
- The content of specific phenolic compounds changes in response to CA, indicating their involvement in the metabolic changes happening in red clover after exposure to cold temperatures.
- We found a very rapid LD decay red clover accessions used for the GWAS, confirming the high level of genetic diversity present within and between red clover accessions of different origin.
- The novel approach that we adopted conducting GWAS with haplotype data in addition to SNP data was successful for the identification of loci associated with FT in the red clover genome.

- We identified candidate genes which may help to further understand the biochemical and physiological mechanisms related to CA and FT in red clover.

References

Abberton, M.T., Marshall, A.H., 2005. Progress in breeding perennial clovers for temperate agriculture. The Journal of Agricultural Science 143, 117–135. https://doi.org/10.1017/S0021859605005101

Agati, G., Azzarello, E., Pollastri, S., Tattini, M., 2012. Flavonoids as antioxidants in plants: Location and functional significance. Plant Science 196, 67–76. https://doi.org/10.1016/j.plantsci.2012.07.014

Alcázar, R., Cuevas, J.C., Planas, J., Zarza, X., Bortolotti, C., Carrasco, P., Salinas, J., Tiburcio, A.F., Altabella, T., 2011. Integration of polyamines in the cold acclimation response. Plant Science, Plant and Microbe Adaptation to Cold 180, 31–38. https://doi.org/10.1016/j.plantsci.2010.07.022

Amdahl, H., 2016. Improving the seed yield potential of tetraploid red clover (Trifolium pratense L.). Norwegian University of Life Sciences, Ås.

Annicchiarico, P., Barrett, B., Brummer, E.C., Julier, B., Marshall, A.H., 2015. Achievements and Challenges in Improving Temperate Perennial Forage Legumes. Critical Reviews in Plant Sciences 34, 327–380. https://doi.org/10.1080/07352689.2014.898462

Annicchiarico, P., Pagnotta, M.A., 2012. Agronomic value and adaptation across climatically contrasting environments of Italian red clover landraces and natural populations. Grass and Forage Science 67, 597–605. https://doi.org/10.1111/j.1365-2494.2012.00887.x

Ashraf, B.H., Jensen, J., Asp, T., Janss, L.L., 2014. Association studies using family pools of outcrossing crops based on allele-frequency estimates from DNA sequencing. Theoretical and Applied Genetics 127, 1331–1341. https://doi.org/10.1007/s00122-014-2300-4

Bekele, W.A., Wight, C.P., Chao, S., Howarth, C.J., Tinker, N.A., 2018. Haplotype-based genotyping-by-sequencing in oat genome research. Plant Biotechnology Journal 16, 1452–1463. https://doi.org/10.1111/pbi.12888

Bélanger, G., Castonguay, Y., Bertrand, A., Dhont, C., Rochette, P., Couture, L., Drapeau, R., Mongrain, D., Chalifour, F.-P., Michaud, R., 2006. Winter damage to perennial forage crops in eastern Canada: Causes, mitigation, and prediction. Canadian Journal of Plant Science 86, 33–47. https://doi.org/10.4141/P04-171

Bertrand, A., Bipfubusa, M., Castonguay, Y., Rocher, S., Szopinska-Morawska, A., Papadopoulos, Y., Renaut, J., 2016. A proteome analysis of freezing tolerance in red clover (Trifolium pratense L.). BMC Plant Biology 16, 65. https://doi.org/10.1186/s12870-016-0751-2

Bertrand, A., Bipfubusa, M., Claessens, A., Rocher, S., Castonguay, Y., 2017. Effect of photoperiod prior to cold acclimation on freezing tolerance and carbohydrate metabolism in alfalfa (Medicago sativa L.). Plant Science 264, 122–128. https://doi.org/10.1016/j.plantsci.2017.09.003

Bertrand, A., Rocher, S., Claessens, A., Bipfubusa, M., Papadopoulos, Y., Castonguay, Y., 2020. Biochemical and molecular responses during overwintering of red clover populations recurrently selected for improved freezing tolerance. Plant Science 292, 110388. https://doi.org/10.1016/j.plantsci.2019.110388

Biazzi, E., Nazzicari, N., Pecetti, L., Brummer, E.C., Palmonari, A., Tava, A., Annicchiarico, P., 2017. Genome-Wide Association Mapping and Genomic Selection for Alfalfa (Medicago sativa) Forage Quality Traits. PLOS ONE 12, e0169234. https://doi.org/10.1371/journal.pone.0169234

Blanco-Pastor, J.L., Barre, P., Keep, T., Ledauphin, T., Escobar-Gutiérrez, A., Roschanski, A.M., Willner, E., Dehmer, K.J., Hegarty, M., Muylle, H., Veeckman, E., Vandepoele, K., Ruttink, T., Roldán-Ruiz, I., Manel, S., Sampoux, J.-P., n.d. Canonical correlations reveal adaptive loci and phenotypic responses to climate in perennial ryegrass. Molecular Ecology Resources n/a. https://doi.org/10.1111/1755-0998.13289

Boller, B., Schubiger, F.X., Kölliker, R., 2010. Red Clover, in: Boller, B., Posselt, U.K., Veronesi, F. (Eds.), Fodder Crops and Amenity Grasses, Handbook of Plant Breeding. Springer, New York, NY, pp. 439–455. https://doi.org/10.1007/978-1-4419-0760-8_18

Brozynska, M., Furtado, A., Henry, R.J., 2016. Genomics of crop wild relatives: expanding the gene pool for crop improvement. Plant Biotechnology Journal 14, 1070–1085. https://doi.org/10.1111/pbi.12454

Byrne, S., Czaban, A., Studer, B., Panitz, F., Bendixen, C., Asp, T., 2013. Genome Wide Allele Frequency Fingerprints (GWAFFs) of Populations via Genotyping by Sequencing. PLOS ONE 8, e57438. https://doi.org/10.1371/journal.pone.0057438

Castonguay, Y., Bertrand, A., Michaud, R., Laberge, S., 2011. Cold-Induced Biochemical and Molecular Changes in Alfalfa Populations Selectively Improved for Freezing Tolerance. Crop Science 51, 2132–2144. https://doi.org/10.2135/cropsci2011.02.0060

Castonguay, Y., Nadeau, P., Lechasseur, P., Chouinard, L., 1995. Differential Accumulation of Carbohydrates in Alfalfa Cultivars of Contrasting Winterhardiness. Crop Science 35, cropsci1995.0011183X003500020038x. https://doi.org/10.2135/cropsci1995.0011183X003500020038x

Cavanagh, C., Morell, M., Mackay, I., Powell, W., 2008. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. Current Opinion in Plant Biology, Genome studies and Molecular Genetics, edited by Juliette de Meaux and Maarten Koornneef / Plant Biotechnology, edited by Andy Greenland and Jan Leach 11, 215–221. https://doi.org/10.1016/j.pbi.2008.01.002

Dalmannsdottir, S., Jørgensen, M., Rapacz, M., Østrem, L., Larsen, A., Rødven, R., Rognli, O.A., 2017. Cold acclimation in warmer extended autumns impairs freezing tolerance of perennial ryegrass (Lolium perenne) and timothy (Phleum pratense). Physiologia Plantarum 160, 266–281. https://doi.org/10.1111/ppl.12548

Daugstad, K., Løken, B., 2015. Bevaring ved bruk – nye lokalsortar i engvekstene timotei, engsvingel og raudkløver 10, 28.

De Vega, J. J., S. Ayling, M. Hegarty,D. Kudrna, J. L. Goicoechea, Å. Ergon, O. A. Rognli, C. Jones, M: Swain, R. Geurts, C. Lang, K. F. X. Mayer, S. Rössner, S. Yates, K. J. Webb, I. S. Donnison, G. E. D. Oldroyd, R. A. Wing, M. Caccamo1, W. Powell, M. T. Abberton and L. Skøt. 2015. Red clover (Trifolium pratense L.) draft genome provides a platform for trait improvement. Sci Rep, 5.

Dhont, C., Castonguay, Y., Nadeau, P., Bélanger, G., Chalifour, F.-P., 2003. Alfalfa Root Nitrogen Reserves and Regrowth Potential in Response to Fall Harvests. Crop Science 43, 181–194. https://doi.org/10.2135/cropsci2003.1810

Dias, P.M.B., Julier, B., Sampoux, J.-P., Barre, P., Dall'Agnol, M., 2008. Genetic diversity in red clover (Trifolium pratense L.) revealed by morphological and microsatellite (SSR) markers. Euphytica 160, 189–205. https://doi.org/10.1007/s10681-007-9534-z

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One 6. https://doi.org/10.1371/journal.pone.0019379

Ergon, Å., 2017. Optimal Regulation of the Balance between Productivity and Overwintering of Perennial Grasses in a Warmer Climate. Journal of Agronomy and Crop Science.

Ergon, Å., Seddaiu, G., Korhonen, P., Virkajärvi, P., Bellocchi, G., Jørgensen, M., Østrem, L., Reheul, D., Volaire, F., 2018. How can forage production in Nordic and Mediterranean Europe adapt to the challenges and opportunities arising from climate change? European Journal of Agronomy 92, 97–106. https://doi.org/10.1016/j.eja.2017.09.016

Ergon, Å., Skøt, L., Sæther, V.E., Rognli, O.A., 2019. Allele Frequency Changes Provide Evidence for Selection and Identification of Candidate Loci for Survival in Red Clover (Trifolium pratense L.). Front. Plant Sci. 10. https://doi.org/10.3389/fpls.2019.00718

Frame, J., Charlton, J.F.L., Laidlaw, A.S., 1997. Temperate forage Legumes.

Fürtauer, L., Weiszmann, J., Weckwerth, W., Nägele, T., 2019. Dynamics of Plant Metabolism during Cold Acclimation. International Journal of Molecular Sciences 20, 5411. https://doi.org/10.3390/ijms20215411

Greene, S.L., Gritsenko, M., Vandemark, G., 2004. Relating Morphologic and RAPD Marker Varlation to Collection Site Environment in wild Populations of Red Clover (Trifolium Pratense L.). Genetic Resources and Crop Evolution 51, 643–653. https://doi.org/10.1023/B:GRES.0000024655.48989.ab

Grime, J. P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. The American Naturalist, 111, 1169–1194. https://doi. org/10.1086/283244

Hamblin, M.T., Jannink, J.-L., 2011. Factors Affecting the Power of Haplotype Markers in Association Studies. The Plant Genome 4. https://doi.org/10.3835/plantgenome2011.03.0008

Helgadóttir, Á., Larsen, A., Marum, P., Fritsen, H., Lindvall, E., Miettinen, E., 2000. Prebreeding of Red Clover (Trifolium pratense L.) for Northern Areas. Acta Agriculturae Scandinavica, Section B - Soil & Plant Science 50, 187–190. https://doi.org/10.1080/090647100750374368

Helgadóttir Á., Frankow-Lindberg B. E., Seppänen M. M., Søegaard K. & Østrem L. (2014). European grasslands overview: Nordic region. In Hopkins et al (Eds.), EGF at 50: The future of European grasslands. Grassland Science in Europe. 19, 15–28.

Helgadóttir, Á., Suter, M., Gylfadóttir, T.Ó., Kristjánsdóttir, T.A., Lüscher, A., 2018. Grass–legume mixtures sustain strong yield advantage over monocultures under cool maritime growing conditions over a period of 5 years. Ann Bot 122, 337–348. https://doi.org/10.1093/aob/mcy074

Herrmann, D., Boller, B., Studer, B., Widmer, F., Kölliker, R., 2008. Improving Persistence in Red Clover: Insights from QTL Analysis and Comparative Phenotypic Evaluation. Crop Science 48, 269–277. https://doi.org/10.2135/cropsci2007.03.0143

Inostroza, L., Bhakta, M., Acuña, H., Vásquez, C., Ibáñez, J., Tapia, G., Mei, W., Kirst, M., Resende, M., Munoz, P., 2018. Understanding the Complexity of Cold Tolerance in White Clover using Temperature Gradient Locations and a GWAS Approach. The Plant Genome 11, 170096. https://doi.org/10.3835/plantgenome2017.11.0096

Isobe, S., Klimenko, I., Ivashuta, S., Gau, M., Kozlov, N.N., 2003. First RFLP linkage map of red clover (Trifolium pratense L.) based on cDNA probes and its transferability to other red clover germplasm. Theor Appl Genet 108, 105–112. https://doi.org/10.1007/s00122-003-1412-z

Isobe, S., Kölliker, R., Boller, B., Riday, H., 2013. Red Clover, in: Genetics, Genomics and Breeding of Forage Crops. CRC Press.

Isobe, S., Kölliker, R., Hisano, H., Sasamoto, S., Wada, T., Klimenko, I., Okumura, K., Tabata, S., 2009. Construction of a consensus linkage map for red clover (Trifolium pratenseL.). BMC Plant Biol 9, 57. https://doi.org/10.1186/1471-2229-9-57

Ištvánek, J., Dluhošová, J., Dluhoš, P., Pátková, L., Nedělník, J., Řepková, J., 2017. Gene Classification and Mining of Molecular Markers Useful in Red Clover (Trifolium pratense) Breeding. Frontiers in Plant Science 8, 367. https://doi.org/10.3389/fpls.2017.00367

Ištvánek, J., Jaroš, M., Křenek, A., Řepková, J., 2014. Genome assembly and annotation for red clover (Trifolium pratense; Fabaceae). American Journal of Botany 101, 327–337. https://doi.org/10.3732/ajb.1300340

Jensen, E.S., Peoples, M.B., Boddey, R.M., Gresshoff, P.M., Hauggaard-Nielsen, H., J.R. Alves, B., Morrison, M.J., 2012. Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. Agron. Sustain. Dev. 32, 329–364. https://doi.org/10.1007/s13593-011-0056-7

John, R., Anjum, N.A., Sopory, S.K., Akram, N.A., Ashraf, M., 2016. Some key physiological and molecular processes of cold acclimation. Biol Plant 60, 603–618. https://doi.org/10.1007/s10535-016-0648-9

Jones, C., De Vega, J., Lloyd, D., Hegarty, M., Ayling, S., Powell, W., Skøt, L., 2020. Population structure and genetic diversity in red clover (Trifolium pratense L.) germplasm. Scientific Reports 10, 8364. https://doi.org/10.1038/s41598-020-64989-z

Keep, T., Sampoux, J.-P., Blanco-Pastor, J.L., Dehmer, K.J., Hegarty, M.J., Ledauphin, T., Litrico, I., Muylle, H., Roldán-Ruiz, I., Roschanski, A.M., Ruttink, T., Surault, F., Willner, E., Barre, P., 2020. High-Throughput Genome-Wide Genotyping To Optimize the Use of Natural Genetic Resources in the Grassland Species Perennial Ryegrass (Lolium perenne L.). G3: Genes, Genomes, Genetics 10, 3347–3364. https://doi.org/10.1534/g3.120.401491

Kjærgaard, T., 2003. A Plant that Changed the World: The rise and fall of clover 1000-2000. Landscape Research 28, 41–49. https://doi.org/10.1080/01426390306531

Klimenko, I., Razgulayeva, N., Gau, M., Okumura, K., Nakaya, A., Tabata, S., Kozlov, N.N., Isobe, S., 2010. Mapping candidate QTLs related to plant persistency in red clover. Theor Appl Genet 120, 1253–1263. https://doi.org/10.1007/s00122-009-1253-5

Kölliker, R., Herrmann, D., Boller, B., Widmer, F., 2003. Swiss Mattenklee landraces, a distinct and diverse genetic resource of red clover (Trifolium pratense L.). Theoretical and Applied Genetics 107, 306–315. https://doi.org/10.1007/s00122-003-1248-6

Kouamé, C.N., Quesenberry, K.H., 1993. Cluster analysis of a world collection of red clover germplasm. Genet Resour Crop Evol 40, 39–47. https://doi.org/10.1007/BF00053463

Lee, M.R.F., Evans, P.R., Nute, G.R., Richardson, R.I., Scollan, N.D., 2009. A comparison between red clover silage and grass silage feeding on fatty acid composition, meat stability and sensory quality of the M. Longissimus muscle of dairy cull cows. Meat Science 81, 738–744. https://doi.org/10.1016/j.meatsci.2008.11.016

Li, Z., Zhang, Y., Peng, D., Peng, Y., Zhang, X., Ma, X., Huang, L., Yan, Y., 2018. The inhibition of polyamine biosynthesis weakens the drought tolerance in white clover (Trifolium repens) associated with the alteration of extensive proteins. Protoplasma 255, 803–817. https://doi.org/10.1007/s00709-017-1186-9

Li, W., Riday, H., Riehle, C., Edwards, A., Dinkins, R., 2019. Identification of Single Nucleotide Polymorphism in Red Clover (Trifolium pratense L.) Using Targeted Genomic Amplicon Sequencing and RNA-seq. Front Plant Sci 10, 1257. https://doi.org/10.3389/fpls.2019.01257

Lin, C.-Y., Xing, G., Xing, C., 2012. Measuring linkage disequilibrium by the partial correlation coefficient. Heredity 109, 401–402. https://doi.org/10.1038/hdy.2012.54

Liu, X., Huang, M., Fan, B., Buckler, E.S., Zhang, Z., 2016. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. PLOS Genetics 12, e1005767. https://doi.org/10.1371/journal.pgen.1005767

Lorenz, A.J., Hamblin, M.T., Jannink, J.-L., 2010. Performance of Single Nucleotide Polymorphisms versus Haplotypes for Genome-Wide Association Analysis in Barley. PLOS ONE 5, e14079. https://doi.org/10.1371/journal.pone.0014079

Lüscher, A., Mueller-Harvey, I., Soussana, J.F., Rees, R.M., Peyraud, J.L., 2014. Potential of legume-based grassland–livestock systems in Europe: a review. Grass and Forage Science 69, 206–228. https://doi.org/10.1111/gfs.12124

Mangin, B., Siberchicot, A., Nicolas, S., Doligez, A., This, P., Cierco-Ayrolles, C., 2012. Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. Heredity 108, 285–291. https://doi.org/10.1038/hdy.2011.73

Marshall, A.H., Collins, R.P., Vale, J., Lowe, M., 2017. Improved persistence of red clover (Trifolium pratense L.) increases the protein supplied by red clover/grass swards grown over four harvest years. European Journal of Agronomy 89, 38–45. https://doi.org/10.1016/j.eja.2017.06.006

Mazur, W.M., Duke, J.A., Wähälä, K., Rasku, S., Adlercreutz, H., 1998. Isoflavonoids and Lignans in Legumes: Nutritional and Health Aspects in Humans 11The method development and synthesis of the standards and deuterium-labelled compounds was supported by National Institutes of Health Grants No. 1 R01 CA56289-01 and No. 2 R01 CA56289-04, and analytical work by the EU research contract FAIR-CT95-0894. The Journal of Nutritional Biochemistry 9, 193–200. https://doi.org/10.1016/S0955-2863(97)00184-8

Menéndez, A.B., Calzadilla, P.I., Sansberro, P.A., Espasandin, F.D., Gazquez, A., Bordenave, C.D., Maiale, S.J., Rodríguez, A.A., Maguire, V.G., Campestre, M.P., Garriz, A., Rossi, F.R., Romero, F.M., Solmi, L., Salloum, M.S., Monteoliva, M.I., Debat, J.H., Ruiz, O.A., 2019. Polyamines and Legumes: Joint Stories of Stress, Nitrogen Fixation and Environment. Frontiers in Plant Science 10. https://doi.org/10.3389/fpls.2019.01415

Meuwissen, T.H.E., Hayes, B.J., Goddard, M.E., 2001. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. Genetics 157, 1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Osterman, J., Hammenhag, C., Ortiz, R., Geleta, M., 2021. Insights Into the Genetic Diversity of Nordic Red Clover (Trifolium pratense) Revealed by SeqSNP-Based Genic Markers. Front Plant Sci 12, 748750. https://doi.org/10.3389/fpls.2021.748750

Pagnotta, M.A., Annicchiarico, P., Farina, A., Proietti, S., 2011. Characterizing the molecular and morphophysiological diversity of Italian red clover. Euphytica 179, 393–404. https://doi.org/10.1007/s10681-010-0333-6

Pál, M., Szalai, G., Janda, T., 2015. Speculation: Polyamines are important in abiotic stress signaling. Plant Science 237, 16–23. https://doi.org/10.1016/j.plantsci.2015.05.003

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J.P., Hyvönen, J., 2013. Advances in plant gene-targeted and functional markers: a review. Plant Methods 9, 6. https://doi.org/10.1186/1746-4811-9-6

Pommerrenig, B., Ludewig, F., Cvetkovic, J., Trentmann, O., Klemens, P.A.W., Neuhaus, H.E., 2018. In Concert: Orchestrated Changes in Carbohydrate Homeostasis Are Critical for Plant Abiotic Stress Tolerance. Plant and Cell Physiology 59, 1290–1299. https://doi.org/10.1093/pcp/pcy037

Raggi, L., Caproni, L., Carboni, A., Negri, V., 2019. Genome-Wide Association Study Reveals Candidate Genes for Flowering Time Variation in Common Bean (Phaseolus vulgaris L.). Front. Plant Sci. 10. https://doi.org/10.3389/fpls.2019.00962

Rapacz, M., Ergon, Å., Höglind, M., Jørgensen, M., Jurczyk, B., Østrem, L., Rognli, O.A., Tronsmo, A.M., 2014. Overwintering of herbaceous plants in a changing climate. Still more questions than answers. Plant Science 225, 34–44. https://doi.org/10.1016/j.plantsci.2014.05.009

Reckling, M., Bergkvist, G., Watson, C.A., Stoddard, F.L., Zander, P.M., Walker, R.L., Pristeri, A., Toncea, I., Bachinger, J., 2016. Trade-Offs between Economic and Environmental Impacts of Introducing Legumes into Cropping Systems. Front. Plant Sci. 7. https://doi.org/10.3389/fpls.2016.00669

Riday, H., n.d. Progress Made in Improving Red Clover (Trifolium pratense L.) Through Breeding 8.

Riday, H., Krohn, A.L., 2010. Increasing Population Hybridity by Restricting Self-Incompatibility Alleles in Red Clover Populations. Crop Science 50, 853–860. https://doi.org/10.2135/cropsci2009.05.0282

Rognli, O.A., Fjellheim, S., Pecetti, L., Boiler, B., 2013. Semi-natural grasslands as a source of genetic diversity. The role of grasslands in a green future: threats and perspectives in less favoured areas. Proceedings of the 17th Symposium of the European Grassland Federation, Akureyri, Iceland, 23-26 June 2013 303–313.

Rognli, O.A., Pecetti, L., Kovi, M.R., Annicchiarico, P., 2021. Grass and legume breeding matching the future needs of European grassland farming. Grass and Forage Science 76, 175–185. https://doi.org/10.1111/gfs.12535

Rosso, B.S., Pagano, E.M., 2005. Evaluation of Introduced and Naturalised Populations of Red Clover(Trifolium pratense L.) at Pergamino EEA-INTA, Argentina. Genetic Resources and Crop Evolution 52, 507–511. https://doi.org/10.1007/s10722-005-0777-z

Sato, S., Isobe, S., Asamizu, E., Ohmido, N., Kataoka, R., Nakamura, Y., Kaneko, T., Sakurai, N., Okumura, K., Klimenko, I., Sasamoto, S., Wada, T., Watanabe, A., Kohara, M., Fujishiro, T., Tabata, S., 2005. Comprehensive Structural Analysis of the Genome of Red Clover (Trifolium pratense L.). DNA Research 12, 301–364. https://doi.org/10.1093/dnares/dsi018

Saviranta, N.M., Julkunen-Tiitto, R., Oksanen, E., Karjalainen, R.O., 2010. Red clover (Trifolium pratense L.) isoflavones: root phenolic compounds affected by biotic and abiotic stress factors. Journal of the Science of Food and Agriculture 90, 418–423. https://doi.org/10.1002/jsfa.3831

Saviranta, N.M.M., Julkunen-Tiitto, R., Oksanen, E., Karjalainen, R.O., 2010. Leaf phenolic compounds in red clover (Trifolium pratense L.) induced by exposure to moderately elevated ozone. Environmental Pollution 158, 440–446. https://doi.org/10.1016/j.envpol.2009.08.029

Schaumont, D., Veeckman, E., Jeugt, F.V. der, Haegeman, A., Glabeke, S. van, Bawin, Y., Lukasiewicz, J., Blugeon, S., Barre, P., Leyva-Pérez, M. de la O., Byrne, S., Dawyndt, P., Ruttink, T., 2022. Stack Mapping Anchor Points (SMAP): a versatile suite of tools for read-backed haplotyping. https://doi.org/10.1101/2022.03.10.483555

Schulz, E., Tohge, T., Zuther, E., Fernie, A.R., Hincha, D.K., 2016. Flavonoids are determinants of freezing tolerance and cold acclimation in Arabidopsis thaliana. Sci Rep 6. https://doi.org/10.1038/srep34027

Schulz, E., Tohge, T., Zuther, E., Fernie, A.R., Hincha, D.K., 2015. Natural variation in flavonol and anthocyanin metabolism during cold acclimation in Arabidopsis thaliana accessions. Plant, Cell & Environment 38, 1658–1672. https://doi.org/10.1111/pce.12518

Segura, V., Vilhjálmsson, B.J., Platt, A., Korte, A., Seren, Ü., Long, Q., Nordborg, M., 2012. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nature Genetics 44, 825–830. https://doi.org/10.1038/ng.2314

Semerikov, V., Belyaev, A., Lascoux, M., 2002. The origin of Russian cultivars of red clover (Trifolium pratense L.) and their genetic relationships to wild populations in the Urals. Theor Appl Genet 106, 127–132. https://doi.org/10.1007/s00122-002-1033-y

Sengupta, S., Mukherjee, S., Basak, P., Majumder, A.L., 2015. Significance of galactinol and raffinose family oligosaccharide synthesis in plants. Front. Plant Sci. 6. https://doi.org/10.3389/fpls.2015.00656

Solberg, S.O., Yndgaard, F., Palmè, A., 2017. Morphological and phenological consequences of ex situ conservation of natural populations of red clover (Trifolium pratense L.). Plant Genetic Resources 15, 97–108. https://doi.org/10.1017/S1479262115000416

Steinshamn, H., Nesheim, L., Bakken, A.K., 2016. Grassland production in Norway.

Strobl, C., Wickelmaier, F., Zeileis, A., 2011. Accounting for Individual Differences in Bradley-Terry Models by Means of Recursive Partitioning. Journal of Educational and Behavioral Statistics 36, 135–153. https://doi.org/10.3102/1076998609359791

Szabados, L., Savouré, A., 2010. Proline: a multifunctional amino acid. Trends in Plant Science 15, 89–97. https://doi.org/10.1016/j.tplants.2009.11.009

Taylor, N.L., 2008. A Century of Clover Breeding Developments in the United States. Crop Science 48, 1. https://doi.org/10.2135/cropsci2007.08.0446

Taylor, N.L., Quesenberry, K.H., 1996. Red Clover Science. Springer Science & Business Media.

Theocharis, A., Clément, C., Barka, E.A., 2012. Physiological and molecular changes in plants grown at low temperatures. Planta 235, 1091–1105. https://doi.org/10.1007/s00425-012-1641-y

Tucak, M., Popović, S., Čupić, T., Španić, V., Meglič, V., 2013. Variation in yield, forage quality and morphological traits of red clover (Trifolium pratense L.) breeding populations and cultivars. Zemdirbyste-Agriculture 100, 63–70. https://doi.org/10.13080/z-a.2013.100.009

Ulloa, O., Ortega, F., Campos, H., 2003. Analysis of genetic diversity in red clover (Trifolium pratense L.) breeding populations as revealed by RAPD genetic markers. Genome 46, 529–535. https://doi.org/10.1139/g03-030

van Tienderen, P.H., de Haan, A.A., van der Linden, C.G., Vosman, B., 2002. Biodiversity assessment using markers for ecologically important traits. Trends in Ecology & Evolution 17, 577–582. https://doi.org/10.1016/S0169-5347(02)02624-1

Vestad, R. 1990. Red clover in Norwegian pastures. Past and future. In: O.A. Rognli (ed.) Present status and future prospects of Norwegian plant breeding. Basic aspects of plant breeding. Norsk landbruksforskning, Suppl. No. 9, pp. 165-172.

Vilhjálmsson, B.J., Nordborg, M., 2013. The nature of confounding in genome-wide association studies. Nat Rev Genet 14, 1–2. https://doi.org/10.1038/nrg3382

Visser, C.L.M. de, Schreuder, R., Stoddard, F., 2014. The EU's dependency on soya bean import for the animal feed industry and potential for EU produced alternatives. OCL 21, D407. https://doi.org/10.1051/ocl/2014021

Voisin, A.-S., Guéguen, J., Huyghe, C., Jeuffroy, M.-H., Magrini, M.-B., Meynard, J.-M., Mougel, C., Pellerin, S., Pelzer, E., 2014. Legumes for feed, food, biomaterials and bioenergy in Europe: a review. Agron. Sustain. Dev. 34, 361–380. https://doi.org/10.1007/s13593-013-0189-y

Waalen, W.M., Tanino, K.K., Olsen, J.E., Eltun, R., Rognli, O.A., Gusta, L.V., 2011. Freezing Tolerance of Winter Canola Cultivars is Best Revealed by a Prolonged Freeze Test. Crop Science 51, 1988–1996. https://doi.org/10.2135/cropsci2011.02.0098

Wexelsen, H. 1937. Undersøkelser over norsk rødkløver. Stammeundersøkelser. Tidsskrift for det norske landbruk, 2: 41-67.

Zanotto, S., Palmé, A., Helgadóttir, Á., Daugstad, K., Isolahti, M., Öhlund, L., Marum, P., Moen, M.A., Veteläinen, M., Rognli, O.A., Ergon, Å., n.d. Trait characterization of genetic resources reveals useful variation for the improvement of cultivated Nordic red clover. Journal of Agronomy and Crop Science n/a. https://doi.org/10.1111/jac.12487

Zhao, K., Tung, C.-W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., McClung, A.M., Bustamante, C.D., McCouch, S.R., 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. Nat Commun 2, 467. https://doi.org/10.1038/ncomms1467

Zhou, Y., Foka, I.C.K., Li, H., 2021. Tonoplast Inositol Transporters: Roles in Plant Abiotic Stress Response and Crosstalk with Other Signals. https://doi.org/10.20944/preprints202106.0515.v1

Paper I

Received: 25 August 2020 | Revised: 10 February 2021 | Accepted: 11 February 2021

DOI: 10.1111/jac.12487

CHILLING AND FREEZING STRESS

Journal of Agronomy and Crop Science | WILEY

Trait characterization of genetic resources reveals useful variation for the improvement of cultivated Nordic red clover

Stefano Zanotto[1] | Anna Palmé[2] | Áslaug Helgadóttir[3] | Kristin Daugstad[4] | Mika Isolahti[5] | Linda Öhlund[6] | Petter Marum[7] | Maria Ahlin Moen[1] | Merja Veteläinen[5] | Odd Arne Rognli[1] | Åshild Ergon[1]

[1]Faculty of Biosciences, Dept. of Plant Sciences, Norwegian University of Life Sciences, As, Norway

[2]Nordic Genetic Resource Centre (NordGen), Alnarp, Sweden

[3]Faculty of Agricultural and Environmental Sciences, Agricultural University of Iceland, Reykjavík, Iceland

[4]Division for Food Production and Society, Norwegian Institute of Bioeconomy Research (NIBIO), Volbu, Heggenes, Norway

[5]Boreal Plant Breeding Ltd., Jokioinen, Finland

[6]Lantmännen Lantbruk, Svalöv, Sweden

[7]Graminor AS, Ridabu, Norway

Correspondence
Åshild Ergon, Faculty of Biosciences, Dept. of Plant Sciences, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway.
Email: ashild.ergon@nmbu.no

Funding information
Nordisk Ministerråd; Nansenfondene

Abstract

Red clover (*Trifolium pratense* L.) is the most important forage legume in the Nordic region, but its utilization is limited by poor persistency. The improvement of cultivated red clover can potentially take advantage of the numerous wild populations and landraces conserved in gene banks; however, there is often limited information available on the phenotypic and genetic characteristics of this material. We characterized 48 populations conserved at NordGen for a number of traits and compared them to commercial cultivars. The material was evaluated in field trials at four locations over two years and in an experiment under controlled conditions. Considerable variation was identified, with stem length, growth type and flowering date having the highest broad sense heritabilities. Traits related to plant size were strongly associated with late flowering and upright growth and differed between landraces/cultivars on the one hand and wild populations on the other. There was a large genotype by environment interaction on winter survival, which only partially correlated with freezing tolerance under controlled conditions. A majority of gene bank accessions exceeded the commercial cultivars in winter survival and freezing tolerance and can therefore be a genetic resource for future improvement of these traits. The phenotypic variation among accessions was associated with two main axes of climatic variation at the collection site. Petiole length of young plants under controlled conditions as well as plant size in the field increased with increasing summer temperature and decreasing summer precipitation, while number of leaves and an apparent vernalization requirement, recorded under controlled conditions, increased with decreasing annual and winter temperature. We discuss the implications these results have for collection, conservation and utilization of red clover genetic resources in the Nordic region.

KEYWORDS

breeding, canonical correlation analysis, freezing tolerance, local adaptation, *Trifolium pratense*, winter survival

1 | INTRODUCTION

Agriculture in large parts of Sweden, Finland, Norway and Iceland is grassland-based because the cold climate, short growing season, shallow soils and/or steep slopes make the land less suitable for other crops. Due to the warming effect of the North Atlantic Current this region has, however, a mild climate compared to other regions at the same latitudes (Helgadóttir et al., 2014). The unique combination of annual cycles of photoperiod and temperature necessitates the availability of cultivars of grassland species that are particularly adapted to this region. Temporal grasslands that are regularly re-sown cover 48%, 40% and 28% of the total agricultural area in Norway, Sweden and Finland, respectively (Helgadóttir et al., 2014; Statistics Sweden, 2013; Steinshamn et al., 2016). In Iceland, most of the agricultural area is covered by grasslands, but the majority of this is natural grasslands (Helgadóttir et al., 2014). The Nordic region is, like the rest of Europe, heavily dependent on imports of protein-rich biomass for feed and food (De Visser et al., 2014; Voisin et al., 2014). Greater cultivation of forage legumes can help reduce this protein deficit, and further increase sustainability, as their ability to make use of atmospheric N_2 through symbiotic N_2 fixation reduces the requirement for nitrogen fertilizer in the cropping system (Jensen et al., 2012; Lüscher et al., 2014; Reckling et al., 2016).

Red clover is the most important forage legume cultivated in northern Europe (Annicchiarico et al., 2015), including the Nordic region, where it is cultivated almost exclusively in mixtures with grasses over a wide range of climatic conditions (Helgadóttir et al. 2014; Steinshamn et al., 2016). Like most perennial forages, red clover is a cross-pollinated species with a high level of genetic variation within populations. It is thought to have originated near the Mediterranean basin (Boller et al., 2010). Domesticated red clover from Spain spread to other parts of Europe and reached the Nordic countries in the 18th century (Kjærgaard, 2003). As the use of red clover spread across Europe, a common practice was to re-sow seed harvested from a restricted area, often within a single farm. This led to the development of locally adapted landraces as a result of semi-conscious selection by the farmers, who exploited the natural selection by producing seeds from the surviving plants in the field (Boller et al., 2010). The use of landraces in Europe declined around 1970 with the increased use of highly productive cultivars developed by modern breeding (Boller et al., 2010). In the Nordic region, red clover breeding was first initiated in Sweden in the beginning of the last century, with several cultivars produced (Olsson, 1997). The first Finnish red clover cultivar 'Tammisto' was released in 1948, prior to which mainly Swedish cultivars were used in Finland (Valle, 1958). Although there was red clover breeding in Norway from the middle of the last century, the landrace 'Moldstad' dominated until the 1980s (Vestad, 1990). In Iceland, cultivars bred in other Nordic countries are cultivated. Current red clover breeding in Norway, Sweden and Finland is based mostly on progeny testing and phenotypic selection of full-sib families, sometimes combined with phenotypic mass selection, followed by creation of synthetic populations. Population improvement through recurrent phenotypic selection is also used. Persistency is an important breeding goal, along with dry matter yield, disease resistance, seed yield and forage quality. Phenotypic evaluation and selection for persistency is complex, since many biotic and abiotic factors affect persistency (Annicchiarico et al., 2015; Ergon et al., 2019; Herrmann et al., 2008). Throughout the Nordic region, the main reason for limited persistency is the various stresses which affect the plants during the winter (Helgadóttir et al. 2014). Given the environmental variability and limited seed market in the Nordic region, the development of cultivars with wide adaptation is desired (Helgadóttir et al., 2000). Cultivars with wide adaptation will be even more important considering the high variability in the climatic conditions expected in the future (Ergon et al., 2018; Rapacz et al. 2014).

The genetic resources found in landraces and wild populations are generally underutilized in plant breeding, both in the Nordic countries and in the rest of Europe (Brozynska et al., 2016; Rognli et al., 2013). Red clover is distributed all over the Nordic countries, but rarely in the alpine regions (Mossberg & Stenberg, 2012). Wild populations have probably either naturalized from cultivated populations or are the result of hybridization between wild and cultivated populations (Daugstad, 2016). In addition to over 100 red clover cultivars, NordGen (the Nordic Genetic Resource Centre) currently holds over 400 other red clover accessions in its collection. These accessions may be a source of variation for traits that are lacking or need improvement in cultivated material, such as superior winter survival or resistance to specific abiotic or biotic stresses. However, introducing such material in a breeding programme is challenging. It is likely to carry undesirable traits, most notably low biomass yield, which needs to be enhanced through subsequent breeding. The various accessions are poorly characterized, so identification of the most suitable ones is difficult. While molecular marker data may be useful for germplasm characterization (Greene et al., 2004; Kölliker et al., 2003; Pagnotta et al., 2011), characterizing the phenotypic variation in agronomically relevant traits is particularly valuable for plant breeding and germplasm conservation, since selectively neutral molecular genetic variation tend not to coincide with phenotypic variation (Riday, 2010; Rognli et al., 2013). Based on the premise that traits displayed by an accession reflect the selection pressures of the environment from which it originates, climatic variables at the collection sites can be used to predict the phenotype of gene bank accessions, helping breeders and gene bank managers to more effectively seek germplasm with specific phenotypic characteristics (e.g. Endresen et al., 2011; Hijmans et al., 2003; Khazaei et al., 2013). Adding high-density genomic data to the analyses can help identify valuable germplasm and be used in development of selection tools for breeding (Blanco-Pastor et al. 2020).

Large genetic variation for morphological and performance traits has been described in red clover material of various origin (Annicchiarico & Pagnotta, 2012; Dias et al., 2008; Greene et al., 2004; Kouamé & Quesenberry, 1993; Pagnotta et al., 2011; Rosso & Pagano, 2005; Tucak et al., 2013). However, only a few studies have evaluated red clover material of Nordic origin. Helgadóttir et al. (2000) evaluated yield and persistency of 13 cultivars of red

clover in pure stand and in mixture with timothy across five Nordic locations, while Solberg et al. (2015) characterized morphology and development of 16 Norwegian accessions from NordGen, grown as single plants in a greenhouse. The aim of the present study was to evaluate accessions of red clover in the Nordic collection at NordGen and compare them to commonly used cultivars in the region. We present results from spaced plant trials carried out at four locations across the Nordic countries, as well as from an experiment under controlled conditions, and describe i) variance components and broad sense heritability estimates, ii) correlations between traits and iii) relationships between phenotypic traits and climate variables at the collection sites.

2 | MATERIAL AND METHODS

2.1 | Plant material

A total of 48 accessions of red clover were selected from the NordGen collection, representing the geographical distribution of the species in the Nordic region (Figure 1, Supplementary Table 1). This included 19 landraces, 2 semi-wild and 27 wild populations. The majority of Finnish populations in the collection are landraces, while the majority of Norwegian populations are wild, resulting in an uneven geographical distribution of population types. In addition, the commercial cultivars Lea, Linus and Gandalf (Graminor, Norway), SW Ares and SW Yngve (Lantmännen, Sweden) and Saija (Boreal, Finland) were included in the experiment under controlled conditions, while Lea, SW Yngve and Saija were included in the field trials.

2.2 | Field trials

Spaced plant trials were established at Jokioinen, Finland (60°48'N, 23°29'E, 115 m a.s.l.), Korpa, Iceland (64°09'N, 21°84'W, 35 m a.s.l.), Løken, Norway (61°12'N, 9°06'E, 545 m a.s.l.) and Lännäs, Sweden (63°08'N, 17°43'E, 25 m a.s.l.) (Figure 1). Data describing the climate during the experiment (summer 2014 - spring 2016) at each location are presented in Supplementary Table 2. Plants were raised from seed in a greenhouse in spring 2014 and transplanted into bare ground during summer or early autumn at all locations except at Korpa, where plants were transplanted into a field which had been sown with a mixture of timothy (*Phleum pratense* L.), cv. Snorri, and smooth meadow grass (*Poa pratensis* L.), cv. Kupol. The different experimental layout at Korpa was implemented to avoid the very high winter mortality of red clover planted into bare ground seen in previous experiments at this location. Strong competition from the grass species was avoided by applying only a low level of N fertilizer in the sowing year. The experimental design was a complete randomized block with four replicates. Populations formed plots within each block, each made up of a group of five different genotypes. Each population was thus represented by 20 genotypes in total. The distance between individual plants within and between plots was 1 m. The plants were phenotyped for the following traits: autumn status

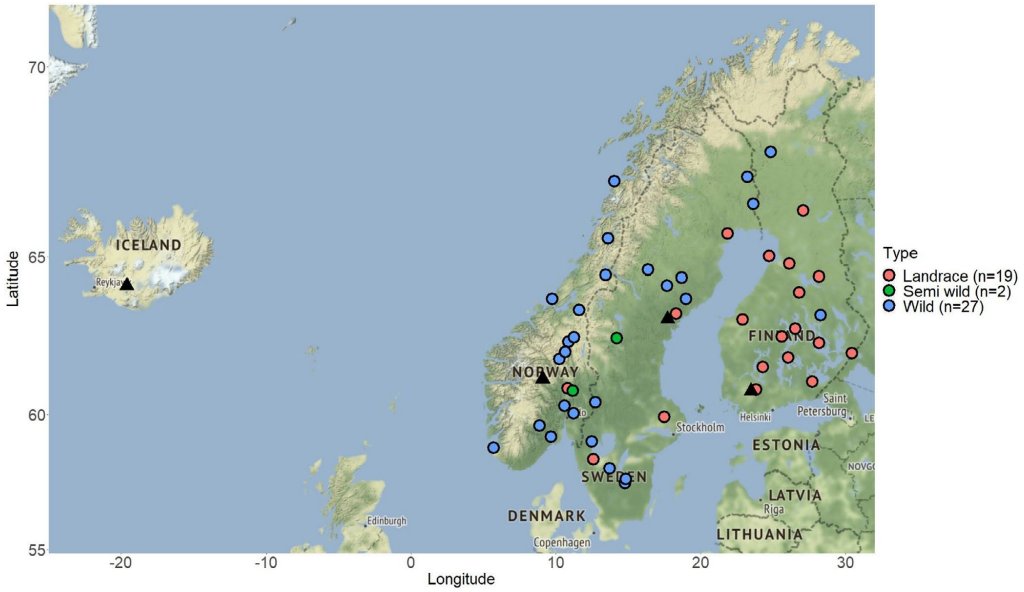


FIGURE 1 Collection sites of the landraces (19), semi-wild (2) and wild (27) red clover populations (colored coded), and test locations where the field trials were conducted (black triangles) [Colour figure can be viewed at wileyonlinelibrary.com]

in 2014; leaf area, stem length, flowering date, flowering score and yield score in 2015; winter survival and spring status in 2015 and 2016. A detailed description of all traits recorded is provided in Supplementary Table 3.

2.3 | Experiment under controlled conditions

Seeds were scarified with sandpaper, sown in a peat soil in early January and germinated in a greenhouse (59°40′ N, 10°47′ E) at 16°C with natural light supplied with additional metal halide lamps (approximately 100–125 μmol/m^{2}s^{-1} PAR) to a photoperiod of 12 hr. After germination, individual young seedlings were transplanted into pots (28 cm^3) filled with peat soil and grown under the same conditions. When the plants were six weeks old, the number of leaves and the length of the longest petiole was recorded, and the plants were moved to cold acclimating conditions at 3–4°C, 12 hr photoperiod and 110 μmol/m^{2}s^{-1} PAR, for two weeks. For determination of freezing tolerance (LT50; temperature required to kill 50% of the plants), cold acclimated plants were exposed to four different test temperatures. Six plants of each population were placed in a programmed freezing chamber initially set at 2°C. The temperature was first lowered from 2°C to −3°C at 1 °C h^{-1} and kept at this level for 12 hr to ensure even freezing, after which the temperature was lowered again by 1 °C h^{-1} down to the test temperature; −8°C, −11°C, −14°C or −17°C. When the temperature reached the test temperature, it was kept there for one hour before the temperature was raised, again by 1 °C h^{-1}, up to 2 °C. After thawing, the pots were returned to the greenhouse and kept there for three weeks, before survival of each plant was recorded. The experiment was repeated three times, resulting in approximately 18 plants per population and test temperature. The data from the three experiments were pooled before LT50 was calculated by probit analysis using PROC PROBIT in SAS 9.2 (SAS Institute, Inc., Car, NC, USA). For determination of timing of reproductive development, 9–10 non-acclimated seedlings of most populations were transplanted into 1-L pots with peat soil (one plant per pot) and kept in the greenhouse (conditions as above, but temperature occasionally exceeding 16 °C) until they were approximately 3 months old. The photoperiod was then increased to 23 hr, and the number of days until the beginning of stem elongation was recorded. Stem elongation is the earliest sign of reproductive development in red clover, and we defined days to elongation as the number of days at 23 hr photoperiod until the longest internode measured at least 2 cm. Recording was done three times a week until 57 days had passed.

2.4 | Data analysis

Variance component analyses of the field data were conducted using the residual maximum likelihood (REML) procedure in DeltaGen as described by Jahufer and Luo (2018). The linear mixed model used for analysis was:

$$Y_{ijm} = M + g_i + l_j + (gl)_{ij} + b_{jm} + \varepsilon_{ijm}$$

where Y_{ijm} is the value of a trait measured in population i in replicate m at location j; M is the overall mean; g_i is the random effect of population i, $N(0, \sigma_g^2)$; l_j is the fixed effect of location j; $(gl)_{ij}$ is the random effect of the interaction between population i and location j, $N(0, \sigma_{gl}^2)$; b_{jm} is the random effect of replicate m within location j, $N(0, \sigma_b^2)$; and ε_{ijm} is the residual effect for population i in replicate m at location j, $N(0, \sigma_\varepsilon^2)$. The statistical significance of the variance components was further assessed using the likelihood ratio test. An estimate of entry-mean broad sense heritability (h_b^2) (Falconer, 1989, hereafter referred to as just heritability), of the considered traits, was calculated by selecting the heritability option provided in DeltaGen (Jahufer & Luo, 2018), using the following formula:

$$h_b^2 = \frac{\sigma_g^2}{\sigma_g^2 + \frac{\sigma_{gl}^2}{n_l} + \frac{\sigma_\varepsilon^2}{n_l n_b}}$$

where σ_g^2, σ_{gl}^2, and σ_ε^2 are the population, population by location and pooled error variance components and n_l and n_b are the number of locations and blocks per location, respectively. Genotypic correlations were estimated using the MANOVA option in DeltaGen (Jahufer & Luo, 2018) using a model which included the random effects of population, location and block within location for the correlation among different traits, and the random effects of population and block within location for the correlation of single traits among locations. These are genotypic and not genetic correlations (according to Falconer, 1989) because our experimental design did not allow the estimation of the additive genetic component of variance. Phenotypic correlations were calculated on averaged data from each population and location using the rcorr function (Harrell, 2020) in RStudio version 1.1.463 (RStudio: Integrated Development for R. RStudio, Inc., Boston).

A canonical correlation analysis (CCA) was performed to test the association of each of the phenotypic traits (population means) recorded, with variables describing the climate at the collection sites of the populations. Annual mean temperature, temperature seasonality, maximum temperature of the warmest month, minimum temperature of the coldest month, mean temperature in the warmest quarter, mean temperature in the coldest quarter, annual precipitation, precipitation seasonality, precipitation in the warmest quarter and precipitation in the coldest quarter at the place of origin of each of the 48 populations from the NordGen collection were retrieved from the Worldclim database (Worldclim.org) using the diva-gis software (diva-gis.org). In addition, we used latitude and altitude as descriptors of climate variation. Altitude and historical climate data for the time period 1950–2000 were obtained with 2.5 arcosin minutes resolution (the highest resolution available with this software), corresponding to an area of 4.5 x 4.5 km. More information on the climatic variables used is provided in Supplementary Table 4. The CCA was performed using the function cca of the R package vegan (Oksanen et al. 2019) in RStudio

version 1.1.463. The squared canonical correlation coefficients were used to assess the proportion of the total variation explained by each canonical variable. The canonical variable loadings express the magnitude of each of the original trait's correlation with the derived canonical trait variables, and the original climate variable's correlation with the derived canonical climate variables (Vicario et al., 1989).

3 | RESULTS

Almost all plants established successfully at all locations, and there was both a significant variation among populations, and a relatively strong population by location interaction (GxE) on autumn status in the establishment year (Supplementary Table 5). Winter survival varied markedly between locations and years (Figure 2a). In

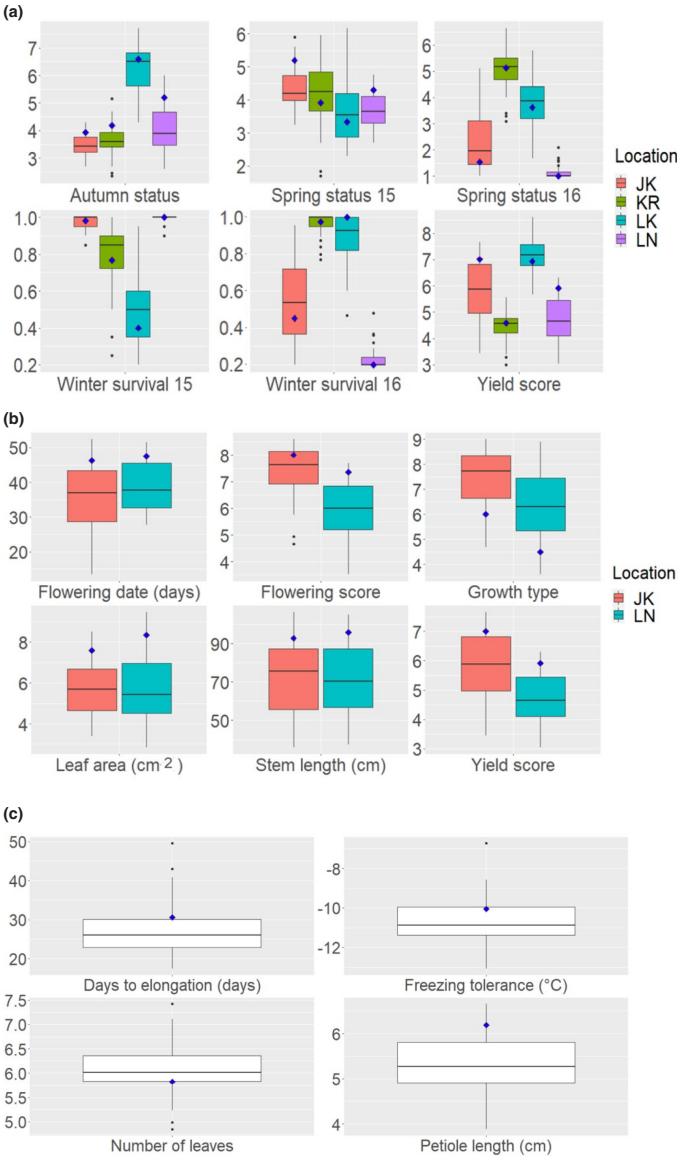


FIGURE 2 Boxplots showing the variation in traits among Nordic red clover populations included in this study. (A), Autumn status, spring status, yield score (all on a scale from 1 to 9) and winter survival (as a proportion from 0 to 1) recorded at four locations; (B), traits recorded during the growing season at the two locations with good winter survival, and thus good representation of all populations (flowering score, growth type and yield score on a scale from 1 to 9; for growth type higher values mean more prostrate growth); (C), traits recorded in an experiment under controlled conditions. Data were averaged for each population prior to plotting. The blue dots correspond to averages of the cultivars included. Trait descriptions are given in Supplementary Table 3. JK, Jokioinen; KR, Korpa; LK, Løken; LN, Lännäs [Colour figure can be viewed at wileyonlinelibrary.com]

the spring of 2015, a very good survival was observed at Jokioinen and Lännäs (an average of all populations of 98%–99%), while it was lower at Korpa (80%) and Løken (49%). During the next winter, the survival rate among the remaining plants was higher at Løken (86%) and Korpa (96%), while it was low at Jokioinen (44%) and almost zero at Lännäs. There was significant GxE for spring status and winter survival in 2015, and no significant variance between populations for these traits (Supplementary Table 5). For the yield score at the first cut in 2015, there was significant variation among populations and significant GxE. The heritability for yield score and autumn status was 0.45 and 0.52, respectively. Genotypic correlations between trait values at the different locations showed that the GxE was largely due to different responses of the populations to the growing conditions at Korpa compared to the other locations, as well as different responses to the winters at Korpa, Løken and Jokioinen/Lännäs (Table 1). The different behaviour of populations at Korpa was also evident from the fact that the proportion of plants that flowered during the 2015 growing season was considerably lower (35%) than at the other locations, where all plants flowered.

TABLE 1 Genotypic correlation coefficients ($p < .05$) for autumn status, spring status and yield score in 2014–2015 recorded in single plant trials with 51 Nordic red clover populations at four test locations. See Supplementary Table 3 for a description of the traits

Autumn status 2014	Jokioinen	Lännäs	Løken
Lännäs	0.70	-	
Løken	0.60	0.67	-
Korpa	0.22	0.01	0.18
Spring status 2015	**Jokioinen**	**Lännäs**	**Løken**
Lännäs	=0.64	-	
Løken	−0.13	−0.44	-
Korpa	0.20	0.39	0.14
Yield Score 2015	**Jokioinen**	**Lännäs**	**Løken**
Lännäs	0.80	-	
Løken	0.11	−0.25	-
Korpa	−0.04	−0.01	0.43

3.1 | Traits expressed during the growing season

Due to winter mortality and subsequent missing data, we based further analysis of traits expressed during the growing season on data collected after the non-selective winters at Jokioinen and Lännäs in 2015, in addition to autumn status at those locations in 2014. There was significant GxE for all traits except growth type, but the variance between populations was considerably larger than GxE for all traits except autumn and spring status (Supplementary Table 6, see also Figure 2b). The highest entry-mean broad sense heritabilities were estimated for stem length, growth type and flowering date (Table 2). The genotypic correlations between the different traits were all moderately high to high, with absolute correlation coefficients ranging from 0.76 to 0.97 (Table 2). Erect growth habit and late flowering were associated with high values for all growth and size-related traits. These traits were also positively correlated with petiole length (e.g. r = 0.80 for yield score), and somewhat negatively correlated with number of leaves, recorded in the experiment under controlled conditions (Supplementary Table 7). Notably, the number of days to elongation recorded in the experiment under controlled conditions was not significantly correlated with flowering date in the field.

Wild/semi-wild populations had a more prostrate growth type, earlier flowering and generally smaller plant size on average than landraces and cultivars, and cultivars did not outperform the best landraces in yield score (Table 3, Figure 2b).

3.2 | Winter survival and spring status after selective winters

To focus on informative data sets, and avoid many missing values, analysis of variation in winter survival was based on the first selective winter at Korpa, Løken and Jokioinen (2015, 2015 and 2016, respectively). Winter survival and spring status were strongly correlated within all three data sets (r = 0.95–0.98). Heritability for winter survival and spring status was very low (Supplementary Table 8). While there was a moderate genotypic correlation between winter survival at Løken and Jokioinen (r = 0.55), winter survival at Korpa

TABLE 2 Genotypic correlation coefficients ($p < .05$) and entry-mean broad sense heritability estimates ($p < .05$) for the traits expressed during the growing season in single plant trials with 51 Nordic red clover populations at the two test locations that were not affected by winter mortality during the first winter (Jokioinen and Lännäs). See Supplementary Table 3 for a description of the traits

	Autumn status	Spring status	Flowering date	Flowering score	Growth type	Leaf area	Stem length	Yield score
Spring status	0.92	-						
Flowering date	0.84	0.76	-					
Flowering score	0.83	0.86	0.89	-				
Growth type	−0.96	−0.88	−0.88	−0.83	-			
Leaf area	0.96	0.89	0.88	0.79	−0.96	-		
Stem length	0.91	0.87	0.96	0.93	−0.95	0.92	-	
Yield score	0.90	0.90	0.90	0.95	−0.92	0.90	0.97	-
h_b^2	0.45 ± 0.10	0.47 ± 0.08	0.89 ± 0.02	0.69 ± 0.06	0.89 ± 0.02	0.65 ± 0.06	0.94 ± 0.01	0.75 ± 0.05

TABLE 3 Analysis of variance to assess the difference between population types of the Nordic red clover populations used in the field trials (**A**) and in the controlled experiment (**B**). Group means followed by different letters are significantly different following the Tukey's test for multiple groups comparisons

A				
Trait	**Cultivars ($n = 3$)**	**Landraces ($n = 19$)**	**Wild ($n = 29$)**	***P*-value**
Autumn status	4.57^{a}	4.10^{a}	3.49^{b}	***
Spring status 15	4.75 ^{a}	4.20^{a}	3.81^{b}	**
Flowering date	46.92^{a}	43.19^{a}	32.47^{b}	***
Flowering score	7.69^{a}	7.13^{a}	6.33^{b}	**
Growth type	5.25^{a}	6.16^{a}	7.56^{b}	***
Leaf area	7.98^{a}	6.62^{a}	5.00^{b}	***
Stem length	94.40^{a}	84.10^{a}	62.10^{b}	**
Yield score	6.50^{a}	5.90^{a}	4.81^{b}	***
Winter survival Korpa 2015	0.77	0.84	0.78	NS
Winter survival Løken 2015	0.40	0.45	0.52	NS
Winter survival Jokioinen 2016	0.32^{ab}	0.25^{a}	0.57^{b}	***
B				
Trait	**Cultivars ($n = 6$)**	**Landraces ($n = 19$)**	**Wild ($n = 29$)**	***P*-value**
Days to elongation	30.58	26.51	27.08	NS
Number of leaves	5.83	6.01	6.14	NS
Petiole length	6.19^{a}	5.67^{a}	4.92^{b}	***
Freezing tolerance	−10.06	−10.48	−10.98	NS

*, $p < .05$; **, $p < .01$; ***, $p < .001$; NS, not significant

had a weak negative genotypic correlation with winter survival at the two other locations (r = −0.38 and −0.23, Supplementary Table 8). Similarly, freezing tolerance (i.e. low LT50 value) was only partly correlated with winter survival at Jokioinen in 2016 (phenotypic, r = −0.49), and not with winter survival at Løken or Korpa in 2015 (Supplementary Table 7).

Wild/semi-wild populations had significantly better winter survival at Jokioinen than landraces, while cultivars were intermediate. There were no significant differences between population types in winter survival at Løken or Korpa or in freezing tolerance (Table 3). However, the range in population averages of these traits was large, and cultivars were among the poorest 50% (Figure 2b, 2c).

3.3 | Relationships between winter survival and traits expressed during the growing season and between traits and the environment at the collection sites

The CCA of the phenotypic traits and the climatic variables resulted in two significant canonical correlations with R^2-values of 0.83 ($p < .003$) and 0.79 ($p < .05$), respectively. The first canonical variable for the phenotypic data set had the strongest correlation with petiole length recorded under controlled conditions (Table 4). Flowering time, field traits related to plant size (except autumn status at Korpa) and LT50 were moderately correlated to both canonical variables, but more so to the first than to the second. The number of leaves and days to elongation under controlled conditions, as well as winter survival at all locations, were more correlated with the second canonical variable than with the first. For the climate data set, the first canonical variable had the highest correlations with temperature and precipitation during summer. This variable may partly be explained by the difference between wild populations (largely from Norway) and landraces (largely from Finland) (Figure 1, Supplementary Table 1). Thus, considering the sign of the coefficients, populations from locations with warm summers and low precipitation tended to have long petioles, and to some extent other traits related to large plant size (and vice versa). The second canonical variable had the strongest correlations with mean annual temperature and temperature during winter. Thus, populations originating from locations with low annual temperatures and in particular cold winter temperatures tended to have many leaves, late stem elongation (probably due to a facultative vernalization requirement; see Discussion), good winter survival at Løken and Jokioinen, but low winter survival at Korpa. Interestingly, the correlations of autumn status and winter survival at Korpa had the opposite sign compared to those of the same traits at the other two locations, for both canonical variables. It is

TABLE 4 Canonical correlation analysis of phenotypic traits and climatic conditions at the collection site of 48 Nordic wild, semi-wild or landrace red clover populations. The correlations between each phenotypic trait and climate variable and their respective first two canonical variables (Can 1 and Can 2). Phenotypic traits recorded in an experiment under controlled conditions were included, together with traits recorded in field trials at four locations. See Supplementary Table 3 for a description of the traits and Supplementary Table 4 for a description of climate variables. When location is not indicated, field traits are averages of data recorded before and after the non-selective winter at Jokioinen and Lännäs in 2014–2015. Correlations in bold are at least twice as strong for the indicated canonical variable than for the other

Phenotype			Climate		
Original variable	**Can 1**	**Can 2**	**Original variable**	**Can 1**	**Can 2**
Autumn status	−0.64	0.34	Altitude	0.40	−0.50
Autumn status Korpa	0.18	−0.18	Latitude	0.22	−0.34
Autumn status Løken	−0.61	0.32	Mean annual temp.	−0.19	**0.62**
Flowering date	−0.68	0.46	Temp. seasonality	−0.35	−0.30
Flowering score	−0.52	0.35	Max. temp. warmest month	**−0.73**	0.28
Growth type	0.59	−0.50	Min. temp. coldest month	0.06	**0.47**
Leaf area	−0.64	0.34	Mean temp. warmest quarter	−0.65	0.49
Stem length	−0.66	0.48	Mean temp. coldest quarter	0.06	**0.51**
Yield score	−0.53	0.42	Annual precipitation	0.57	0.32
Spring status 2015	−0.50	0.31	Precip. seasonality	0.01	**−0.15**
Winter survival Jokioinen 2016	0.49	−0.61	Precip. warmest quarter	**0.58**	0.21
Winter survival Korpa 2015	−0.16	**0.42**	Precip. coldest quarter	0.55	0.31
Winter survival Løken 2015	0.02	**−0.40**			
Days to elongation	0.11	**−0.49**			
Number of leaves	0.13	**−0.61**			
Petiole length	**−0.79**	0.34			
Freezing tolerance	−0.31	0.21			

also notable that while winter survival at Jokioinen was to some extent correlated with both canonical variables, winter survival at Løken was only correlated to the second variable. These results are in agreement with the strong GxE observed for winter survival and are supported by the variation in correlations between winter survival in the different locations and other traits (Supplementary Table 7 and Supplementary Table 8).

4 | DISCUSSION

4.1 | Traits expressed during the growing season

Positive correlations between erect growth, plant size, stem length and leaf size have also been found in very diverse material of red clover from different parts of the world (Annicchiarico & Pagnotta, 2012; Greene et al., 2004; Herrmann et al., 2008; Pagnotta et al., 2011; Rosso & Pagano, 2005; Solberg et al., 2015; Tucak et al., 2013). In these studies, erect growth habit and large plants are sometimes associated with early flowering, sometimes with late flowering, probably depending on both the origin of the material, and the conditions at the test site.

Interestingly, the number of days to stem elongation in the controlled experiment was not correlated with flowering time in the field. Red clover flowers in response to long photoperiods and increasing temperatures; the first visible sign of reproductive development is the initiation of stem elongation. The photoperiod requirement for Nordic red clover is around 16–18 hr (Lunnan, 1989; Van Dobben, 1964). There is little requirement for vernalization in red clover, but in some populations flowering is stimulated by a cold treatment (Fejer, 1960; Van Dobben, 1964). In a previous study of vernalized plants of three Norwegian cultivars, it was found that the temperature requirement for onset of stem elongation was saturated between 10 and 14°C, while the temperature requirement of flower bud initiation was higher (Ergon et al., 2016). Days to elongation in our controlled experiment was measured at 16°C and 23 hr photoperiod. These conditions are likely to have been saturating for the stimulation of elongation. It therefore appears that the variation in timing of stem elongation reflects either variation in a facultative vernalization requirement (as the plants were not vernalized), or variation in loss of juvenility. Variation in flowering date in the field trials in 2015 (older and vernalized plants) probably reflects variation in photoperiod and temperature responses, controlled by other genes. Given the weak requirement for vernalization in red clover, all plants were probably fully vernalized at all locations. Despite this, a large proportion of plants did not flower at Korpa in 2015. Korpa is the most Northern location with the longest photoperiod in summer,

so photoperiod cannot be the reason for this. Thus, the low summer temperature (mean of 9°C) remains as the most likely cause of limited flowering at Korpa. Autumn status was another trait that was expressed differently at Korpa compared to the other three locations. Korpa belongs to a different agroclimatic zone than Jokioinen (Björnsson, 1993; Helgadóttir & Björnsson, 1994). In addition to being the northernmost location, Korpa is characterized by an oceanic climate with cooler and wetter summers than the other locations; this was also the case in 2014 (Supplementary Table 2). Analysis of GxE on yield in perennial ryegrass also showed that populations behaved differently at Korpa compared to other Nordic locations with more continental climates (Helgadòttir et al., 2018). Both the oceanic climate and the rapid shortening of daylength in the fall may have affected growth as well as growth cessation of the various populations differently at Korpa relative to the other locations. Also, at Korpa the red clover single plants were planted in a field previously sown with a mixture of timothy and smooth meadow grass. Although grass growth was limited by low N fertilization, we cannot rule out that clover-grass interactions may have affected the growth of red clover plants belonging to different populations differently.

4.2 | Winter survival and its relationship to other traits

Winter survival at Jokioinen and to some extent at Løken was negatively correlated with plant size. This is in agreement with earlier studies in Italian red clover (Annicchiarico & Pagnotta, 2012) as well as in Nordic white clover populations (Aasmo Finne et al., 2000). It is possible that the populations with more above-ground growth (particularly if they grow late in the autumn) tend to allocate less resources to storage tissues in the tap root and crown. Lack of organic reserves can reduce the ability to survive the winter and regrow the following spring. Thus, the variation we observed may represent the differences between 'competitors', allocating resources to growth and 'stress tolerators', allocating resources to storage and maintenance, as described in Grime's classical plant strategy theory (Grime, 1977). Winter survival is a very complex trait. Winter stresses include freezing, ice encasement, soil movements, waterlogging, drought, fungal pathogens and lack of light (Bélanger et al., 2006; Rapacz et al., 2014). Winter survival depends on various stress resistance mechanisms and is strongly influenced by cold acclimation and growth cessation in autumn, as well as deacclimation and growth resumption in spring (Rapacz et al., 2014). Freezing tolerance, measured under controlled conditions, can be used as a proxy for cold acclimation status and winter survival ability in some grass species (Gusta et al., 2001; Hulke et al., 2008; Waldron et al., 1998) and in white clover (Annicchiarico et al. 2001). However, freezing tolerance is not always correlated with winter survival, because other stresses than freezing may be predominant, and because resistance mechanisms towards the various stress factors are multiple and likely to be, to some extent, genetically independent. This was also shown in our experiments; there was considerable GxE for winter survival, and freezing tolerance was correlated to some extent with winter survival at Jokioinen in 2016, but not at Løken or Korpa in 2015, suggesting that freezing was an important winter stress at Jokioinen but that other winter stresses were prevalent at the two other locations, in the respective years. The lack of correlations between LT50 and winter survival in the field could also be due to the method of testing freezing tolerance, such as cold acclimation conditions and duration of the freezing treatment. A prolonged freezing test, LD50 (lethal duration time for 50% kill), may be better correlated to winter survival in the field. Waalen et al. (2011) found that the LD50 test was more effective than the LT50 test in identifying in freezing tolerance in canola cultivars. Surprisingly, winter survival at Korpa was negatively correlated with winter survival at Jokioinen and Løken, both according to genotypic and canonical correlations. The oceanic climate typical for Iceland is characterized by less severe winter temperatures, and plants are subjected to stresses such as frequent freeze-thawing cycles, waterlogging and ice encasement. The different winter survival at Korpa may also be related to the different expression of autumn status at Korpa, as discussed above. Moreover, soil moisture and excessive rainfall during the acclimation was demonstrated to negatively affect the capacity of plants to obtain a full acclimation (Bélanger et al., 2006).

4.3 | Breeding prospects and management of gene bank material

Due to high GxE for winter survival, breeding of cultivars with improved persistency and broad adaptation to the Nordic region requires testing at several locations over several years. Alleles involved in various processes and mechanisms will have to be combined, either at individual or population level, to ensure adaptation to the variation in Nordic winters (e.g. as initiated by Helgadóttir et al., 2000). This may be even more important in the prospect of increasing variability in the winter climate expected in the future at Nordic latitudes (Rapacz et al. 2014).

Controlled experiments aiming at characterizing LT50 and other specific winter survival-related traits may speed up the breeding process and help ensure that several traits are combined. For example, successful enhancement of freezing tolerance and winter survival in the field was realized through the recurrent selection of alfalfa and red clover plants surviving controlled freezing treatments (Bertrand et al., 2016; Castonguay et al., 2009). However, it is clear from our results that characterizing only LT50 is not sufficient when breeding for Nordic conditions, as it explains only a portion of the variation in winter survival. We found very strong positive correlations between yield score and stem length in the field, and petiole length measured on young plants in the greenhouse. The latter trait could be used in rapid evaluation of a large number of individuals.

Several gene bank accessions outperformed the cultivars, particularly for winter survival, but also for yield and flowering scores,

suggesting that they can be utilized in further improvement of persistency, biomass and seed yield. Also, the large variation in morphological traits such as growth type, leaf area and stem length that we found can be exploited for breeding cultivars with various morphologies, adapted to different end uses and agronomic systems (e.g. for grazing).

A gene bank collection should be as broad as possible, representing most of the existing variation. At the same time, there are economical limitations, so the redundancy should be kept low. Our results suggest that the Nordic red clover collection should at least contain accessions which capture the variation along the two identified phenotype climate gradients in order to cover the variability present in the region. Additional gradients may be identified in future studies with more accessions, capturing larger variation in climatic conditions. Development of predictive models based on environmental information, phenotypic data as well as genomic data, would further increase the precision in conservation and utilization of Nordic red clover germplasm.

ACKNOWLEDGMENTS

This work received funding from the Nordic Council of Ministers' Arctic Co-operation Programme and Nansenfondet.

AUTHOR CONTRIBUTIONS

Anna Palmé, Áslaug Helgadóttir, Merja Veteläinen and Petter Marum planned the field trials. Áslaug Helgadóttir, Linda Öhlund, Kristin Daugstad and Mika Isolahti conducted the field trials. Åshild Ergon and Maria Ahlin Moen planned and conducted the experiment under controlled conditions. Stefano Zanotto analysed the data. Åshild Ergon and Odd Arne Rognli supervised the data analysis. Stefano Zanotto and Åshild Ergon wrote the paper. All authors commented on the manuscript and approved the final version.

ORCID

Stefano Zanotto https://orcid.org/0000-0001-6588-7237
Áslaug Helgadóttir https://orcid.org/0000-0003-0107-0177
Åshild Ergon https://orcid.org/0000-0003-1275-0450

REFERENCES

Aasmo Finne, M., Rognli, O. A., & Schjelderup, I. (2000). Genetic variation in a Norwegian germplasm collection of white clover (*Trifolium repens* L.). *Euphytica*, *112*, 33–44. https://doi.org/10.1023/A:1003818727631

Annicchiarico, P., Barrett, B., Brummer, E. C., Julier, B., & Marshall, A. H. (2015). Achievements and challenges in improving temperate perennial forage legumes. *Critical Reviews in Plant Science*, *34*, 327–380. https://doi.org/10.1080/07352689.2014.898462

Annicchiarico, P., & Pagnotta, M. A. (2012). Agronomic value and adaptation across climatically contrasting environments of Italian red clover landraces and natural populations. *Grass and Forage Science*, *67*, 597–605. https://doi.org/10.1111/j.1365-2494.2012.00887.x

Annichiarico, P., Collins, R. P., Fornasier, F., & Rhodes, I. (2001). Variation in cold tolerance and spring growth among Italian white clover populations. *Euphytica*, *122*, 407–416. https://doi.org/10.1023/A:1012918100232

Bélanger, G., Castonguay, Y., Bertrand, A., Dhont, C., Rochette, P., Couture, L., Drapeau, R., Mongrain, D., Chalifour, F.-P., & Michaud, R. (2006). Winter damage to perennial forage crops in eastern Canada: Causes, mitigation, and prediction. *Canadian Journal of Plant Science*, *86*, 33–47. https://doi.org/10.4141/P04-171

Bertrand, A., Bipfubusa, M., Castonguay, Y., Rocher, S., Szopinska-Morawska, A., Papadopoulos, Y., & Renaut, J. (2016). A proteome analysis of freezing tolerance in red clover (*Trifolium pratense* L.). *BMC Plant Biology*, *16*, 65. https://doi.org/10.1186/s12870-016-0751-2

Björnsson, H. (1993). Zones for performance testing of timothy *(Phleum pratense* L.) in the Nordic countries. *Acta Agriculturae Scandinavica - Section B, Soil and Plant Science*, *43*, 97–113.

Blanco-Pastor, L., Barre, P., Keep, T., Ledauphin, T., Escobar-Gutièrrez, A., Roschaski, A. M., Willner, E., Dehmer, K. J., Hegarty, M., Muylle, H., Veeckman, E., Vandepoele, K., Ruttink, T., Rolàn-Ruiz, I., Manel, S., & Sampoux, J.-P. (2020). Canonical correlations reveal adaptive loci and phenotypic responses to climate in perennial ryegrass. *Molecular Ecology Resources*, *00*, 1–22. https://doi.org/10.1111/1755-0998.13289

Boller, B., Schubiger, F. X., & Kölliker, R. (2010). Red Clover. In B. Boller, U. K. Posselt, & F. Veronesi (Eds.), *Fodder Crops and Amenity Grasses*, *Handbook of Plant Breeding* (pp. 439–455). Springer.

Brozynska, M., Furtado, A., & Henry, R. (2016). Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnology Journal*, *4*, 1070–1085. https://doi.org/10.1111/pbi.12454

Castonguay, Y., Michaud, R., Nadeau, P., & Bertrand, A. (2009). An Indoor screening method for improvement of freezing tolerance in alfalfa. *Crop Science*, *49*, 809–818. https://doi.org/10.2135/cropsci2008.09.0539

Daugstad, K. (2016). Bevaring ved bruk-nye lokalsortar i engvekstene timotei, engsvingel og raudkløver. *NIBIO Rapport*. Vol *2*, 87. (In Norwegian)

De Visser, C. L. M., Schreuder, R., & Stoddard, F. (2014). The EU's dependency on soya bean import for the animal feed industry and potential for EU produced alternatives. *OCL*, *21*, D407. https://doi.org/10.1051/ocl/2014021

Dias, P. M. B., Julier, B., Sampoux, J.-P., Barre, P., & Dall'Agnol, M., (2008). Genetic diversity in red clover (*Trifolium pratense* L.) revealed by morphological and microsatellite (SSR) markers. *Euphytica*, *160*, 189–205. https://doi.org/10.1007/s10681-007-9534-z

Endresen, D. T., Street, K., Mackay, M., Bari, A., & De Pauw, E. (2011). Predictive association between biotic stress traits and eco-geographic data for wheat and barley landraces. *Crop Science*, *51*, 2036–2055. https://doi.org/10.2135/cropsci2010.12.0717

Ergon, Å., Seddaiu, G., Korhonen, P., Virkajärvi, P., Bellocchi, G., Jørgensen, M., Østrem, L., Reheul, D., & Volaire, F. (2018). How can forage production in Nordic and Mediterranean Europe adapt to the challenges and opportunities arising from climate change? *European Journal of Agronomy*, *92*, 97–106. https://doi.org/10.1016/j.eja.2017.09.016

Ergon, Å., Skøt, L., Sæther, V. E., & Rognli, O. A. (2019). Allele frequency changes provide evidence for selection and identification of candidate loci for survival in red clover (*Trifolium pratense* L.). *Frontiers in Plant Science*, *10*, 718. https://doi.org/10.3389/fpls.2019.00718

Ergon, Å., Solem, S., Uhlen, A. K., & Bakken, A. K. (2016). Generative development in red clover in response to temperature and photoperiod. In I. Roldán-Ruiz, J. Baert, D. Reheul, (Eds.), *Breeding in a World of Scarcity*. Springer International Publishing (pp. 243–247).

Falconer, D. S. (1989). *Introduction to quantitative genetics*. Longman Sci. Tech.

Fejer, S. O. (1960). Response of some New Zealand pasture species to vernalization. *New Zealand Journal of Agricultural Research*, *3*, 656–662. https://doi.org/10.1080/00288233.1960.10427145

Greene, S. L., Gritsenko, M., & Vandemark, G. (2004). Relating morphologic and RAPD marker variation to collection site environment in wild populations of red clover (*Trifolium pratense* L.). *Genetic Resources and Crop Evolution*, *51*, 643–653. https://doi.org/10.1023/B:GRES.0000024655.48989.ab

Grime, J. P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist*, *111*, 1169–1194. https://doi.org/10.1086/283244

Gusta, L. V., O'Connor, B. J., Gao, Y.-P., & Jana, S. (2001). A re-evaluation of controlled freeze-tests and controlled environment hardening conditions to estimate the winter survival potential of hardy winter wheats. *Canadian Journal of Plant Science*, *81*, 241–246. https://doi.org/10.4141/P00-068

Harrel, F., & Doupont, C. (2020). Hmisc: Harrell Miscellaneous. R package version 4.4-1. https://CRAN.R-project.org/package=Hmisc

Helgadóttir, Á., Aavola, R., Isolahti, M., Marum, P., Persson, C., Aleliunas, A. et al. (2018). Adaptability and phenotypic stability of *Lolium perenne* L. cultivars of diverse origin grown at the margin of the species distribution. *Journal of Agronomy and Crop Science*, *204*, 493–504. https://doi.org/10.1111/jac.12273

Helgadóttir, Á., & Björnsson, H. (1994). Cooperative breeding for the northern marginal areas. In O. A. Rognli, E. Solberg, & I. Schjelderup (Eds.), *Developments in Plant Breeding* (pp. 129–141). Springer, Netherlands.

Helgadóttir Á., Frankow-Lindberg B. E., Seppänen M. M., Søegaard K. & Østrem L. (2014). European grasslands overview: Nordic region. In Hopkins et al (Eds.), *EGF at 50: The future of European grasslands. Grassland Science in Europe*. *19*, 15–28.

Helgadóttir, Á., Larsen, A., Marum, P., Fritsen, H., Lindvall, E., & Miettinen, E. (2000). Prebreeding of red clover (*Trifolium pratense* L.) for northern areas. *Acta Agriculturae Scandinavica, Section B - Soil and Plant Science*, *50*, 187–190. https://doi.org/10.1080/09064710050374368

Herrmann, D., Boller, B., Studer, B., Widmer, F., & Kölliker, R. (2008). Improving persistence in red clover: Insights from QTL analysis and comparative phenotypic evaluation. *Crop Science*, *48*, 269–277. https://doi.org/10.2135/cropsci2007.03.0143

Hijmans, R. J., Jacobs, M., Bamberg, J. B., & Spooner, D. M. (2003). Frost tolerance in wild potato species: Assessing the predictivity of taxonomic, geographic, and ecological factors. *Euphytica*, *130*, 47–59. https://doi.org/10.1023/A:1022344327669

Hulke, B. S., Watkins, E., Wyse, D. L., & Ehlke, N. J. (2008). Freezing tolerance of selected perennial ryegrass (*Lolium perenne* L.) accessions and its association with field winterhardiness and turf traits. *Euphytica*, *163*, 131–141. https://doi.org/10.1007/s10681-007-9631-z

Jahufer, M. Z. Z., & Luo, D. (2018). DeltaGen: A comprehensive decision support tool for plant breeders. *Crop Science*, *58*, 1118–1131. https://doi.org/10.2135/cropsci2017.07.0456

Jensen, E. S., Peoples, M. B., Boddey, R. M., Gresshoff, P. M., Hauggaard-Nielsen, H., J.R. Alves, B., & Morrison, M. J., (2012). Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. *Agronomy for Sustainable Development*, *32*, 329–364. https://doi.org/10.1007/s13593-011-0056-7

Khazaei, H., Street, K., Bari, A., Mackay, M., & Stoddard, F. L. (2013). The FIGS (Focused Identification of Germplasm Strategy) approach identifies traits related to drought adaptation in *Vicia faba* genetic resources. *PLoS One 8*. e63107.

Kjærgaard, T. (2003). A plant that changed the world: The rise and fall of clover 1000–2000. *Landscape Research*, *28*, 41–49. https://doi.org/10.1080/01426390306531

Kölliker, R., Herrmann, D., Boller, B., & Widmer, F. (2003). Swiss Mattenklee landraces, a distinct and diverse genetic resource of red clover (*Trifolium pratense* L.). *Theoretical and Applied Genetics*, *107*, 306–315. https://doi.org/10.1007/s00122-003-1248-6

Kouamé, C. N., & Quesenberry, K. H. (1993). Cluster analysis of a world collection of red clover germplasm. *Genetic Resources and Crop Evolution*, *40*, 39–47. https://doi.org/10.1007/BF00053463

Lunnan, T. (1989). Effects of photoperiod, temperature and vernalization on flowering and growth in high latitude populations of red clover. *Norwegian Journal of Agricultural Sciences*, *3*, 201–210.

Lüscher, A., Mueller-Harvey, I., Soussana, J. F., Rees, R. M., & Peyraud, J. L. (2014). Potential of legume-based grassland–livestock systems in Europe: A review. *Grass and Forage Science*, *69*, 206–228. https://doi.org/10.1111/gfs.12124

Mossberg, B., & Stenberg, L. (2012). *Gyldendals store nordiske flora*. Gyldendal Norsk Forlag AS. ISBN 978-82-05-42485-2. (In Norwegian)

Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. R. H., Szoecs, E., & Wagner, H. (2019). vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

Olsson, G. (1997). *Den svenska växtförädlingens historia: jordbruksväxternas utveckling sedan 1880-talet*. Kungl. Skogs- och Lantbruksakad. (In Swedish)

Pagnotta, M. A., Annicchiarico, P., Farina, A., & Proietti, S. (2011). Characterizing the molecular and morphophysiological diversity of Italian red clover. *Euphytica*, *179*, 393–404. https://doi.org/10.1007/s10681-010-0333-6

Rapacz, M., Ergon, Å., Höglind, M., Jørgensen, M., Jurczyk, B., Østrem, L., Rognli, O. A., & Tronsmo, A. M. (2014). Overwintering of herbaceous plants in a changing climate. Still more questions than answers. *Plant Science*, *225*, 34–44. https://doi.org/10.1016/j.plantsci.2014.05.009

Reckling, M., Bergkvist, G., Watson, C. A., Stoddard, F. L., Zander, P. M., Walker, R. L., Pristeri, A., Toncea, I., & Bachinger, J. (2016). Trade-offs between economic and environmental impacts of introducing legumes into cropping systems. *Frontiers in Plant Science*, *7*, 669. https://doi.org/10.3389/fpls.2016.00669

Riday, H. (2010). Progress made in improving red clover (*Trifolium pratense* L.) through breeding. *International Journal of Plant Breeding*, *1*, 22–29.

Rognli, O. A., Fjellheim, S., Pecetti, L., & Boller, B. (2013). Semi-natural grasslands as a source of genetic diversity. The role of grasslands in a green future: Threats and perspectives in less favoured areas. In Helgadóttir et al (Eds.), *The Role of Grasslands in a green future. Threats and perspectives in less favoured areas. Grassland Science in Europe*. *19*, 303–313.

Rosso, B. S., & Pagano, E. M. (2005). Evaluation of introduced and naturalized populations of red clover (*Trifolium pratense* L.) at Pergamino EEA-INTA, Argentina. *Genetic Resources and Crop Evolution*, *52*, 507–511. https://doi.org/10.1007/s10722-005-0777-z

Solberg, S. O., Yndgaard, F., & Palmé, A. (2015). Morphological and phenological consequences of *ex situ* conservation of natural populations of red clover (*Trifolium pratense* L.). *Plant Genetic Resources*, *15*, 97–108. https://doi.org/10.1017/S1479262115000416

Statistics Sweden. Yearbook of agricultural statistics 2013 (Chapter 3). https://djur.jordbruksverket.se/statistik/statistikomr/jordbrukasstatistisksammanstallning/jordbruksstatistiskarsbok2013.4.765a35dc13f7d0bf7c43ab.html (In Swedish)

Steinshamn H., Nesheim L., Bakken A. K. (2016). Grassland production in Norway. In Höglind M. et al (Eds.), *The multiple roles of grassland in the European bioeconomy. Grassland Science in Europe. 21*, 811–813.

Tucak, M., Popović, S., Čupić, T., Španić, V., & Meglič, V. (2013). Variation in yield, forage quality and morphological traits of red clover (*Trifolium pratense* L.) breeding populations and cultivars. *Zemdirbyste-Agriculture*, *100*, 63–70. https://doi.org/10.13080/z-a.2013.100.009

Valle, O. (1958). Experiences with Canadian Altaswede and Swedish commercial red clover in Finland. *Agricultural and Food Science*, *30*, 293–299. https://doi.org/10.23986/afsci.71463

Van Dobben, W. H. (1964). Influence of photoperiod and temperature on the flowering of red clover. Instituut voor Biologish en Scheikundig

Onderzoek van Landbouwgewassen, Wageningen. Medeling 241. Jaarboek 1964, 77–85.

Vestad, R. (1990). Rødkløver i norsk engdyrking. Fortid og framtid. In O. A. Rognli (Ed.), *Norsk planteforedling i nåtid og fortid, Norsk Landbruksforsking (ISSN 0801–5333), Suppl. No. 9* (pp. 165–172). (In Norwegian)

Vicario, A., Mazón, L. I., Aguirre, A., Estomba, A., & Lostao, C. (1989). Relationships between environmental factors and morph polymorphism in Cepaea nemoralis, using canonical correlation analysis. *Genome*, *32*, 908–912. https://doi.org/10.1139/g89-528

Voisin, A.-S., Guéguen, J., Huyghe, C., Jeuffroy, M.-H., Magrini, M.-B., Meynard, J.-M., Mougel, C., Pellerin, S., & Pelzer, E. (2014). Legumes for feed, food, biomaterials and bioenergy in Europe: A review. *Agronomy for a Sustainable Development*, *34*, 361–380. https://doi.org/10.1007/s13593-013-0189-y

Waldron, B. L., Ehlke, N. J., Vellekson, D. J., & White, D. B. (1998). Controlled freezing as an indirect selection method for field winter-hardiness in turf-type perennial ryegrass. *Crop Science*, *38*, 811–816. https://doi.org/10.2135/cropsci1998.0011183X003800030032x

Waleen, W. M., Tanino, K. K., Olsen, J. E., Eltun, R., Rognli, O. A., & Gusta, L. V. (2011). Freezing tolerance of winter canola cultivars is best revealed by a prolonged freeze test. *Crop Science*, *51*, 1989–1996. https://doi.org/10.2135/cropsci2011.02.0098

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Zanotto S, Palmé A, Helgadóttir Á, et al. Trait characterization of genetic resources reveals useful variation for the improvement of cultivated Nordic red clover. *J Agro Crop Sci.* 2021;207:492–503. https://doi.org/10.1111/jac.12487

SUPPLEMENTARY MATERIAL

Supplementary Table 1. Information about the 54 Nordic red clover accessions used in the study. Entry 1 to 48 were included in all the analyses in this study, entry 49-51 were included in all analyses except the canonical correlation analysis, and entry 52-54 were included only in the experiment under controlled conditions. Altitude data were retrieved from the WorldClim database (Worldclim.org) using the diva-gis software (diva-gis.org) with 2.5 arcosin minutes resolution, corresponding to an area of 4.5 x 4.5 km; therefore there may be some deviation from the actual altitude at the collection site in areas with high variation in altitude within short distances.

Entry	Entry name	NordGen ID[†]	Type	Country of origin	Collection site	Latitude	Longitude	Altitude (m a.s.l.)
1	RAUHALA HM0204	NGB14425	Landrace	Finland		N 60° 50'	E 23° 50'	135
2	PEUHKURI ME0502	NGB169	Landrace	Finland	Peuhkuri, Lemi	N 61° 7'	E 27° 43'	99
3	LAHTUA HM0402	NGB14438	Landrace	Finland		N 61° 36'	E 24° 17'	123
4	HAAPALEHTO HM0102	NGB14451	Landrace	Finland		N 61° 54'	E 26° 3'	128
5	HATUNVAARA ME0601	NGB162	Landrace	Finland	Hatunvaara, Kitee	N 62° 3'	E 30° 26'	118
6	MÄNTYSUO ME0101	NGB1154	Landrace	Finland	Mäntysuo, Leppävirta	N 62° 22'	E 28° 11'	140
7	HIETAMA AP0201	NGB2844	Landrace	Finland	Hietama, Äänekoski	N 62° 35'	E 25° 36'	130
8	SONKARI ME0501 "MATTILAN KANTA"	NGB1144	Landrace	Finland	Sonkari, Vesanto	N 62° 50'	E 26° 32'	114
9	VARPULA AP0402	NGB2814	Landrace	Finland	Varpula, Kauhava	N 63° 7'	E 22° 54'	41
10	PELLONPÄÄ ME0501	NGB1145	Landrace	Finland	Pellonpää, Tervo	N 63° 57'	E 26° 49'	160
11	SAUKKAVAARA ME0201	NGB1139	Landrace	Finland	Saukkavaara, Paltamo	N 64° 25'	E 28° 11'	197
12	HYRKÄS ME0101	NGB1146	Landrace	Finland	Hyrkäs, Muhos	N 64° 48'	E 26° 7'	65
13	HAILUOTO ME0101 "HAAPANIEMI"	NGB1133	Landrace	Finland	Hailuoto, Hailuoto	N 65° 1'	E 24° 44'	11
14	AUTTI EH0203	NGB1032	Landrace	Finland	Vaulumäki, Autti	N 66° 19'	E 27° 5'	157
15	VUOTJÄRVI HM0201	NGB14442	Wild	Finland		N 63° 15'	E 28° 18'	152
16	KÖNGÄS PH0101	NGB1025	Wild	Finland	Köngäs, 28 km N Kittilä	N 67° 53'	E 24° 50'	189
17	KOLSTAD	NGB2198	Landrace	Norway		N 60° 53'	E 10° 49'	246
18	HEIGRE 09-7-62-1	NGB7116	Wild	Norway	Heigre, Sandnes	N 58° 50'	E 5° 42'	40
19	ROLIGHETEN 09-6-54-1	NGB6736	Wild	Norway	Roligheten, Skien	N 59° 13'	E 9° 40'	319

20	09-6-53-8	NGB13451	Wild	Norway	Fetveit, Vinje	N 59° 37'	E 7° 53'	325
21	Sørum 0403	NGB15623	Wild	Norway		N 60° 2'	E 11° 14'	131
22	09-6-52-4	NGB6738	Wild	Norway	Kraggeli i Lunner, Roa	N 60° 19'	E 10° 36'	313
23	09-5-43-9	NGB7117	Wild	Norway		N 63° 25'	E 11° 36'	310
24	SLOEM 09-5-42-2	NGB6742	Wild	Norway	Sloem, Bjugn	N 63° 45'	E 9° 45'	50
25	MURRUMOEN 09-5-40-4	NGB13990	Wild	Norway	Murrumoen	N 64° 28'	E 13° 26'	452
26	TRØA 09-6-48-6	NGB1575	Wild	Norway	Trøa, Atnbrua	N 61° 51'	E 10° 14'	815
27	TORRUD 09-6-48-2	NGB1572	Wild	Norway	Torrud, Alvdal	N 62° 6'	E 10° 39'	731
28	TØEN 09-6-48-1	NGB1571	Wild	Norway	Tøen, Vingelen	N 62° 25'	E 10° 52'	890
29	09-6-48-3	NGB13447	Wild	Norway	Langøien Nedre (setra), Os	N 62° 33'	E 11° 15'	792
30	LEKANGER 092156	NGB14111	Wild	Norway		N 67° 7'	E 14° 2'	39
31	KLEVIMOEN 0921915	NGB14113	Wild	Norway		N 65° 32'	E 13° 37'	256
32	Olsrud 0103	NGB15595	Semi-wild	Norway		N 60° 48'	E 11° 10'	200
33	SIDENSJÖ	NGB2487	Landrace	Sweden		N 63° 18'	E 18° 18'	126
34	ULTUNA	NGB4089	Landrace	Sweden		N 59° 55'	E 17° 29'	43
35	FRUGÅRDEN	NGB2392	Landrace	Sweden		N 58° 25'	E 12° 36'	57
36	SÖDRA SUNDERBYN PH0102	NGB1022	Landrace	Sweden	Södra Sunderbyn	N 65° 40'	E 21° 52'	22
37	STENSJÖN IB0104	NGB14193	Wild	Sweden	Karlsberg 350 m S, Stensjön, Nässjö	N 57° 33'	E 14° 47'	238
38	HAMNARYDSJÖN IB0101	NGB14186	Wild	Sweden	Hamnarydssjön E, Norra Solberga kyrka, Nässjö	N 57° 41'	E 14° 51'	265
39	NOLGÅRDEN NÄS FO0501	NGB16965	Wild	Sweden		N 58° 5'	E 13° 42'	224
40	BRÄCKETORP FO0502	NGB16982	Wild	Sweden		N 59° 3'	E 12° 29'	182
41	JUHOLA FO0102	NGB16990	Wild	Sweden		N 60° 25'	E 12° 44'	385
42	VÄSTANSJÖ GB0103	NGB1420	Wild	Sweden	Bjurholm, Västansjö	N 63° 45'	E 18° 59'	159
43	BAKSJÖNÄS GB0101	NGB1414	Wild	Sweden	Storsjön, Baksjönäs	N 64° 9'	E 17° 40'	443
44	NYGÅRD GB0201	NGB1401	Wild	Sweden	Vänjaurbäck, Nygård	N 64° 23'	E 18° 42'	357
45	RAUNAVAARA AK0402	NGB1011	Wild	Sweden	Raunavaara	N 66° 30'	E 23° 37'	163
46	ERKHEIKKI ME0102	NGB1012	Wild	Sweden	Erkheikki, Pajala	N 67° 13'	E 23° 13'	175
47	REMBACKA GB0201	NGB1406	Wild	Sweden	Vilhelmina, Rembacka	N 64° 37'	E 16° 22'	401

48	Klövsjö HAJ0402	NGB18499	Semi-wild	Sweden	Road bank, Klövsjö	N 62° 31'	E 14° 12'	491
49	Saija		Cultivar	Finland				
50	SW Yngve		Cultivar	Sweden				
51	Lea		Cultivar	Norway				
52	Linus		Cultivar	Norway				
53	Gandalf		Cultivar	Norway				
54	SW Ares		Cultivar	Norway				

†ID code in NordGen (the Nordic Genetic Resource Centre)

Supplementary Table 2. Climatic conditions during the field trials. The data were obtained from climatic stations present at each test location.

Location	Summer† 2014		Winter‡ 2014-2015					Summer 2015		Winter 2015-2016				
	Mean temp.	Total precip.	Mean temp.	Min. temp.	Days with snow cover	Number of frost days	Total precip.	Mean temp.	Total precip.	Mean temp.	Min. temp.	Days with snow cover	Number of frost days	Total precip.
Løken	12.0°C	287mm	-1.0°C	-23.8°C	NR	171	332mm	10.1°C	394mm	-2.5°C	-25.5°C	NR	180	297mm
Lännäs	13.2°C	251mm	-0.6°C	-23.1°C	111	164	341mm	11.8°C	257mm	-2.0°C	-26.9°C	114	185	220mm
Jokioinen	13.9°C	315mm	1.2°C	-23.7°C	88	124	359mm	12.8°C	305mm	0.2°C	-30.0°C	85	131	351mm
Korpa	10.4°C	494mm	1.6°C	-11.9°C	99	137	834mm	9.0°C	287mm	1.8°C	-13.0°C	67	124	775mm

† May – September, ‡ October - April

Supplementary Table 3. Description of the traits recorded in the field trials (A) and the experiment under controlled conditions (B).

Trait	Description
A	
Autumn status	General impression of plant status at the end of the establishment season (2014), scored from 1 (dead) – 9 (highest biomass)
Growth type	Visual estimate of the angle that the outer shoots make with the horizontal in early summer 2015, scored from 1 (erect) – 9 (prostrate)
Leaf area	Calculated as the product of length and width of the medial leaflet on the third leaf from the top of the longest stem, at the time when most plants had flowered in 2015 (cm^2)
Stem length	Length of the longest stem at the time when most plants had flowered in 2015 (cm)
Flowering date	Date when three heads has started getting pink in 2015
Flowering score	Evaluation of how many flowers were present in each plant at the time when most plants had flowered in 2015, scored from 1 (none) – 9 (most)
Yield score	Visual estimate of biomass accumulation at the time of the first cut in 2015, scored from 1 (low) – 9 (high)
Winter survival 2015, 2016	Ratio between the number of plants alive in the spring of 2015 or 2016 and the number of plants alive in the previous autumn
Spring status 2015, 2016	General impression of plant status two weeks after the growth started in 2015 and 2016, scored from 1 (dead) – 9 (highest biomass)
B	
Number of leaves	Number of leaves
Petiole length	Length of the longest petiole (cm)
Days to elongation	The number of days until stem elongation was observed (internode of > 2 cm) after transfer to 23 h photoperiod in a greenhouse set at 16 °C
Freezing tolerance (LT50)	The estimated temperature at which 50 % of the plants died. Note that the value of LT50 decreases with increasing freezing tolerance

Supplementary Table 4. Description of the climatic variables used in this study. Variable descriptions follow the definitions present in the WorldClim database (Worldclim.org).

Climatic Variable	Description
BIO1	Annual Mean Temperature
BIO4	Temperature Seasonality (weekly temperature standard deviation *100)
BIO5	Max Temperature of Warmest Month
BIO6	Min Temperature of Coldest Month
BIO10	Mean Temperature of Warmest Quarter†
BIO11	Mean Temperature of Coldest Quarter
BIO12	Annual Precipitation
BIO15	Precipitation Seasonality (Coefficient of Variation for precipitation)
BIO18	Precipitation of Warmest Quarter
BIO19	Precipitation of Coldest Quarter

†The definition of a quarter is any 13 consecutive weeks, (or any consecutive 3 months if running with a monthly time step). For example, the driest quarter will be the 13 consecutive weeks that are drier than any other set of 13 consecutive weeks.

Supplementary Table 5. Variance component analysis (P<0.05) for autumn status, winter survival, spring status and yield score in 2014-2015 recorded in single plant trials with 51 Nordic red clover populations at four test locations. See Supplementary Table 3 for a description of the traits; note that winter survival has a smaller scale than the other traits. Variance component analysis presents estimated population (σ_g^2), population by location (σ_{gl}^2), block within location ($\sigma_{l/b}^2$), and pooled error (σ_ε^2) variance components with associated standard errors (± SE), and estimated entry-mean broad sense heritability (h_b^2).

Source of variation	Autumn status	Spring status	Winter survival	Yield score
$\boldsymbol{\sigma_g^2}$	0.15 ± 0.05	NS[1]	NS	0.18 ± 0.07
$\boldsymbol{\sigma_{gl}^2}$	0.25 ± 0.04	0.39 ± 0.06	0.007 ± 0.001	0.45 ± 0.33
$\boldsymbol{\sigma_{l/b}^2}$	0.05 ± 0.02	0.15 ± 0.05	0.003 ± 0.001	0.09 ± 0.03
$\boldsymbol{\sigma_\varepsilon^2}$	1.13 ± 0.03	2.92 ± 0.07	0.026 ± 0.002	1.18 ± 0.05
$\boldsymbol{h_b^2}$	0.52 ± 0.08	-	-	0.45 ± 0.10

[1] not significant

Supplementary Table 6. Variance component analysis (P<0.05) for traits expressed during the growing season in single plant trials with 51 Nordic red clover populations at the two test locations that were not affected by winter mortality during the first winter (Jokioinen and Lännäs). See Supplementary Table 3 for a description of the traits; note that the traits vary in scale. Variance component analysis presents estimated population (σ_g^2), population by location (σ_{gl}^2), block within location ($\sigma_{l/b}^2$), and pooled error (σ_ε^2) variance components with associated standard errors (± SE), and estimated entry-mean broad sense heritability (h_b^2).

Source of variation	Autumn status	Spring status	Flowering date	Flowering score	Growth type	Leaf area	Stem length	Yield score
$\boldsymbol{\sigma_g^2}$	0.18 ± 0.06	0.17 ± 0.05	63.90 ± 13.43	0.74 ± 0.17	1.50 ± 0.30	1.65 ± 0.39	3.44 ± 0.70	0.78 ± 0.18
$\boldsymbol{\sigma_{gl}^2}$	0.17 ± 0.04	0.10 ± 0.03	6.82 ± 1.73	0.16 ± 0.06	NS	0.24 ± 0.12	0.07 ± 0.03	0.16 ± 0.05
$\boldsymbol{\sigma_{l/b}^2}$	0.04 ± 0.02	0.06 ± 0.03	1.05 ± 0.19	NS[1]	NS	0.30 ± 0.15	NS	0.08 ± 0.04
$\boldsymbol{\sigma_\varepsilon^2}$	1.10 ± 0.04	1.13 ± 0.04	35.50 ± 1.18	2.06 ± 0.07	1.37 ± 0.05	6.16 ± 0.20	1.62 ± 0.05	1.50 ± 0.05
$\boldsymbol{h_b^2}$	0.45 ± 0.10	0.47 ± 0.08	0.89 ± 0.02	0.69 ± 0.06	0.89 ± 0.02	0.65 ± 0.06	0.94 ± 0.01	0.75 ± 0.05

[1] not significant

Supplementary Table 7. Phenotypic correlation coefficients (*r*) between various traits from the field and controlled experiment. See Supplementary Table 3 for a description of the traits. Yield score is the mean of Yield score at Jokioinen and Lännäs were the traits were highly correlated. Population averages were used for the analysis (N=51).

	Autumn status					Winter survival					
	Løken	Lännäs	Jokioinen	Stem length	Yield Score	Løken 2015	Korpa 2015	Jokioinen 2016	Days to elongation	Number of leaves	Petiole length
Autumn status Lännäs	0.61***	-									
Autumn status Jokioinen	0.51***	0.56***	-								
Stem length	0.69***	0.71***	0.69***	-							
Yield score	0.66***	0.80***	0.70***	0.95***	-						
Winter survival Løken 2015	-0.41***	-0.41**	-0.47***	-0.41**	NS	-					
Winter survival Korpa 2015	NS	NS	NS	NS	NS	NS	-				
Winter survival Jokioinen 2016	-0.78**	-0.78***	-0.52***	-0.85***	-0.81 ***	0.48***	NS	-			
Days to elongation	-0.41**	NS	-0.55***	NS	-0.30*	0.40**	-0.30*	0.34*	-		
Number of leaves	NS	NS	NS	-0.35*	NS	NS	NS	0.38**	NS	-	
Petiole length	0.66***	0.80***	0.61***	0.87***	0.80***	-0.33*	NS	-0.70***	NS	-0.37**	-
Freezing tolerance (LT50)	0.42**	0.38**	0.36**	0.53***	0.49***	NS	NS	-0.49***	NS	NS	0.47***

*, p < .05; **, p < .01; ***, p < .001; NS, not significant

Supplementary Table 8. Variance component analysis (**A**) and genotypic correlations (**B**) for spring status and winter survival recorded after three selective winters (Jokioinen 2016, Løken and Korpa 2015). Variance component analysis shows population (σ_g^2), population by location (σ_{gl}^2), blocks within locations ($\sigma_{l/b}^2$), and pooled error (σ_ε^2) variance components, their associated standard errors (± SE), and entry-mean broad sense heritability (h_b^2) for the 51 Nordic red clover populations (48 wild/semi-wild/landrace population + 3 reference cultivars). Averages of each population, block and location were used for the analysis. Significance level for correlation coefficients and variance components was set at P < 0.05.

A

Source of variation	Spring status	Winter survival
$\boldsymbol{\sigma_g^2}$	NS[1]	NS
$\boldsymbol{\sigma_{gl}^2}$	0.70 ± 0.11	0.03 ± 0.01
$\boldsymbol{\sigma_{l/b}^2}$	0.18 ± 0.07	0.01 ± 0.01
$\boldsymbol{\sigma_\varepsilon^2}$	0.98 ± 0.86	0.05 ± 0.01
$\boldsymbol{h_b^2}$	-	0.13 ± 0.21

[1] not significant

B

Spring status	Løken 2015	Korpa 2015
Korpa 2015	-0.24	-
Jokioinen 2016	0.46	-0.28

Winter survival	Løken 2015	Korpa 2015
Korpa 2015	-0.38	1
Jokioinen 2016	0.55	-0.23

Paper II

Biochemical changes after cold acclimation in Nordic red clover (*Trifolium pratense* L.) accessions with contrasting levels of freezing tolerance

Stefano Zanotto[1*], Annick Bertrand[2], Randy W. Purves[3,4], Jorunn E. Olsen[1], Fatma M. Elessawy[4], Åshild Ergon[1]

[1]Norwegian University of Life Sciences, Faculty of Biosciences, Dept. of Plant Sciences, Ås, Norway

[2]Agriculture and Agri-Food Canada, Québec City, QB, Canada

[3]Centre for Veterinary Drug Residues, Canadian Food Inspection Agency, Saskatoon, SK, Canada

[4]College of Pharmacy and Nutrition, University of Saskatchewan, Saskatoon, SK, Canada

*** Correspondence:** Stefano Zanotto, e-mail: stefano.zanotto@nmbu.no

Keywords: amino acids, carbohydrates, persistency, phenolic compounds, abiotic stress, winter survival

Abstract

The ability to tolerate low freezing temperatures is one component of winter survival and field persistence of red clover grown in the Nordic countries. Cold acclimation (CA) is a key process allowing plants to acquire higher level of freezing tolerance. Little is known about the biochemical responses to cold and the importance of such changes in freezing tolerance in Nordic red clover. To shed light on this, five freezing tolerant (FT) and five freezing susceptible (FS) red clover accessions of Nordic origin were subjected to short-term cold acclimation under controlled experimental conditions, and the effect on the contents of carbohydrates, amino acids and phenolic compounds in the plant crowns were assessed. Starch depletion and the accumulation of soluble sugars and free amino acids were the most evident responses to cold. Sucrose, arginine, asparagine and proline had the highest absolute increase after cold acclimation. The content of six phenolic compounds increased during CA, while the content of six other phenolic compounds were higher in non-acclimated plants. Among these, L- typtophan, deoxyclovamide and dideoxyclovamide had the highest concentration; the first increased during CA while the latter two decreased. Freezing tolerant (FT) accessions had higher content of raffinose, pinitol, arginine, serine, alanine, valine, phenylalanine and one phenolic compound (a pinocembrin hexoside derivative) than freezing susceptible (FS) accessions, suggesting a role for these compounds in the freezing tolerance of red clover. FS accessions, on the other hand, had higher content of starch, glucose, fructose and three phenolic compounds.

Introduction

Red clover (*Trifolium pratense* L.) is widely cultivated in temperate regions and is the main forage legume grown in Northern Europe, where it is mainly used in mixtures with grasses to produce silage (Helgadóttir et al., 2014; Annicchiarico et al. 2015). Red clover is a protein-rich feed, particularly suited for ensiling and use in ruminants' rations because of its high polyphenol oxidase (PPO) activity, which reduces nitrogen losses (Lee, 2014; Hart et al., 2016). Legumes can, in symbiosis with rhizobia, fix atmospheric nitrogen, resulting in decreased application of nitrogen fertilizers and more sustainable agricultural systems (Jensen et al., 2012; Lüscher et al., 2014). The adoption of forage legumes in crop rotations is particularly important in regions characterized by cold climates, such as parts of Northern Europe and Canada where the cultivation of grain legumes may be hampered by low temperatures and short growing seasons. In such areas, the utilization of red clover may be limited due to its poor persistence, mainly because of the high mortality during winter (Abberton and Marshall, 2005). Winter survival is a complex trait affected by many biotic and abiotic factors (Bélanger et al. 2006). Freezing temperature is a main cause of winter mortality at continental locations characterized by harsh winters, and the ability of plants to acquire a high level of freezing tolerance is a key trait for winter survival at these locations. Freezing tolerance is achieved during a period of cold acclimation (CA) at low above freezing temperatures, and can be improved by selection (Bertrand et al., 2016). In Nordic local red clover accessions freezing tolerance was recently shown to be correlated with low annual temperatures at their place of origin, suggesting that the freezing tolerance was affected by selection pressure for this trait in that plant material (Zanotto et al., 2021).

CA in winter annual and perennial herbaceous plants adapted to a temperate climate is associated with huge changes at the transcriptional, proteomic and metabolomic levels (Theocharis et al., 2012; John et al., 2016; Fürtauer et al., 2019). Metabolic changes during CA (as well as acclimation to other types of abiotic stress) typically include an increase in the content (on a dry matter basis) of soluble sugars, sugar alcohols, organic acids, amino acids and amino acid derivatives, and other substrates for secondary metabolites. Soluble carbohydrates are thought to have multiple functions in freezing tolerance, such as stabilization of membranes and proteins, lowering of the freezing temperature within the cell, and quenching or scavenging of reactive oxygen species (Pommerrenig et al., 2018). Amino acids such as proline and glycine betaine have been shown to accumulate under stress in a wide range of species (Giri, 2011; Szabados and Savouré, 2010). In addition to functioning as osmolytes, glycine betaine protects photosystem II under stress, while proline stabilizes membranes, protects against oxidative stress and acts as a signaling molecule (Szabados and Savouré, 2009). The amino acid arginine is a precursor in the

synthesis of polyamines, which play roles in abiotic and biotic stress resistance by stabilizing membranes, acting as scavengers of reactive oxygen species, and by participating in complex signaling networks that regulate stress responses (Alcázar et al., 2011; Pál et al., 2015 Menéndez et al., 2019). Phenolic compounds, including flavonoids, play important roles in the defense towards abiotic and biotic stress and have been shown to act as antioxidants (Agati et al. 2012). A study of mutant lines affected in the first steps of flavonoid biosynthesis showed that flavonoids are important for freezing tolerance in *Arabidopsis thaliana* (Schultz et al. 2016). Red clover is rich in phenolic compounds, in particular isoflavones, which have been shown to be involved in oxidative stress response in both roots and leaves (Saviranta et al., 2009; Saviranta et al., 2010).

Proteome analyses of red clover crowns showed that CA induced marked changes in protein abundance. Proteins involved in carbohydrate metabolism, energy production and amino acid metabolism were upregulated in response to cold acclimation but varied in abundance between different genetic backgrounds (Bertrand et al. 2016). Biochemical analyses of populations recurrently selected for freezing tolerance showed that the accumulation of soluble sugars and of specific amino acids in plant crowns after a prolonged period of CA are associated with the acquisition of freezing tolerance in red clover and alfalfa (Castonguay et al., 2011; Bertrand et al., 2020). Some responses seem to be more general as they were similar for all genetic backgrounds within the two species. Particularly sucrose increased markedly during cold acclimation and reached higher concentrations in populations with higher levels of freezing tolerance in most genetic backgrounds of both species. Among the amino acids, the content of proline, arginine and asparagine showed the largest increase in response to CA, but the effect of selection for improved FT on the relative amounts of these amino acids differed between the genetic backgrounds. In addition, the identification of up- and down-regulated proteins involved in the biosynthesis of phenolic compounds after CA (Bertrand et al., 2016), such as the isoflavone-reductase family proteins, suggests that phenolic compounds may play a role in the acquisition of higher freezing tolerance in red clover.

The aims of this work were to characterize biochemical changes induced by short-term CA in red clover accessions of Nordic origin and investigate if the CA response at the biochemical level differs between accessions with contrasting levels of freezing tolerance. We studied plant crowns (transition between shoot and root) as this is the most critical tissue for winter survival, giving rise to shoot regrowth after winter. We hypothesized that: 1) significant changes in content of specific carbohydrates, amino acids and phenolic compounds occur in red clover crowns during short-term CA; and 2) red clover accessions with different levels of freezing tolerance differ in these responses to short-term CA.

Material and methods

Plant materials, growing conditions, and experimental treatments

Ten red clover accessions of Nordic origin were selected from the NordGen (the Nordic Genetic Resource Centre) collection; five freezing tolerant (FT) accessions and five freezing susceptible (FS) accessions. These accessions had the most extreme freezing tolerance (high or low), assessed as LT50 (temperature required to kill 50% of the plants), in our previous screening of 48 red clover gene bank accessions of Nordic origin (Zanotto et al.2021). In the same study we found that the level of freezing tolerance was significantly correlated with the winter survival recorded in the field after a winter characterized by severe temperatures. Information regarding the origin of the accessions, their freezing tolerance and winter survival are provided in Table 1. FT and FS accessions were strongly significantly different in their level of freezing tolerance and winter survival according to analyses of variance ($p<0.05$).

Seeds were scarified with sandpaper, sown in a peat soil ("Gartnerjord", pH 5.5-6.5, L.O.G. AS, Oslo, Norway) in September 2019. After germination, seedlings (3 per pot) were transplanted into pots (420 cm3; 7.5 cm wide, 7.5 cm height) with the same peat soil. The plants were grown in a greenhouse at Ås, Norway (59°40' N, 10°47' E) at 16 °C, supplied with light from metal halide lamps (HQI lamps, OSRAM GmbH, Ausburg, Germany) providing a photosynthetic photon flux density (PPFD) of approximately 100-125 µmol m^{-2} s^{-1} in the top of the canopy under a photoperiod of 12 h. When the plants were eight weeks old, half of them were moved to cold acclimating conditions in a separate cold room at 3-4 °C, 12 h photoperiod and a PPFD of 110 µmol m^{-2} s^{-1} from metal halide lamps (HQI lamps, OSRAM) without natural light, for three weeks (cold acclimation treatment, hereafter referred to as CA), while the others continued to grow in the same conditions as described above for two more weeks (no acclimation treatment, hereafter referred to as NA). The experiment had a complete randomized design with four replicated blocks. Within each block there were 10 pots per accession with three plants per pot, giving a total of 2400 plants (30 plants x 10 accessions x 4 blocks x 2 treatments).

Sampling of crown tissue

Plant crowns, a 2-cm transition zone between shoots and roots, were separated from the roots and shoots after washing, and flash-frozen in liquid nitrogen. Crowns of the 30 plants of each accession, were pooled after harvesting, resulting in a total of 80 samples (10 accessions x 4 blocks x 2 treatments). Samples were freeze-dried with a LYOQUEST -55 freeze dryer (Telstar, Madrid, Spain) and milled to a fine powder with a Cyclotec 1093 sample mill (Foss A/S, Hillerød, Denmark) with a 1-mm sieve. The samples were stored at -20 °C until the analyses described below.

Determination of carbohydrate and amino acid content and composition

Soluble sugars and amino acids were extracted from 0.2 g freeze-dried crown tissue per sample with 7 mL of methanol : chloroform : water (MCE, 12:5:3, *v*:*v*:*v*). The extracts were heated for 20 min at 65 °C to stop enzymatic activity, cooled on ice, and left overnight at 4°C for optimal extraction. After a 10-min centrifugation at 4°C and 1500 x g, 1-mL of the supernatant was collected, and 0.25 mL of water was added to the tubes. After homogenization on a vortex mixer, samples were left to rest for 10 min and were subsequently centrifuged at 13 000 x *g* for 3 min for complete phase separation. A 0.7-mL subsample of the aqueous phase was collected and evaporated to dryness overnight using a centrifugal evaporator (Savant Speedvac plus SC210A, Thermo Fisher Scientific). Pellets were dissolved in water (0.7 mL), homogenized on a vortex mixer and centrifuged for 3 min at 13 000 x g prior to HPLC analyses.

Soluble sugars were separated and quantified on a Waters HPLC analytical system controlled by the Empower II software (Waters, Milford, MA, USA) as described in Bertrand et al. (2020). An HPX-87P column (Bio-Rad, Hercules, CA, USA) was used for sugar separation at 80 °C at a flow rate of 0.5 mL min^{-1} with water as the mobile phase and detected on a Refractive Index Detector (Waters, Model 2414). Peak identities and quantities of raffinose, stachyose, sucrose, pinitol, glucose and fructose were determined by comparison to standards. The quantity of total soluble sugars (Tot.SS) was calculated as the sum of raffinose, stachyose, sucrose, pinitol, glucose and fructose. Starch was hydrolyzed by adding 3 mL of digestion buffer (200 mM sodium acetate, pH 4.5) with amyloglucosidase (AGS, 15 U mL^{-1}, Sigma-Aldrich) into the tubes containing the remaining supernatant and pellet after extraction of soluble carbohydrates. Potato starch (Sigma-Aldrich S2630) quantities of 0, 2.5, 5.0, 7.5, and 10 mg were included in the hydrolysis batch to establish the linear regression between the quantity of glucose yielding from starch hydrolysis (R^2 =0.9997) of known concentrations were included in the hydrolysis batch to establish the linear regression between glucose and starch concentrations. The difference in glucose concentration obtained after and before AGS hydrolysis was fitted into the regression equation to calculate the starch content of each sample expressed on a 105°C DM basis (mg g^{-1} DM).

Twenty-one amino acids (alanine, arginine, asparagine, aspartate, glutamate, glutamine, glycine, γ-aminobutyric acid (GABA), α- aminobutyric acid (AABA), histidine, proline, methionine, lysine, serine, leucine, isoleucine, ornithine, phenylalanine, threonine, tyrosine and valine) were separated and quantified using Waters ACQUITY UPLC analytical system controlled by the Empower II software (Waters). Amino acid derivatization was performed using AccQ Tag Ultra reagent (6-aminoquinolyl-N-hydroxysuccinimidyl carbamate, Cohen 2003). Derivatives were separated on an AccQ Tag Ultra column (2.1 × 100 mm; Waters) and detected with Waters ACQUITY Tunable UV detector at 260 nm. Peak identity and amino acid quantity were determined

by comparison to a standard mix containing the 21 amino acids with a range of detection from 2.5 to 40 pmol/µL and a R^2 of 0.999. The content of total amino acids (Tot.AA) was calculated as the sum of the 21 amino acids.

Determination of polyphenol content and composition

Red clover crown extracts were prepared using a procedure developed by Mirali et al. (2014) and modified by Elessawy et al. (Elessawy et al., 2020). However, since this is an untargeted study (these previous studies were targeted) an important modification was that internal standards were not used. Instead, as is common for untargeted analyses, a quality control (QC) sample, which is a mixture containing an equal portion (15 µL) of each sample, was used for retention time alignment and to enable relative quantification of the metabolites. A volume of 1 mL of the acetone : water (70 : 30 v/v) extraction solvent was added to 0.2 g of freeze-dried, powdered crown tissue (per sample) and samples were subsequently shaken for 1 h at 23°C on a Thermomixer C (Eppendorf, Germany) at a speed of 3.3 x g. The samples were centrifuged with a Fresco 21 centrifuge (Thermo Fisher, USA) at a speed of 16 200 x g for 15 min, and each supernatant was transferred into a new-labelled tube that was centrifuged at 16 200 x g for 5 min a second time to ensure removal of all pellet material. A 200 µL aliquot of each extract was dried down with a Speed Vac SPD111V (Thermo Fisher, USA) and the dried samples were kept at -20 °C until analysis. Blank samples were prepared in a similar way using empty Sarstedt 2 mL microtubes. Dried samples were reconstituted in 1000 µL of MilliQ-water : methanol (90 : 10 v/v) just prior to analysis.

The untargeted workflow used liquid chromatography (LC) and high-resolution mass spectrometry (HRMS). The LC system was a Dionex 3000 LC equipped with a Diode Array Detector (DAD) that was placed in-line between the column outlet and the inlet to the ion source of the high resolution Thermo Fisher Q-Exactive mass spectrometer. For LC separation, a Waters Acquity HSST3 column (2.1 x 50 mm, 1.8 µm) was used; the temperature of the column chamber was set to 30 °C and the liquid flow rate was 0.35 mL min^{-1}. The run time was 35 min with an injection volume of 4 µL, the gradient is given in Supplementary Table 1. Eluent exiting the column passed through the DAD, which acquired absorbance data from 200-600 nm at 5 Hz, before being introduced to the HESI (heated ESI) source of the mass spectrometer. The Q-Exactive was used to acquire full scan data for the crown samples using a mass resolution (full width at half maximum, FWHM, @ m/z 200) of 140,000 in positive mode and a mass range of 120-1800 m/z. A QC (quality control) sample and four ID (identification samples) were also prepared. The QC sample contained equal amounts of all 78 red clover samples, whereas each ID sample contained equal amounts of samples within the groups NA-FT (ID1), NA-FS (ID2), CA-FT (D3), and CA-FS (ID4). The QC sample was injected every 9-10 runs to account for changes in retention time and/or signal intensity,

thereby enabling relative quantification of the polyphenols. The ID samples were used to obtain fragmentation (MS/MS) spectra using the scan function “Full scan/DDMS2”. DDMS2 (data dependent MS/MS) acquired fragmentation data on the most abundant ions detected in full-scan mode. Mass resolution of the full scan analysis was 140,000 (FWHM @ m/z 200) and MS/MS was carried out on the top (most abundant) 7 peaks at a resolution of 17 500 (FWHM @ m/z 200) from each scan using a stepped collision energy fragmentation (15, 35, and 55 eV). The MS/MS acquisition used an exclusion list (m/z values) of the most intense ions detected from the blank sample. Additional MS/MS spectra were also acquired using stepped collision energies of 10/15 eV and 70/90 eV to assist in compound identification.

Data analyses

Analyses of carbohydrates and amino acids

The contents of carbohydrates and amino acids were subjected to analyses of variance (ANOVA), performed with treatment (CA vs NA), accession group (FT vs FS) and the interaction between them as factors in the analyses. The aov function of the package Stats (R Core Team, 2021) was used for the analyses after the homogeneity of variance was assessed with Levene's test and the normality of the data with the Shapiro-Wilk test, respectively. A Multi Factor Analysis (MFA) was performed with the package FactorMineR (Le et al. 2008) to summarize the contribution of the amino acid and carbohydrate contents to the biochemical changes induced by the acclimation treatment in red clover crowns. This multivariate analysis allows for simultaneous analysis of quantitative and qualitative variables accounting for a group structure in the data. In this case the analysis was used to summarize the contribution of the amino acid and carbohydrate contents to the biochemical changes induced by the acclimation treatment in red clover crowns. The contents of amino acids and carbohydrates were used as active quantitative variables and classified in two different groups, while acclimation treatment and accession group were used as categorical variables (active and supplementary respectively). Visualization of the results of the MFA was obtained with the R package factoextra (Kassambara and Mundt, 2020).

Analyses of polyphenols

A customized untargeted workflow was developed by adapting an existing workflow in the Compound Discoverer 3.1 software (Thermo Fisher) to process LC-MS raw data. The workflow is similar to one reported previously with some modifications (Elessawy et al., 2021). The Compound Discoverer 3.1 parameters used in generating the analysis are incorporated in a “Summaries” window by the software. Tabs within this summary include “Workflow”, “Study”, “Grouping & Ratio”, and “Filters”. Note that separate analyses were carried out for comparing FT vs. FS and NA

vs. CA, but only the "Grouping & Ratio" tab changed between these analyses. Thus, in the supplementary information there are two different summaries for "Grouping & Ratio", namely one for NA/CA and the other for FT/FS. The Compound Discoverer workflow uses the full-scan accurate mass data to determine the possible molecular formula for each m/z value and MS/MS spectra from ID spectra to help identify compounds. In addition to using Thermo's mzCloud library, which contains fragmentation data of over 19 000 compounds analyzed with Thermo Orbitrap instrumentation (www.mzcloud.org), the MS/MS spectra were also compared (using the m/z vault node) to an in-house library at the Core Mass Spectrometry Facility (UofS, Canada). Fragmentation spectra from several other libraries were also used offline, including libraries available in public databases, such as FoodB (foodB.ca) and the human metabolome database (hmdb.ca). The primary focus included the differences between the CA and NA treatment and between the FT and FS accession groups. Two samples did not yield enough tissue for all the analyses, therefore only 78 samples were analyzed for their phenolic profile.

Note that the identification levels of the identified polyphenols follow that used in the work by Elessawy et al. (2021), which was derived from Sumner et al., with the addition of level 2/3 indicating isomeric compounds, including cis/trans (Sumner et al., 2007). For simplicity, we used similar nomenclature for identification as Saviranta et al. (Saviranta et al., 2010a; Saviranta et al., 2010b), that is, A=aglycone, G=glucoside or galactoside, M=malonate. We did not have authentic standards for any of the compounds and therefore, no compound identities were confirmed, although several assignments are putative or isomeric. Several of the peaks show characteristic losses of m/z 162 (-G), 248 (-G-M), and 334 (-G-M-M) in their MS/MS spectra leaving the aglycone. In many cases, comparison of fragmentation of the aglycone to MS/MS libraries was pivotal in identification.

Results

Carbohydrate and amino acid content

The first two dimensions of the MFA accounted for 73 % and 15 % of the variation in the data set, respectively (Figure 1A, Supplementary Figure 1). The first dimension was dominated by the treatment effect (CA or NA), while the second captured the variation due to the accession group (FT or FS). Crown samples from CA plants had a high score for the first dimension of the MFA, while NA plants scored low (right panel of Figure 1A). The content of most of the analyzed carbohydrates and amino acids was positively correlated with the first MFA dimension (Figure 1B and Supplementary Table 2), but the content of tyrosine, GABA and starch was negatively correlated to it (Figure 1C and Supplementary Table 2). Accessions belonging to the FS group had

higher score for the second dimension of the MFA than the FT group (Figure 1A). The second MFA dimension was also partially associated to the accumulation of most carbohydrates (Figure 1B). Fructose and glucose content was highly significantly correlated to the second dimension, while tyrosine content was negatively correlated to it (Figure 1C and Supplementary Table 2).

The content of all the analyzed carbohydrates and amino acids, except glucose, were clearly affected by the CA treatment (Figure 2,3 and Supplementary Table 2,3). Most of the metabolites were found in higher content in the CA samples, and only starch, GABA and tyrosine were less concentrated in the CA than in the NA samples. Among the soluble carbohydrates, sucrose had the strongest absolute increase (in mg g^{-1} DW) in response to cold acclimation. However, raffinose showed the highest fold change with 12 times increased content in the CA samples as compared to the NA samples, where this sugar was almost absent. Stachyose remained at a low undetectable level throughout the experiment. Among the amino acids, proline, arginine and asparagine had the strongest increase in response to CA and ended up as the most abundant amino acids in CA crowns. Proline and arginine accounted for around 40% of total amino acid content in CA crowns while they represented only 10% of the amino acids in NA crowns.

Interestingly, accession group also had a significant effect on the content of soluble sugar and only sucrose did not differ significantly between the FT and FS accession groups. There was a significantly higher content of starch, total soluble sugars, fructose and glucose in the FS group, while there was a higher content of pinitol and raffinose in the FT group. The largest differences were seen for fructose, which had 80% higher concentration in the FS group, and raffinose, which had a 50% higher concentration in the FT group. The effect of accession group was weaker for the amino acids. However, serine, arginine, alanine, GABA, tyrosine, valine and phenylalanine differed between accession groups, all being significantly more concentrated in the FT than in the FS group. Phenylalanine and tyrosine showed the highest difference of about 40% while the other amino acids generally differed by about 10%. The interaction effect between treatment and accession group was significant only for raffinose and fructose among the carbohydrates and for valine among the amino acids. The increase in the content of raffinose and valine after cold acclimation were more pronounced in the FT group with 7% and 13% higher content in FT compared to FS, respectively, while that of fructose was 14% higher in the FS group than in the FT group. A summary of the differences in carbohydrate and amino acid contents between freezing susceptible (FS) and freezing tolerant (FT) accessions before (NA) and after (CA) 3 weeks cold acclimation treatment is shown in Figure 4.

Polyphenol content

The 31 highest peak intensities (measured as area under the curve) for the phenolic compounds detected by LC-MS analyses across all the ten Nordic red clover accessions in this study are shown in Figure 4. A detailed description of the identification of compounds is presented in Supplementary file 1. All these compounds were clearly found in all the samples, therefore the spectra of a randomly chosen accession is shown. The identities and details of the respective compounds are provided in Table 1, while the MS/MS spectra for the compounds are presented in Supplementary Figure 2. The peaks with the highest intensities were identified as pseudobaptigenin glycoside malonate (G-M) (peak 23), formononetin-G-M (peak 24) and maackiain-G-M (peak 26).

Twelve and four peaks had significantly different intensities between acclimation treatments and accession groups, respectively (Figure 5). The MS/MS spectra and the details of the peak identities are presented in Supplementary Figure 3 and Table 2. Only three of these, peak 1 (L- typtophan), 4 (deoxyclovamide) and 7 (dideoxyclovamide), were among the high intensity peaks shown in Figure 4; the first was more intense in the CA samples and the latter two more intense in the NA samples. The other 13 compounds were detected only at relatively low intensities and are therefore not distinguishable in the chromatograms presented in Figure 4. Among the compounds that distinguished FT from FS accessions, peak 37 (pinocembrin isomer-G derivative) was found with higher peak intensity in FT, while peaks 34, 35 and 36 (dihydrokaempferol, biochanin A-G-M-M and a phloretin isomer) had higher peak intensities in FS accessions.

Discussion

Changes in carbohydrate and amino acid content in response to cold acclimation

The sharp increase in soluble sugars, mainly sucrose, after short-term CA is likely due to the breakdown of starch, which declined. Starch is the primary carbohydrate reserve in forage legumes and its breakdown that results in accumulation of soluble sugars in plant cells is a well-known mechanism involved in the acquisition of freezing tolerance during CA (Bélanger et al., 2006). In particular, cold stress is known to trigger starch degradation through increased expression of β-amylase (Dong et al. 2019). The increase in total soluble sugars in mg per g DW was higher than the corresponding decrease in starch. It is likely that during CA, plant crowns function as sinks, accumulating sugars photosynthesized in other tissues, in addition to those arising from starch breakdown. Among the different soluble sugars that we analyzed, sucrose accounted for the highest amount. Sucrose accumulation in CA plants could be due to a reduced catabolism in CA plants as emphasized in a proteome analyses of red clover crowns (Bertrand et al. 2016), showing that the abundance of the sucrose-cleaving enzyme sucrose synthase decreased in response to CA.

Unexpectedly, the content of starch and total soluble sugars was higher in the FS than in the FT accession groups. This is not consistent with previous findings by Bertrand et al. (2020), who studied Canadian red clover cultivars after long-term cold acclimation under semi-natural conditions. In that study, subpopulations recurrently selected for improved freezing tolerance had higher starch level in the autumn before CA and accumulated higher total soluble sugar and sucrose levels during CA, than the non-selected cultivars. Results with alfalfa varied among studies and genetic backgrounds of the material. While Castonguay et al. (1995) found no association between sucrose accumulation and freezing tolerance of cultivars, a later study found that two non-selected alfalfa cultivars accumulated less sucrose during CA than populations obtained after several cycles of recurrent selection for improved FT also showing a marked enhancement of FT (4-5°C) (Castonguay et al. 2011). The lack of a significant difference between the FT and FS accession groups in our study suggests that starch, total soluble sugars and sucrose contents after short-term CA are not related to the difference in freezing tolerance in the accessions under study. The strong association between freezing susceptibility and a high content of glucose and fructose, suggests that these monosaccharides are not linked with higher levels of freezing tolerance, and that they are probably used as building blocks for the production of other valuable carbohydrates by the FT accessions, of which some are discussed below.

The raffinose and pinitol contents increased during CA, were higher in the FT than in the FS group and increased more sharply in FT than in FS (significant for raffinose only), suggesting that these carbohydrates have a role to play in the freezing tolerance of red clover. Bertrand et al. (2020) reported an increased content of pinitol with CA which was maintained during the whole winter and proposed that pinitol enhances freezing tolerance in red clover through its osmoprotective function. The trisaccharide raffinose is the smallest molecule of the raffinose family oligosaccharides (RFOs), which are α-1,6-galactosyl extensions of sucrose (Sengupta et al. 2015), produced by adding galactinol molecules to an initial molecule of sucrose. RFOs are thought to have several functions such as carbohydrate storage and transport, signaling and abiotic stress response (Van den Ende 2013; Sengupta et al. 2015; Vinson et al. 2020), including protection against chilling stress (Nishizawa et al. 2008). Bertrand et al. (2020) reported an increase in raffinose content during the CA of red clover but reported that the positive effect of selection for freezing tolerance on the accumulation of raffinose after CA was only significant in one of the two genetic backgrounds under study. On the other hand, Castonguay et al. (1995) found that differences in the maximum level of freezing tolerance between non-hardy and winter-hardy cultivars of alfalfa was more related to the capacity of the plants to accumulate the RFOs stachyose and raffinose, than sucrose. Bertrand et al. (2016) did not find genes specifically involved in the biosynthesis of raffinose to be up- or down-regulated during CA in their proteome analyses of red

clover crowns. However, in alfalfa crowns, a several-fold increase of the expression level of galactinol synthase (GaS) was reported after exposure to low temperature (Bertrand et al., 2017). Galactinol is involved in the synthesis of raffinose from sucrose (Sengupta et al. 2015). Taken together, these reports hint at a role for RFOs in the freezing tolerance of red clover. It seems, however, that this legume species has a limited capacity to accumulate raffinose in some genetic backgrounds, and an inability to accumulate stachyose. This may possibly explain its lower capacity to reach high levels of freezing tolerance as compared to alfalfa.

The increased accumulation of free amino acids during CA that we observed is a well-known stress response in plants. Arginine, asparagine and proline, which had the highest increase after CA, were all previously found to be involved in the cold stress response in red clover and alfalfa (Castonguay et al. 2011, Bertrand et al. 2020). However, among these, only arginine was significantly more concentrated in the FT than the FS group, suggesting that arginine may play a role in determining the difference in freezing tolerance between the two groups. Arginine is a key storage compound in higher plants and the close relationship between its abundance in taproots of alfalfa during overwintering and shoot regrowth in the spring suggest that this amino acid may play a determinant role as a N reserve in perennial plants (Dhont et al. 2003). Arginine is also a precursor of polyamines, which in turn can be catabolized into GABA. Both polyamines and GABA are involved in stress response and tolerance in legumes (Menéndez et al., 2019), for example, polyamines are involved in freezing tolerance in M. falcata (Zhou et al. 2018) and drought tolerance in T. repens (Yong et al., 2017; Li et al., 2018). In our study we found a decreased content of GABA after CA. In contrast, Bertrand et al. (2020) found that during winter GABA was at its highest level in December and as such could be involved in the acquisition of maximum freezing tolerance in red clover. It may be possible that our experimental conditions with short-term acclimation treatment with above zero temperature caused a slight decrease of GABA concentration instead of an accumulation since this intermediate of polyamine synthesis had to be entirely used as building block for the synthesis of other more critical compounds. The different genetic backgrounds of Nordic European and Canadian red clover may also be a cause for the different responses in our study and that of Bertrand et al. (2020). Serine, valine, tyrosine, alanine and phenylalanine are other amino acids which had higher content in the FT than in the FS group and that could therefore be involved in the differences in freezing tolerance between the two groups. Of these, the tyrosine content decreased during CA while the others increased. The role of phenylalanine in the acquisition of freezing tolerance, as a precursor of the flavonoid pathway is discussed below.

Changes in polyphenol content in response to cold acclimation

This is to our knowledge the first work that investigated the phenolic profile in crowns of red clover, while several studies focused on roots and aerial tissue (Lin et al. 2000; Tsao et al. 2006; Sivesind and Seguin, 2005; Saviranta et al. 2008; Saviranta et al. 2010a; Saviranta et al. 2010b; Chiriac et al., 2020). Crowns are the meristematic tissues providing root and shoot regrowth in the spring and, as such, are vital for the winter survival of red clover that likely allocates most of its energy in protecting crowns during cold acclimation. Most of the detected compounds belong to the isoflavones, a class of phenolic compounds particularly abundant in legumes (Mazur et al., 1998). An overview of the biosynthetic pathway leading to the synthesis of isoflavones is shown in Supplementary Figure 4. We identified 31 major phenolic compounds in red clover crowns. These findings are consistent with previous studies that identified 28 and 31 phenolic compounds in red clover roots and leaves, respectively (Saviranta et al. 2010a; Saviranta et al. 2010b). We found that pseudobaptigenin-G-M, formononetin-G-M and maackiain-G-M were the most abundant phenolic compounds in red clover crowns. These three compounds were also the most abundant phenolic compounds found in roots of field-grown red clover by Saviranta et al. (2010a). Roots of field-grown plants, and leaves, had a somewhat different profile, with higher levels of formononetin aglycone (A) and biochanin A in roots of pot-grown plants (Saviranta et al., 2010a), and higher levels of biochanin A-G-M) and phaselic acid in leaves (Saviranta et al., 2010b).

We hypothesized that the accumulation of phenolic compounds in clover crowns would be significantly increased by the CA treatment because of a cold-induced stress response. However, the effect of CA was only minor as only twelve identified phenolic compounds were significantly different in their abundance between CA and NA plants, of which only three were among the compounds with the highest intensities. Of these, the amino acid L-tryptophan was significantly more abundant in CA crowns. Interestingly, another aromatic amino acid, L-phenylalanine, and a third compound with an amino acid-like structure N-(phenylacetyl) aspartic acid, were among the compounds significantly more concentrated in CA crowns (the latter two were identified with low abundance). A recent study involving application of L- phenylalanine and tyrosine to wheat plants revealed that these amino acids are precursors of the flavonoid pathway (Feduraev et al., 2020). Thus, the increased level of L- phenylalanine in the CA crowns in our study may reflect the increased metabolic activity towards the accumulation of phenolic compounds in red clover. However, we found a decreased level of tyrosine after CA, which is somehow in contrast with this hypothesis. The three other phenolic compounds showing significantly higher abundance in CA than NA crowns, daidzein, genistein and an isomer of biochanin A, were all previously identified in red clover roots (Saviranta et al., 2008), and were suggested to play a role in the defense in the rhizosphere (Dinkins et al., 2021). Among the major phenolics in our study were deoxidated forms

of clovamide, a phenolic compound from the class of hydroxy-cinnamic acid amides previously found in red clover roots after treatment with jasmonic acid and that was proposed to be involved in the plant defense system (Tebayashi et al., 2000).

In general, the higher total number of peaks distinguishing the CA vs. NA and the fact that some of the major peaks were significantly different between NA and CA, indicated that the cold acclimation treatment had a stronger effect than accession group in the accumulation of phenolic compounds in red clover crowns. Schulz et al. (2015) found significantly higher amounts of phenolic compounds in the classes of flavonols and anthocyanins in A. thaliana leaves after 14 days of CA at 4°C. However, in that study, the response differed between accessions with different genetic backgrounds, which is similar to the situation in red clover grown in a multi-year multi-location field trial (Sivesind and Seguin, 2005). These observations indicate a strong genetic control of the accumulation of phenolic compounds under CA. Furthermore, Schulz et al. (2015) showed that the levels of only a few phenolic compounds were significantly correlated to freezing tolerance. A subsequent study (Schulz et al., 2016) employing plants mutated in different genes in the phenolic compound biosynthesis pathway, suggested a redundancy effect of different flavonoids on the level of freezing tolerance. These authors also suggested that a combined contribution of a set of different phenolic compounds is more likely than the specific action of a particular compound. Furthermore, Saviranta et al. (2010a) found only a minor effect of ozone stress in the accumulation of single phenolic compounds in red clover roots. Similar results were reported in a study on red clover leaves (Saviranta et al. 2010b), where ozone stress significantly increased the total phenolic content, while single compounds were only marginally affected. Furthermore, Saviranta et al. (2008) showed that the concentration of four different isoflavones in red clover was tissue specific, and that these differences also depended on the developmental stage, something which was also confirmed in a later study (Saviranta et al. 2010b).

Author contributions

SZ conceived the idea. SZ, ÅE and AB planned the experiment. SZ conducted the growth experiment, extracted phenolic compounds and prepared samples for carbohydrate extraction and amino acid extraction. AB conducted analysis of carbohydrate and amino acid content. FME and RWP conducted the analysis of phenolic compounds. SZ and RWP produced figures and tables. SZ wrote the paper under supervision and co-writing by ÅE and JEO. AB and RWP contributed in the writing of the paper. All authors read and approved the final version.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We would like to thank Øyvind Jørgensen, Liv Berge, Josée Bourassa and Sandra Delaney for their technical assistance.

References

Abberton, M.T., Marshall, A.H., 2005. Progress in breeding perennial clovers for temperate agriculture. The Journal of Agricultural Science 143, 117–135. doi.org/10.1017/S0021859605005101

Agati, G., Azzarello, E., Pollastri, S., Tattini, M., 2012. Flavonoids as antioxidants in plants: Location and functional significance. Plant Science 196, 67–76. doi.org/10.1016/j.plantsci.2012.07.014

Alcázar, R., Cuevas, J.C., Planas, J., Zarza, X., Bortolotti, C., Carrasco, P., Salinas, J., Tiburcio, A.F., Altabella, T., 2011. Integration of polyamines in the cold acclimation response. Plant Science, Plant and Microbe Adaptation to Cold 180, 31–38. doi.org/10.1016/j.plantsci.2010.07.022

Annicchiarico, P., Barrett, B., Brummer, E.C., Julier, B., Marshall, A.H., 2015. Achievements and Challenges in Improving Temperate Perennial Forage Legumes. Critical Reviews in Plant Sciences 34, 327–380. doi.org/10.1080/07352689.2014.898462

Bélanger, G., Castonguay, Y., Bertrand, A., Dhont, C., Rochette, P., Couture, L., Drapeau, R., Mongrain, D., Chalifour, F.-P., Michaud, R., 2006. Winter damage to perennial forage crops in eastern Canada: Causes, mitigation, and prediction. Canadian Journal of Plant Science 86, 33–47. doi.org/10.4141/P04-171

Bertrand, A., Bipfubusa, M., Castonguay, Y., Rocher, S., Szopinska-Morawska, A., Papadopoulos, Y., Renaut, J., 2016. A proteome analysis of freezing tolerance in red clover (Trifolium pratense L.). BMC Plant Biology 16, 65. doi.org/10.1186/s12870-016-0751-2

Bertrand, A., Bipfubusa, M., Claessens, A., Rocher, S., Castonguay, Y., 2017. Effect of photoperiod prior to cold acclimation on freezing tolerance and carbohydrate metabolism in alfalfa (Medicago sativa L.). Plant Science 264, 122–128. doi.org/10.1016/j.plantsci.2017.09.003

Bertrand, A., Rocher, S., Claessens, A., Bipfubusa, M., Papadopoulos, Y., Castonguay, Y., 2020. Biochemical and molecular responses during overwintering of red clover populations recurrently selected for improved freezing tolerance. Plant Science 292, 110388. doi.org/10.1016/j.plantsci.2019.110388

Blakeney, A.B., Mutton, L.L., 1980. A simple colorimetric method for the determination of sugars in fruit and vegetables. Journal of the Science of Food and Agriculture 31, 889–897. doi.org/10.1002/jsfa.2740310905

Castonguay, Y., Bertrand, A., Michaud, R., Laberge, S., 2011. Cold-Induced Biochemical and Molecular Changes in Alfalfa Populations Selectively Improved for Freezing Tolerance. Crop Science 51, 2132–2144. doi.org/10.2135/cropsci2011.02.0060

Castonguay, Y., Nadeau, P., Lechasseur, P., Chouinard, L., 1995. Differential Accumulation of Carbohydrates in Alfalfa Cultivars of Contrasting Winterhardiness. Crop Science 35, 509-516. doi.org/10.2135/cropsci1995.0011183X003500020038x

Chiriac, E.R., Chiţescu, C.L., Borda, D., Lupoae, M., Gird, C.E., Geană, E.-I., Blaga, G.-V., Boscencu, R., 2020. Comparison of the Polyphenolic Profile of Medicago sativa L. and Trifolium pratense L. Sprouts in Different Germination Stages Using the UHPLC-Q Exactive Hybrid Quadrupole Orbitrap High-Resolution Mass Spectrometry. Molecules 25, 2321. doi.org/10.3390/molecules25102321

Cohen S.A. (2003) Amino Acid Analysis Using Pre-Column Derivatization with 6-AminoquinolylN-Hydroxysuccinimidyl Carbamate. In: Smith B.J. (eds) Protein Sequencing Protocols. Methods in Molecular Biology™, vol 211. Humana Press. https://doi.org/10.1385/1-59259-342-9:143

Dhont, C., Castonguay, Y., Nadeau, P., Bélanger, G., Chalifour, F.-P., 2003. Alfalfa Root Nitrogen Reserves and Regrowth Potential in Response to Fall Harvests. Crop Science 43, 181–194. doi.org/10.2135/cropsci2003.1810

Dinkins, R.D., Hancock, J., Coe, B.L., May, J.B., Goodman, J.P., Bass, W.T., Liu, J., Fan, Y., Zheng, Q., Zhu, H., 2021. Isoflavone levels, nodulation and gene expression profiles of a CRISPR/Cas9 deletion mutant in the isoflavone synthase gene of red clover. Plant Cell Rep 40, 517–528. doi.org/10.1007/s00299-020-02647-4

Dong, S., Beckles, D.M., 2019. Dynamic changes in the starch-sugar interconversion within plant source and sink tissues promote a better abiotic stress response. Journal of Plant Physiology 234–235, 80–93. doi.org/10.1016/j.jplph.2019.01.007

Elessawy, F.M., Bazghaleh, N., Vandenberg, A., Purves, R.W., 2020. Polyphenol profile comparisons of seed coats of five pulse crops using a semi-quantitative liquid chromatography-mass spectrometric method. Phytochemical Analysis 31, 458–471. doi.org/10.1002/pca.2909

Elessawy, F.M., Vandenberg, A., El-Aneed, A., Purves, R.W., 2021. An Untargeted Metabolomics Approach for Correlating Pulse Crop Seed Coat Polyphenol Profiles with Antioxidant Capacity and Iron Chelation Ability. Molecules 26, 3833. doi.org/10.3390/molecules26133833

Gagné-Bourque, F., Bertrand, A., Claessens, A., Aliferis, K.A., Jabaji, S., 2016. Alleviation of Drought Stress and Metabolic Changes in Timothy (Phleum pratense L.) Colonized with Bacillus subtilis B26. Frontiers in Plant Science 7.

Giri, J., 2011. Glycinebetaine and abiotic stress tolerance in plants. Plant Signaling & Behavior 6, 1746–1751. https://doi.org/10.4161/psb.6.11.17801

Feduraev, P., Skrypnik, L., Riabova, A., Pungin, A., Tokupova, E., Maslennikov, P., Chupakhina, G., 2020. Phenylalanine and Tyrosine as Exogenous Precursors of Wheat (Triticum aestivum L.) Secondary Metabolism through PAL-Associated Pathways. Plants (Basel) 9, 476. doi.org/10.3390/plants9040476

Fürtauer, L., Weiszmann, J., Weckwerth, W., Nägele, T., 2019. Dynamics of Plant Metabolism during Cold Acclimation. International Journal of Molecular Sciences 20, 5411. doi.org/10.3390/ijms20215411

Hart, E.H., Onime, L.A., Davies, T.E., Morphew, R.M., Kingston-Smith, A.H., 2016. The effects of PPO activity on the proteome of ingested red clover and implications for improving the nutrition of grazing cattle. Journal of Proteomics 141, 67–76. doi.org/10.1016/j.jprot.2016.04.023

Helgadóttir Á., Frankow-Lindberg B. E., Seppänen M. M., Søegaard K. & Østrem L. (2014). European grasslands overview: Nordic region. In Hopkins et al (Eds.), EGF at 50: The future of European grasslands. Grassland Science in Europe. 19, 15–28.

Hildebrandt, T.M., 2018. Synthesis versus degradation: directions of amino acid metabolism during Arabidopsis abiotic stress response. Plant Mol Biol 98, 121–135. doi.org/10.1007/s11103-018-0767-0

Jensen, E.S., Peoples, M.B., Boddey, R.M., Gresshoff, P.M., Hauggaard-Nielsen, H., J.R. Alves, B., Morrison, M.J., 2012. Legumes for mitigation of climate change and the provision of

feedstock for biofuels and biorefineries. A review. Agron. Sustain. Dev. 32, 329–364. doi.org/10.1007/s13593-011-0056-7
John, R., Anjum, N.A., Sopory, S.K., Akram, N.A., Ashraf, M., 2016. Some key physiological and molecular processes of cold acclimation. Biol Plant 60, 603–618. doi.org/10.1007/s10535-016-0648-9
Kassambara, A., Mundt, F., (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. CRAN.R-project.org/package=factoextra
Le, S., Josse, J., Husson, F. (2008). FactoMineR: An R Package for Multivariate Analysis. Journal of Statistical Software, 25(1), 1-18. doi.org/10.18637/jss.v025.i01
Lee, M.R.F., 2014. Forage polyphenol oxidase and ruminant livestock nutrition. Front. Plant Sci. 5. doi.org/10.3389/fpls.2014.00694
Li, Z., Zhang, Y., Peng, D., Peng, Y., Zhang, X., Ma, X., Huang, L., Yan, Y., 2018. The inhibition of polyamine biosynthesis weakens the drought tolerance in white clover (Trifolium repens) associated with the alteration of extensive proteins. Protoplasma 255, 803–817.doi.org/10.1007/s00709-017-1186-9
Lin, L.-Z., He, X.-G., Lindenmaier, M., Yang, J., Cleary, M., Qiu, S.-X., Cordell, G.A., 2000. LC-ESI-MS Study of the Flavonoid Glycoside Malonates of Red Clover (Trifolium pratense). J. Agric. Food Chem. 48, 354–365. doi.org/10.1021/jf991002+
Lüscher, A., Mueller-Harvey, I., Soussana, J.F., Rees, R.M., Peyraud, J.L., 2014. Potential of legume-based grassland–livestock systems in Europe: a review. Grass and Forage Science 69, 206–228. doi.org/10.1111/gfs.12124
Mazur, W.M., Duke, J.A., Wähälä, K., Rasku, S., Adlercreutz, H., 1998. Isoflavonoids and Lignans in Legumes: Nutritional and Health Aspects in Humans 11The method development and synthesis of the standards and deuterium-labelled compounds was supported by National Institutes of Health Grants No. 1 R01 CA56289-01 and No. 2 R01 CA56289-04, and analytical work by the EU research contract FAIR-CT95-0894. The Journal of Nutritional Biochemistry 9, 193–200.doi.org/10.1016/S0955-2863(97)00184-8
Menéndez, A.B., Calzadilla, P.I., Sansberro, P.A., Espasandin, F.D., Gazquez, A., Bordenave, C.D., Maiale, S.J., Rodríguez, A.A., Maguire, V.G., Campestre, M.P., Garriz, A., Rossi, F.R., Romero, F.M., Solmi, L., Salloum, M.S., Monteoliva, M.I., Debat, J.H., Ruiz, O.A., 2019. Polyamines and Legumes: Joint Stories of Stress, Nitrogen Fixation and Environment. Frontiers in Plant Science 10. doi.org/10.3389/fpls.2019.01415
Mirali, M., Ambrose, S.J., Wood, S.A., Vandenberg, A., Purves, R.W., 2014. Development of a fast extraction method and optimization of liquid chromatography–mass spectrometry for the analysis of phenolic compounds in lentil seed coats. Journal of Chromatography B 969, 149–161.doi.org/10.1016/j.jchromb.2014.08.007
Nishizawa, A., Yabuta, Y., Shigeoka, S., 2008. Galactinol and Raffinose Constitute a Novel Function to Protect Plants from Oxidative Damage. Plant Physiology 147, 1251–1263. doi.org/10.1104/pp.108.122465
Pál, M., Szalai, G., Janda, T., 2015. Speculation: Polyamines are important in abiotic stress signaling. Plant Science 237, 16–23. doi.org/10.1016/j.plantsci.2015.05.003
Pommerrenig, B., Ludewig, F., Cvetkovic, J., Trentmann, O., Klemens, P.A.W., Neuhaus, H.E., 2018. In Concert: Orchestrated Changes in Carbohydrate Homeostasis Are Critical for Plant Abiotic Stress Tolerance. Plant and Cell Physiology 59, 1290–1299. doi.org/10.1093/pcp/pcy037
R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/
Saviranta, N.M., Anttonen, M.J., Wright, A. von, Karjalainen, R.O., 2008. Red clover (Trifolium pratense L.) isoflavones: determination of concentrations by plant stage, flower colour, plant part and cultivar. Journal of the Science of Food and Agriculture 88, 125–132. doi.org/10.1002/jsfa.3056

Saviranta, N.M., Julkunen-Tiitto, R., Oksanen, E., Karjalainen, R.O., 2010a. Red clover (Trifolium pratense L.) isoflavones: root phenolic compounds affected by biotic and abiotic stress factors. Journal of the Science of Food and Agriculture 90, 418–423. doi.org/10.1002/jsfa.3831

Saviranta, N.M.M., Julkunen-Tiitto, R., Oksanen, E., Karjalainen, R.O., 2010b. Leaf phenolic compounds in red clover (Trifolium pratense L.) induced by exposure to moderately elevated ozone. Environmental Pollution 158, 440–446. doi.org/10.1016/j.envpol.2009.08.029

Schulz, E., Tohge, T., Zuther, E., Fernie, A.R., Hincha, D.K., 2016. Flavonoids are determinants of freezing tolerance and cold acclimation in Arabidopsis thaliana. Sci Rep 6. doi.org/10.1038/srep34027

Schulz, E., Tohge, T., Zuther, E., Fernie, A.R., Hincha, D.K., 2015. Natural variation in flavonol and anthocyanin metabolism during cold acclimation in Arabidopsis thaliana accessions. Plant, Cell & Environment 38, 1658–1672. doi.org/10.1111/pce.12518

Sengupta, S., Mukherjee, S., Basak, P., Majumder, A.L., 2015. Significance of galactinol and raffinose family oligosaccharide synthesis in plants. Front. Plant Sci. 6. doi.org/10.3389/fpls.2015.00656

Sivesind, E., Seguin, P., 2005. Effects of the Environment, Cultivar, Maturity, and Preservation Method on Red Clover Isoflavone Concentration. J. Agric. Food Chem. 53, 6397–6402. doi.org/10.1021/jf0507487

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.-M., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A.N., Lindon, J.C., Marriott, P., Nicholls, A.W., Reily, M.D., Thaden, J.J., Viant, M.R., 2007. Proposed minimum reporting standards for chemical analysis. Metabolomics 3, 211–221. doi.org/10.1007/s11306-007-0082-2

Szabados, L., Savouré, A., 2010. Proline: a multifunctional amino acid. Trends in Plant Science 15, 89–97. doi.org/10.1016/j.tplants.2009.11.009

Tebayashi, S., Ishihara, A., Tsuda, M., Iwamura, H., 2000. Induction of clovamide by jasmonic acid in red clover. Phytochemistry 54, 387–392. doi.org/10.1016/S0031-9422(00)00098-4

Theocharis, A., Clément, C., Barka, E.A., 2012. Physiological and molecular changes in plants grown at low temperatures. Planta 235, 1091–1105. doi.org/10.1007/s00425-012-1641-y

Tsao, R., Papadopoulos, Y., Yang, R., Young, J.C., McRae, K., 2006. Isoflavone Profiles of Red Clovers and Their Distribution in Different Parts Harvested at Different Growing Stages. J. Agric. Food Chem. 54, 5797–5805. doi.org/10.1021/jf0614589

Van den Ende, W., 2013. Multifunctional fructans and raffinose family oligosaccharides. Front. Plant Sci. 4. doi.org/10.3389/fpls.2013.00247

Vinson, C.C., Mota, A.P.Z., Porto, B.N., Oliveira, T.N., Sampaio, I., Lacerda, A.L., Danchin, E.G.J., Guimaraes, P.M., Williams, T.C.R., Brasileiro, A.C.M., 2020. Characterization of raffinose metabolism genes uncovers a wild Arachis galactinol synthase conferring tolerance to abiotic stresses. Scientific Reports 10, 15258. doi.org/10.1038/s41598-020-72191-4

Yong, B., Xie, H., Li, Z., Li, Y.-P., Zhang, Y., Nie, G., Zhang, X.-Q., Ma, X., Huang, L.-K., Yan, Y.-H., Peng, Y., 2017. Exogenous Application of GABA Improves PEG-Induced Drought Tolerance Positively Associated with GABA-Shunt, Polyamines, and Proline Metabolism in White Clover. Frontiers in Physiology 8, 1107. doi.org/10.3389/fphys.2017.01107

Zanotto, S., Palmé, A., Helgadóttir, Á., Daugstad, K., Isolahti, M., Öhlund, L., Marum, P., Moen, M.A., Veteläinen, M., Rognli, O.A., Ergon, Å., 2021. Trait characterization of genetic resources reveals useful variation for the improvement of cultivated Nordic red clover. Journal of Agronomy and Crop Science. doi.org/10.1111/jac.12487

Zhuo, C., Liang, L., Zhao, Y., Guo, Z., Lu, S., 2018. A cold responsive ethylene responsive factor from Medicago falcata confers cold tolerance by up-regulation of polyamine turnover, antioxidant protection, and proline accumulation. Plant, Cell & Environment 41, 2021–2032. doi.org/10.1111/pce.13114

Table 1. Information about the origin, freezing tolerance and winter survival of ten Nordic red clover accessions used in the study. FI, Finland; NO, Norway; SE, Sweden; FS freezing susceptible; FT, freezing tolerant.

Accession	NordGen ID†	Type	Collection site	Latitude	Longitude	Altitude (m a.s.l.)	Freezing tolerance (LT50) ††	%WS†††	Group
HIETAMA AP0201	NGB2844	Landrace	Hietama, Äänekoski (FI)	N 62° 35'	E 25° 36'	130	-6.73	0.17	FS
VUOTJÄRVI HM0201	NGB14442	Wild	(FI)	N 63° 15'	E 28° 18'	152	-8.58	0.33	FS
NOLGÅRDEN NÄS FO0501	NGB16965	Wild	(NO)	N 58° 5'	E 13° 42'	224	-9.08	0.37	FS
SAUKKAVAARA ME0201	NGB1139	Landrace	Saukkavaara, Paltamo (FI)	N 64° 25'	E 28° 11'	197	-9.11	0.14	FS
ULTUNA	NGB4089	Landrace	(SE)	N 59° 55'	E 17° 29'	43	-9.18	0.00	FS
TØEN 09-6-48-1	NGB1571	Wild	Tøen, Vingelen (NO)	N 62° 25'	E 10° 52'	890	-12.15	0.70	FT
SÖDRA SUNDERBYN PH0102	NGB1022	Landrace	Södra Sunderbyn (SE)	N 65° 40'	E 21° 52'	22	-12.16	0.42	FT
ROLIGHETEN 09-6-54-1	NGB6736	Wild	Roligheten, Skien (NO)	N 59° 13'	E 9° 40'	319	-12.33	0.80	FT
RAUNAVAARA AK0402	NGB1011	Wild	Raunavaara (SE)	N 66° 30'	E 23° 37'	163	-12.42	0.95	FT
HYRKÄS ME0101	NGB1146	Landrace	Hyrkäs, Muhos (FI)	N 64° 48'	E 26° 7'	65	-13.09	0.61	FT

†ID code in NordGen (the Nordic Genetic Resource Centre)

††Freezing tolerance as determined in Zanotto et al. (2021)

†††Winter survival at Jokioinen, Finland, in 2016 as determined in Zanotto et al. (2021)

Table 2. Identities of polyphenols labelled in Figure 5.

Peak #	Retention time (min)	Polyphenol[1]	Chemical formula	Measured molecular weight (Da)[2]	Mass Error (ppm)[3]	λ_{max} (nm)[4]	ID level[5]
1	4.53	L-Tryptophan	$C_{11}H_{12}N_2O_2$	204.08958	-1.45	280	2
2	9.54	*cis*-clovamide	$C_{18}H_{17}NO_7$	359.09971	-2.21	290,320	2/3
3	10.63	Calycosin-G	$C_{22}H_{22}O_{10}$	446.12068	-1.38	250,260,285	2/3
4	10.82	Deoxyclovamide (caffeoyl tyrosine)	$C_{18}H_{17}NO_6$	343.10559	-1.50	295,320	2/3
5	11.39	Genistin	$C_{21}H_{20}O_{10}$	432.10510	-1.27	260,318	2
6	11.75	Pratensein 7-*O*-*β*-D-glucoside	$C_{22}H_{22}O_{11}$	462.11555	-1.43	263,328	2
7	12.18	Dideoxyclovamide	$C_{18}H_{17}NO_5$	327.11026	-1.26	299,308	2/3
8	12.24	Daidzein G-M	$C_{24}H_{22}O_{12}$	502.11028	-1.68	262,308	2/3
9	12.41	Calycosin-G-M	$C_{25}H_{24}O_{13}$	532.12097	-1.35	260,290w	2/3
10	12.51	Calycosin-G-M	$C_{25}H_{24}O_{13}$	532.12106	-1.18	260,290w	2/3
11	13.04	Calycosin-G-M	$C_{25}H_{24}O_{13}$	532.12104	-1.22	260,286	2/3
12	13.36	Pseudobaptigenin-G	$C_{22}H_{20}O_{10}$	444.10518	-1.05	250,262,292	2/3
13	13.42	Irilin B/methylorobol-G-M	$C_{25}H_{24}O_{14}$	548.11661	-1.21	263,334w	2/3
14	13.71	Ononin (formononetin 7-*O*-*β*-D-glucoside)	$C_{22}H_{22}O_9$	430.12602	-0.85	260,300	2
15	13.91	Genistin 6”-*O*-malonate	$C_{24}H_{22}O_{13}$	518.10562	-0.82	260,330w	2
16	14.09	Pratensein 7-*O*-*β*-D-glucoside-6”-*O*-malonate	$C_{25}H_{24}O_{14}$	548.11595	-1.19	264,330	2
17	14.75	Irilin B/methylorobol-G-M	$C_{25}H_{24}O_{14}$	548.11595	-1.20	260,290	2/3
18	14.90	Calycosin-G-M-M	$C_{28}H_{26}O_{16}$	618.12108	-1.62	250,285w	2/3
19	15.22	Pseudobaptigenin-G-M	$C_{25}H_{22}O_{13}$	530.10559	-0.85	260w,290	2/3
20	15.41	Calycosin	$C_{16}H_{12}O_5$	284.06817	-1.07	a	2
21	15.49	Irilone/afrormosin-G	$C_{22}H_{20}O_{11}$	460.10004	-1.13	268,340w	2/3
22	15.56	Formononetin-G-M	$C_{25}H_{24}O_{12}$	516.12640	-0.73	260,305	2/3
23	15.69	Pseudobaptigenin-G-M	$C_{25}H_{22}O_{13}$	530.10556	-0.91	260,290	2/3
24	16.01	Formononetin-G-M	$C_{25}H_{24}O_{12}$	516.12638	-0.78	260,305	2/3
25	16.98	Maackiain-G-M+NH3	$C_{25}H_{27}NO_{13}$	549.14736	-1.60	284,309s	3
26	16.98	Maackiain-G-M	$C_{25}H_{24}O_{13}$	532.12092	-1.45	284,309s	2/3
27	17.24	Irilone/afrormosin-G-M	$C_{25}H_{22}O_{14}$	546.10044	-0.95	266,334w	2/3
28	17.46	Irilone/afrormosin-G-M	$C_{25}H_{22}O_{14}$	546.10044	-0.94	268,338w	2/3
29	17.95	Formononetin-G-M-M	$C_{28}H_{26}O_{15}$	602.12636	-1.35	250,300w	2/3
30	18.35	Biochanin A 7-*O*-*β*-D-glucoside-6”-*O*-malonate	$C_{25}H_{24}O_{13}$	532.12121	-0.91	260,322	2
31	18.76	Prunetin-G-M	$C_{25}H_{24}O_{13}$	532.12123	-0.86	260,325w	2/3
32	19.08	Pseudobaptigenin	$C_{16}H_{10}O_5$	282.05276	-0.21	250,293	2
33	19.47	Formononetin	$C_{16}H_{12}O_4$	268.07317	-1.45	250,301	2

[1] A = aglycone, G = glucoside or galactoside, M = malonate; [2] Mass of the neutral compound for the most abundant isotope based on the measured accurate mass (Da); [3] Mass error = ((measured accurate mass – calculated exact mass)/(calculated exact mass)) x 10^6 in parts per million (ppm); [4] a = too small to measure accurately, w = weak, s = strong; [5] Identification level as described in Sumner et al. 2007 with the addition of the 2/3 level in Elessawy et al. 2021

Table 3. Compounds identified in Figure 6, listed with increasing log_2 fold change.

Peak #	Retention time (min)	Polyphenol[1]	Chemical formula	Measured molecular weight (Da) [1]	Mass Error (ppm) [2]	Log_2 fold change	P-value	ID level[3]
		*Figure 5 (**A**): CA vs. NA*[4]						
38	10.60	Formononetin derivative	$C_{31}H_{38}N_2O_{10}$	598.25155	-1.83	-1.59	1.1e-11	3
7	12.18	Dideoxyclovamide	$C_{18}H_{17}NO_5$	327.11026	-1.26	-1.23	6.3e-22	2/3
39	6.00	Tetrahydro-carboline-3-carboxylic acid	$C_{12}H_{12}N_2O_2$	216.08944	-2.03	-1.22	4.1e-5	2/3
40	15.29	Irilone derivative	$C_{29}H_{31}NO_{12}$	585.18364	-1.69	-1.12	2.5e-8	3
4	10.82	Deoxyclovamide (caffeoyl tyrosine)	$C_{18}H_{17}NO_6$	343.10559	-1.50	-1.09	1.8e-18	2/3
41	15.08	Irilone derivative	$C_{29}H_{31}NO_{12}$	585.18364	-1.69	-1.05	2.1e-8	3
42	14.41	Daidzein	$C_{15}H_{10}O_4$	254.05789	-0.07	1.06	1.5e-12	2
43	22.85	Biochanin A isomer	$C_{16}H_{12}O_5$	284.06827	-0.73	1.08	2.4e-5	3
44	2.11	L-phenylalanine	$C_9H_{11}NO_2$	165.07881	-1.04	1.30	1.8e-9	2
1	4.55	L-tryptophan	$C_{11}H_{12}N_2O_2$	204.08958	-1.45	1.31	5.5e-13	2
45	17.24	Genistein	$C_{15}H_{10}O_5$	270.05274	-0.30	1.56	1.9e-15	2
46	7.21	N-(phenylacetyl)aspartic acid	$C_{12}H_{13}NO_5$	251.07915	-0.87	1.68	2.8e-9	2/3
		*Figure 5 (**B**): FT vs. FS*[5]						
34	9.47	Phloretin isomer	$C_{15}H_{14}O_5$	274.08375	-1.35	-1.17	1.2e-8	3
35	8.66	Dihydrokaempferol	$C_{15}H_{12}O_6$	288.06302	-1.27	-1.15	3.8e-7	2
36	20.42	Biochanin A-G-M-M	$C_{28}H_{26}O_{16}$	618.12121	-1.42	-1.04	7.7e-6	2/3
37	7.88	Pinocembrin-G derivative	$C_{21}H_{27}NO_{10}$	453.16272	-1.71	1.63	3.0e-10	3

[1] A = aglycone, G = glucoside or galactoside, M = malonate; [2] Mass of the neutral compound for the most abundant isotope based on the measured accurate mass (Da).; [2]Mass error = ((measured accurate mass – calculated exact mass)/(calculated exact mass)) x 10^6 in parts per million (ppm); [3]identification level as described in Sumner et al. 2007 with the addition of the 2/3 level in Elessawy et al. 2021; [4] *CA* = Cols acclimated, *NA* = non acclimated; [5] *FT* = freezing tolerant, *FS* =freezing susceptible

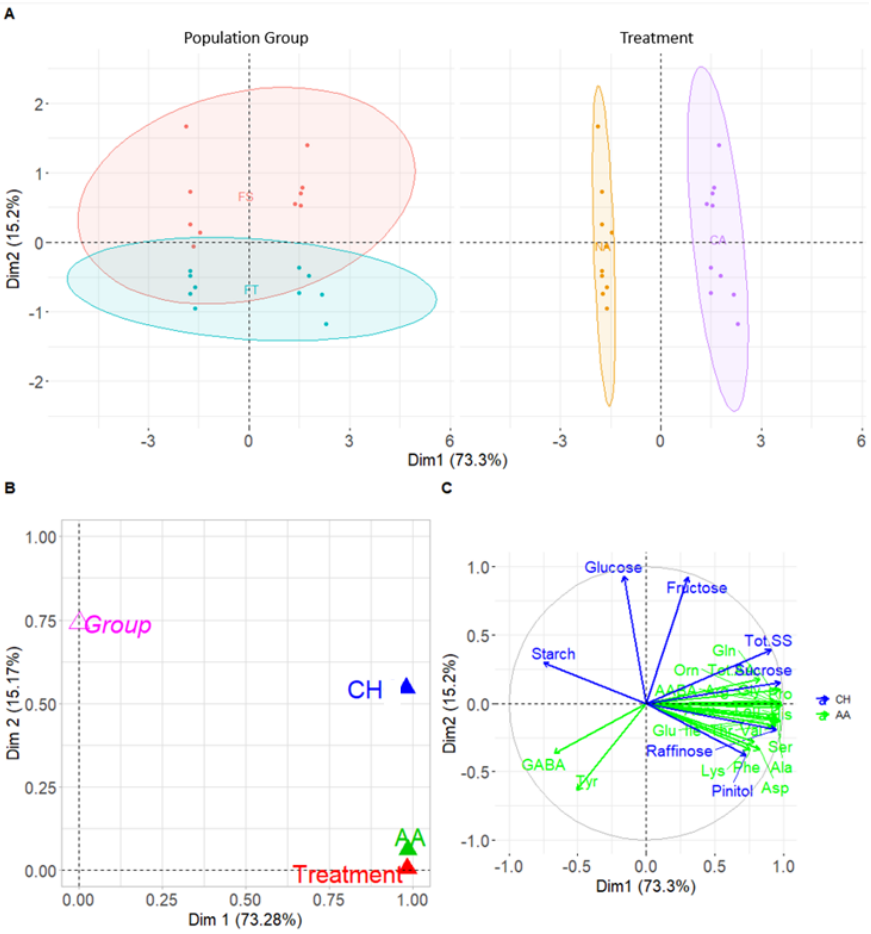


Figure 1. Results of multi-factor analysis (MFA) of carbohydrate (CH) and amino acid (AA) content in crowns of 5 freezing sensitive (FS) and 5 freezing tolerant (FT) red clover accessions that had been either cold acclimated (CA) or non-acclimated (NA). For each accession and acclimation treatment, the content of a set of CHs and AAs was measured in 4 replicate samples, each consisting of 30 individual crowns. The content of individual CHs and AAs were active quantitative variables while acclimation treatment and accession group were active and supplementary categorical variables in the analysis. (**A**), representation of the accessions in the space defined by the first two dimensions of the MFA. Dots representing the accessions are color coded to indicate accession group and treatment on the left and right side, respectively. 95% confidence ellipses are shown. (**B**), projection of the variable groups in the first two dimensions of the MFA. The dots representing the variable groups can assume values between 0 and 1 in the first two MFA dimensions, with a score close to 1 meaning that the variable group is very strongly linked to that dimension of the MFA. (**C**), correlation of quantitative variables belonging to the CH and AA variable groups with the first two dimensions of the MFA, only variables significantly correlated ($P<0.05$) with the first and/or second dimension are shown.

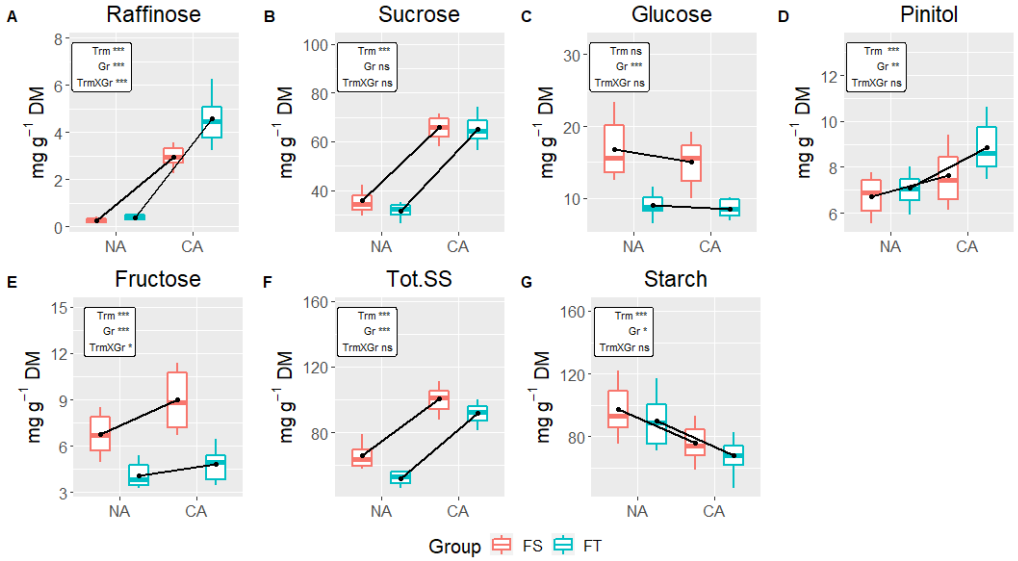


Figure 2. Content of individual and total soluble sugars as well as starch in crowns of 5 freezing sensitive (FS) and 5 freezing tolerant (FT) red clover accessions that had been either cold acclimated (CA) or non-acclimated (NA). Interaction effects are shown by the black lines connecting the means (black dots) of the two treatments within accession group. Results from analysis of variance shown in each panel indicate which factors had a significant effect. Trm; treatment (CA or NA), Gr; accession group (FS or FT), TrmXGr (treatment X accession group interaction, ***; P <0.001, **; P <0.01, *; P <0.05. Note that the scale on the y-axes vary among the different metabolites.

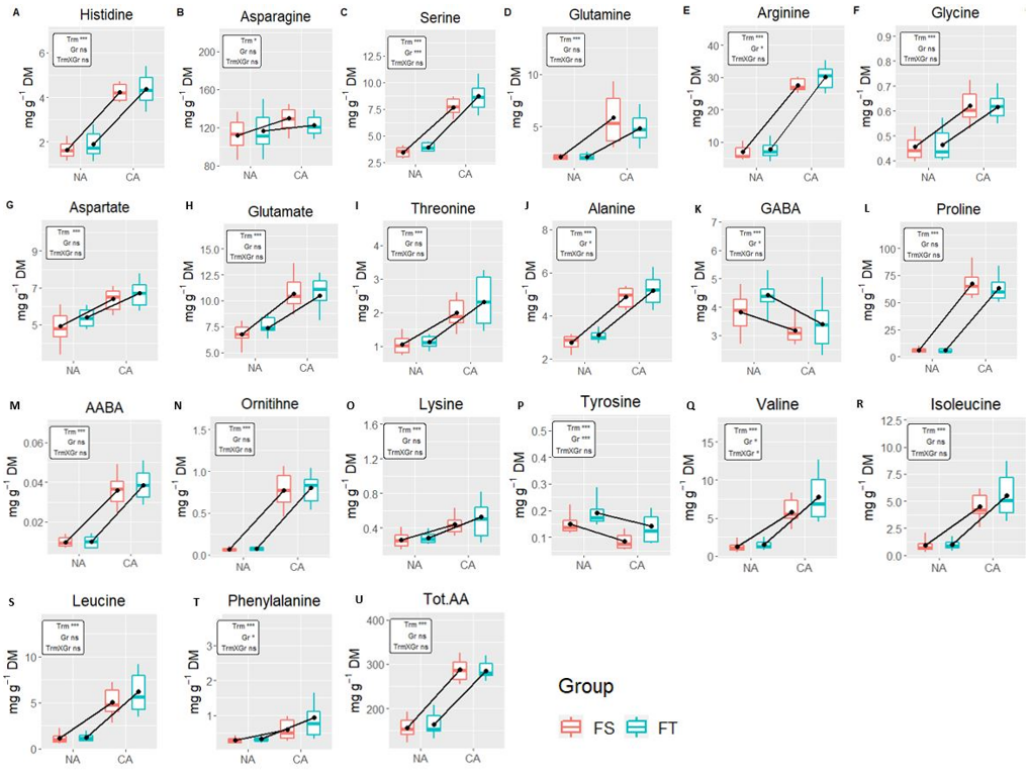


Figure 3. Content of individual and total amino acids in red clover crowns of 5 freezing sensitive (FS) and 5 freezing tolerant (FT) red clover accessions that had been either cold acclimated (CA) or non-acclimated (NA). Interaction effect is also shown by the black lines connecting the means (black dots) of the two treatments within accession group. Results from analysis of variance shown in each panel indicate which factors had a significant effect. Trm; treatment (CA or NA), Gr; accession group (FS or FT), TrmXGr (treatment X accession group interaction, ***; P <0.001, **; P <0.01, *; P <0.05. Note that the scale on the y-axes vary among the different metabolites.

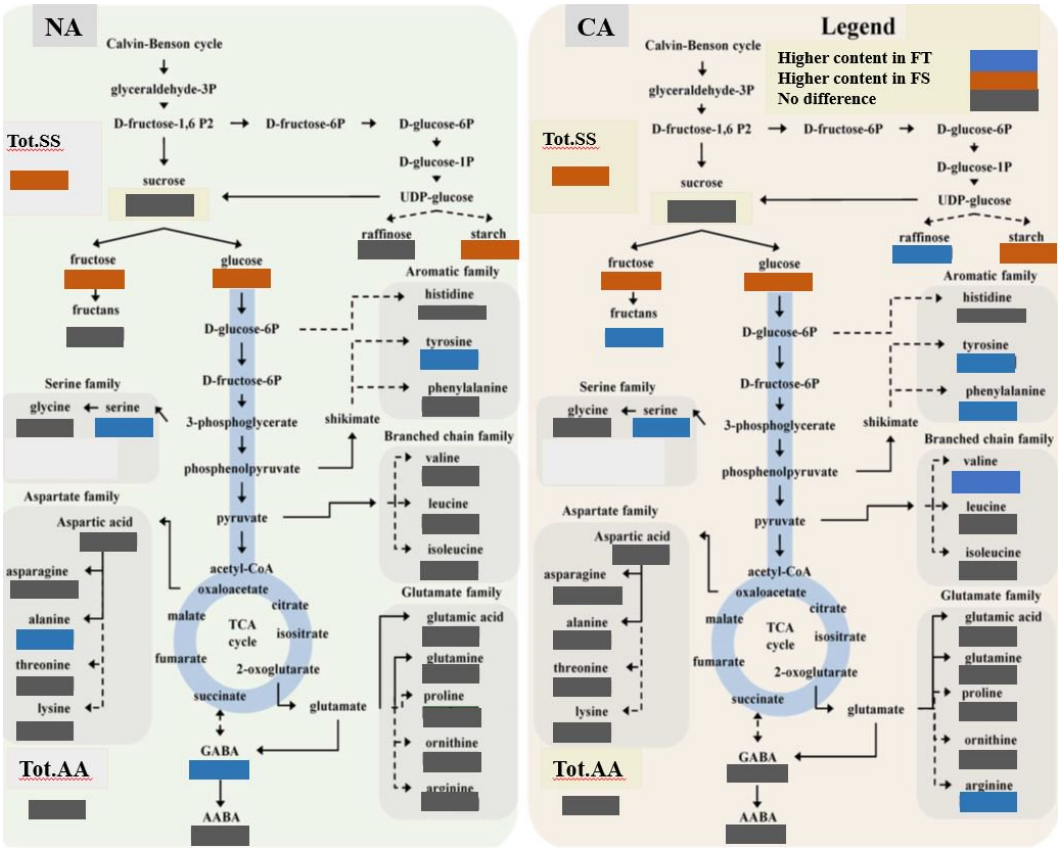


Figure 4. Metabolic pathway maps of differences in carbohydrate and amino acid content between freezing susceptible (FS) and freezing tolerant (FT) accessions which were either non-acclimated (NA) or cold acclimated (CA). Metabolites are color coded to indicate differences between FT and FS accessions, according to analyses of variance within acclimation treatments. Dashed arrows symbolize multistep reactions and solid arrows one-step reactions (figure adapted from Gagné-Bourque et al., 2016)

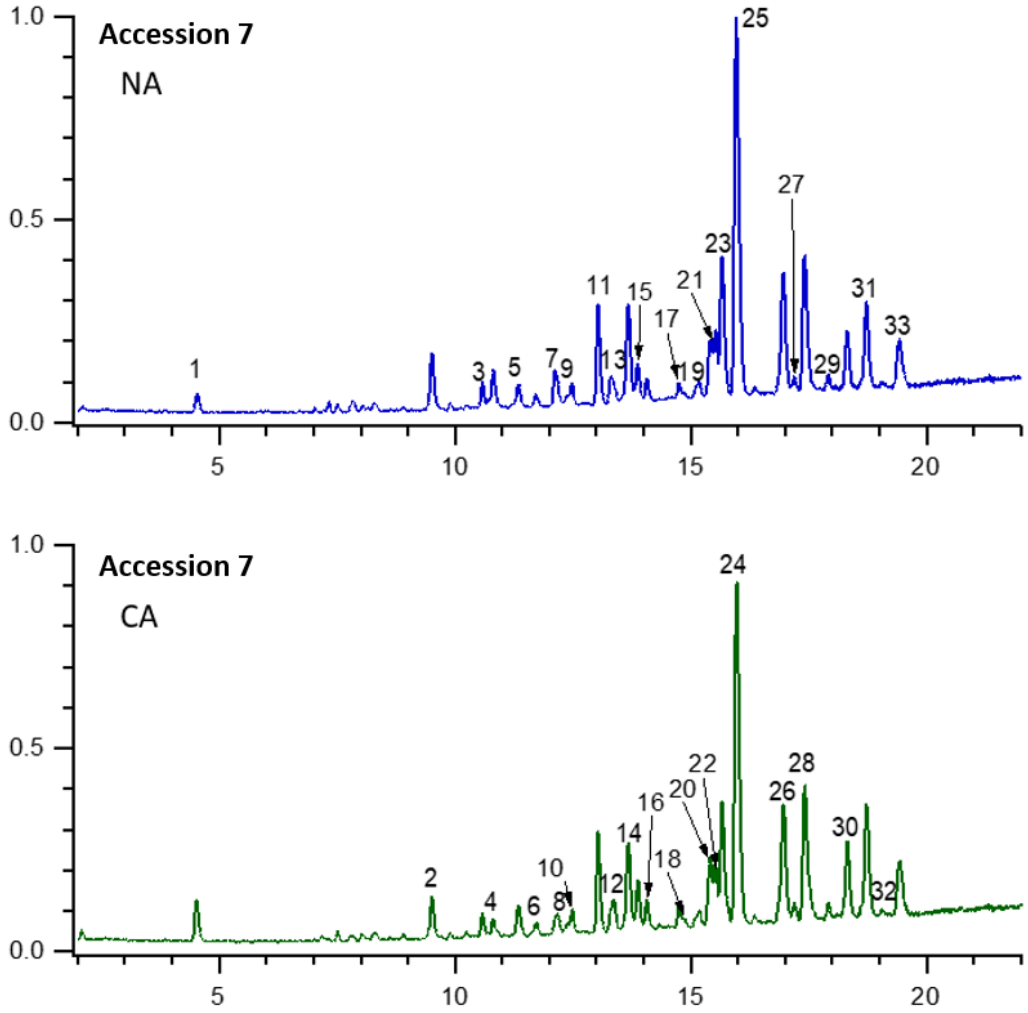


Figure 5. Representative LC-MS trace using accession 7, SÖDRA SUNDERBYN PH0102 (FT), where NA indicates non acclimated and CA indicates cold acclimated (CA). For clarity the odd numbered peaks are labelled in the upper trace and the even numbered peaks are labelled in the lower trace. The peak numbers correspond to the peak numbers given in Table 1.

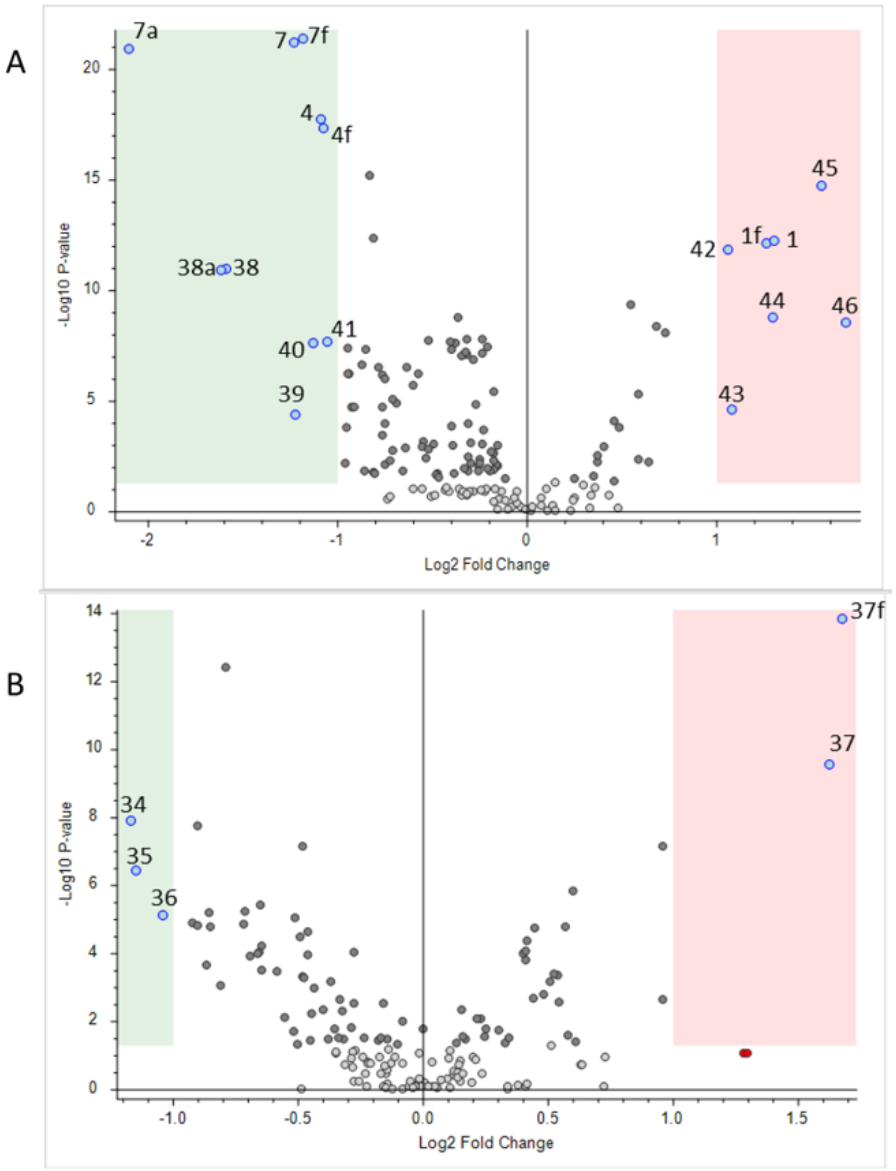


Figure 6. Volcano plots of (**A**) CA vs. NA and (**B**) FT (freezing tolerant) vs. FS (freezing susceptible) accessions. The red and green areas show the compounds that have significantly higher and lower intensities, for the two comparisons (CA vs NA and FT vs FS), respectively. The numbers correspond to the assignments in Tables 1 and 2, however, numbers with a suffix are not in the tables as they are redundant. The suffix 'a' indicates an adduct ion and the suffix "f" indicates a fragment ion.

Paper III

A genome-wide association study of freezing tolerance in red clover (*Trifolium pratense* L.)

Stefano Zanotto[1], **Tom Ruttink**[2], **Marie Pegard**[3], **Leif Skøt**[4], **Christoph Grieder**[5], **Roland Kölliker**[6] and **Åshild Ergon**[1]

[1]Norwegian University of Life Sciences (NMBU), Faculty of Biosciences, P.O. Box 5003, N-1432 Ås, Norway

[2]Flanders Research Institute for Agriculture, Fisheries and Food (ILVO), Plant Sciences Unit, Caritasstraat 39, B-9090 Melle, Belgium

[3]INRAE P3F, 86600 Lusignan, France, [4]IBERS, Aberystwyth University, Plas Gogerddan, Aberystwyth, SY23 3EE, United Kingdom

[5]Group of Fodder Crop Breeding, Agroscope, 8046 Zurich, Switzerland

[6]Molecular Plant Breeding, Institute of Agricultural Sciences, ETH Zürich, Universitätsstr. 2, 8092 Zürich, Switzerland

Abstract

Improvement of freezing tolerance (FT) is an important breeding goal in red clover (*Trifolium pratense* L.), whose field persistency is often affected by poor winter survival. Here we phenotyped 393 red clover accessions, mostly of European origin, for FT and found considerable variation, with LT50-values (temperature required to kill 50% of the plants) ranging from -5.96 to -11.47 °C. Accessions were genotyped as pools of individuals using pool-GBS, generating both SNP and haplotype allele frequency data at accession level. Linkage disequilibrium was determined as a squared partial correlation between the allele frequencies of pairs of SNPs and found to decay at extremely short distances (< 1 kb). SNP- and haplotype-based genome-wide association studies (GWAS) identified altogether 13 loci significantly associated with FT, explaining 30 and 26 % of the phenotypic variation for the SNP- and haplotype-based analyses, respectively. Ten of the loci were found within or at a short distance from genes possibly involved in mechanisms associated with FT. This study paves the way for a better understanding of the genetic control of FT and for the development of molecular markers for the improvement of this trait in red clover.

Introduction

Red clover (*Trifolium pratense* L.) is the most important perennial forage legume in northern Europe (Helgadóttir et al., 2014; Annicchiarico et al., 2015), where it is mainly cultivated in mixture with grasses under mowing regime for the production of silage and hay. Its capability to fix atmospheric nitrogen through symbiotic nitrogen fixation significantly reduces the requirement for nitrogen fertilization (Jensen et al., 2012; Lüscher et al., 2014; Reckling et al., 2016). Thanks to this, and to the high protein content of its forage (Frame et al., 1997), red clover plays an important role in the shift towards a more sustainable agriculture in the Nordic region. Like in the rest of Europe, Nordic agriculture is currently not self-sufficient when it comes to plant-based protein for food and feed (Visser et al., 2014; Voisin et al., 2014). However, the cultivation of red clover at northern latitudes is hampered by its poor field persistency, mainly caused by low winter survival (Abberton and Marshall, 2005), which is a complex trait affected by several biotic and abiotic factors (Bélanger et al., 2006). The improvement of persistency is one of the main breeding goals for the species because it will result in both higher forage yield and protein content of mixed red clover - grass swards (Marshall et al., 2017).

Among the several stresses that reduce plant survival during the winter, freezing temperatures are one of the main causes of mortality at locations with continental climate characterized by harsh winters. The ability to tolerate low freezing temperatures (freezing tolerance, FT) is achieved by a period of cold acclimation (CA) at low above-zero temperatures, which induces several changes at the transcriptomic, proteomic and metabolic level (Bertrand et al., 2020; 2016; Zanotto et al. in press). CA results in the accumulation of soluble sugars and free amino acids in the crowns, which are critical for overwintering and regrowth in red clover plants. Phenolic compounds are also involved in the response to CA, (Zanotto et al., in press). At the proteomic level, CA induces marked changes in the proteome, mainly related to enzymes involved in carbohydrate, amino acid and secondary metabolism as well as energy production (Bertrand et al., 2016). Variation in FT among red clover gene bank material of Nordic origin was recently reported (Zanotto et al., 2021). FT can be improved by selection and breeding (Bertrand et al., 2016), but little is known about the genetic control of FT in red clover. Previous studies identified regions in the red clover genome associated with persistency and survival as affected by various biotic and abiotic stresses (Ergon et al., 2019; Herrmann et al., 2008; Klimenko et al., 2010), but none of these studies specifically analyzed the level of FT of the plant material they used.

Red clover is a natural diploid (2n=14) outbreeding species with a genome size of approximately 420 Mb (Sato et al., 2005). A draft genome for the species (309 Mb, of which 164 Mb are placed on chromosomes) was published in 2015 by De Vega et al., facilitating genomic studies. Red clover is highly self-incompatible, has a very polymorphic genome, and usually a higher genetic diversity within than between populations (Jones et al., 2020). Red clover cultivars are commonly bred as synthetic populations with up to twenty or more parents. A recent study reveals that a high amount of genetic diversity is maintained within Nordic red clover cultivars created by modern breeding programs (Osterman et al., 2021). Furthermore, there is a large amount of genetic variation for several traits available among gene bank material such as landraces, wild populations and old cultivars. This can be exploited to improve cultivars for specific target traits, including traits related to persistency such as winter survival and freezing tolerance (Annicchiarico and Pagnotta, 2012; Dias et al., 2008; Greene et al., 2004; Kouamé and Quesenberry, 1993; Zanotto et al., 2021).

Genotyping by sequencing (GBS) of DNA pools is a time- and cost-effective method for genotyping a large number of accessions and when genetic characterization is more relevant at population level than at the constituent individual level, such as in plant breeding of self-incompatible forage species (Byrne et al., 2013). Genome wide association studies (GWAS) have been successfully used in forage legumes to elucidate the genetic control of complex traits (Biazzi et al., 2017; Inostroza et al., 2018). GWAS analysis may be confounded by population structure and variation in relatedness among accessions (Korte and Farlow, 2013). Therefore, it is important to use GWAS models that can account for this in order to reduce the risk of identifying false positive associations (Liu et al., 2016). In GWAS analysis, haplotype data can complement single nucleotide polymorphism (SNP) data, as haplotypes may reveal associations that are not detectable when using SNP data only (Bekele et al., 2018; Hamblin and Jannink, 2011; Ergon et al, submitted).

In this study we characterized a wide and diverse panel made up of 393 red clover accessions of mainly European origin for their level of FT with the goal of identifying regions in the red clover genome associated with this trait. Accessions were genotyped at population level by GBS on DNA extracted from pools of 200 individual plants. GWAS analyses were conducted using both SNPs and short haplotypes within GBS loci. The presence of relatedness among the accessions and the rapid linkage disequilibrium (LD) decay affecting red clover are important factors to consider when conducting genomic studies to dissect complex quantitative traits, including FT. The implication of these and other aspects for GWAS is discussed.

Material and methods

Plant material and phenotyping

A total of 393 red clover accessions were phenotyped for freezing tolerance (FT) under controlled experimental conditions. The accessions were wild populations, landraces, cultivars and breeding material, mostly of European origin (Supplementary Table 1). FT was determined in young plants after a short cold acclimation treatment (two weeks at 3–4 °C, 12 hr photoperiod and 110 µmol m^2s^{-1} PAR) and expressed as LT50 (temperature required to kill 50% of the plants), similar to the protocol described by Zanotto et al., (2021). The experiment was organized in eight incomplete blocks, each containing half of the accessions. Accessions were randomized among blocks, so that each accession was represented in each of four of the eight blocks. Due to space limitation, only one block could be freeze tested at the same time, every two weeks. Sowing, pricking and cold acclimation prior to the freezing test was therefore staggered in a sequential order, accordingly. Each block consisted of twelve sub-blocks with one plant per accession. Three plants per accession and block were exposed to one of four different testing temperatures within the range from -5 °C, to -17 °C (temperatures were adjusted after the first rounds (blocks) in order to align with the variation in FT). In total, across all blocks, each accession was represented by 48 plants tested at various temperatures. Survival data from the four blocks per accession were pooled before LT50 was estimated using the invest function of the investr package in R (Greenwell et al., 2014).

Genotyping

DNA extraction, SNP calling and filtering

The 393 accessions were genotyped at the accession level using Genotyping-by-Sequencing (GBS, Elshire et al., 2011) of DNA extracted using the DNEasy 96 well kit (QIAGEN) from pools of the single leaf first emerging from 200 seedlings per accession. The seedlings were sown in 96 well plant trays with two or three seedlings per well and germinated in the greenhouse. GBS was performed by the LGC Group (www.lgcgroup.com), using a combination of *Pst*I and *Mse*I for digestion, an adaptation that involves molecular normalization of read depth across loci within samples and size selection (range 100-250 bp, peak around 175 bp), and paired-end sequencing.

Reads were demultiplexed with cutadapt 3.3 (Martin, 2011), and 3’ restriction site remnants, common adapter sequences and 5’ restriction site remnants were removed using a

custom python script, cutadapt 3.3 (Martin, 2011) and FASTX-Toolkit 0.0.14 (Gordon & Hannon, 2010). All reads were merged with a minimum overlap of 10 bp (PEAR, Zhang et al. 2013). Merged reads were quality filtered and reads shorter than 60 bp were discarded. Reads were aligned to the red clover reference genome sequence v2.1 (De Vega, 2015) with the BWA-mem algorithm implemented in BWA 0.7.17 (Li & Durbin, 2009) with default parameters. Alignments were sorted, indexed and filtered on mapping quality 20 (q20) with SAMtools 1.10 (Li et al., 2009).

Mpileup files (converted from BAM format with SAMtools) in which all genome positions with minimum read depth of 30 were concatenated, thus joining the neighboring GBS stacks and excluding the part of the genome without coverage, were used as input to calculate Watterson's θ estimator with NPStat v0.99 (Ferretti et al., 2013). NPStat was run with the following settings: minor allele count equal to one read (MAC1), window-size equal to 10,000 bp, haploid sample size equal to 120 (the maximum number technically acceptable in NPStat), and maximum coverage equal to 500. Per population, a single genome-wide value was calculated as the mean across all windows (about 500 windows per sample). Loci with very high (>500) read depth may be derived from repetitive sequences that are mapped onto a single genomic locus and were thus excluded.

The NPStat-derived theta values per pool-GBS sample were used as diversity prior to the Bayesian SNP calling algorithm implemented in SNAPE-pooled (Raineri et al., 2012) to identify significant SNPs in each population. SNAPE-pooled was run with the following settings: - priortype = informative, -fold = folded, -nchr = 120 for consistency with NPStat.

The relative proportion of reads per SNP allele was used as an estimate of allele frequency. We used a custom python script to apply filters on the SNAPE-pooled reference allele frequency data. Filters were applied in the following order: (i) SNP positions were deleted if the reference allele was not A, C, G, or T, (ii) SNP frequencies were set to missing data when the two observed alleles were both different from the reference allele, or when the sum of the reference and the alternative allele read counts was lower than 27, (iii) using the Bayesian estimates of allele presence provided by SNAPE-pooled, we set the alternative allele frequency (and allele counts) to 0 if $p(freq_{alt} \neq 0) < 0.95$ and the reference allele frequency (and allele counts) to 0 if $p(freq_{ref} \neq 0) < 0.95$, (iv) after all the above filtering, we discarded SNP positions with more than two remaining alleles across the whole set of populations (thus removing potential residual low frequency sequencing errors). (v) One accession (EUC_TP_107) had more than 80% missing values across SNPs and was discarded from further analyses, leaving 392 accessions. (vi) SNPs with more than 5% missing values or

overall MAF < 0.05 were removed. This resulted in a set of 20,156 SNPs of which 12,777 were on chromosomes and 7,379 on unanchored scaffolds. The number of missing values per marker and per individual, as well as the allele frequency distribution, after filtering, is shown in Supplementary Figure 1 and 2. Missing data (0.72% of all sample and SNP genotype calls) in the 20,156 SNP data set were imputed by replacing each missing data point x_{ij} for accession i and SNP j with the mean of the non-missing values for SNP j of the other accessions.

Haplotype calling and filtering

Haplotype variants within GBS loci were called and their relative frequencies estimated with the SMAP package (Schaumont et al., 2022) as explained in Ergon et al. (submitted), using the SNP data set achieved when removing only SNPs with more than 20% missing data in step (vi) above as input. Haplotype variants with an overall MAF < 0.05 were discarded.

We define a haplotype polymorphism (HTP) as a GBS locus that contains two or more haplotype variants. The allele frequencies of a total of 75,247 haplotype variants within 12,036 HTPs (with an average of 6.25 haplotype variants per HTP) were estimated. This initial data set was then filtered, removing HTPs with missing values in more than 5% of the samples. This resulted in a new data set with allele frequencies of 20,745 haplotype variants in 7,477 HTPs with an average of 2.8 variants per HTP. The number of missing values per HTP and per accession, as well as the haplotype allele frequency distribution after filtering, is shown in Supplementary Figure 3 and 4. The data set was then imputed for missing values (0.38% of the HTP) with the mean allele frequencies for each specific haplotype variant as explained above for the SNP data set. Overall, 13,359 haplotype variants (64%) were located on chromosomes and 7,389 on unanchored scaffolds (36%), within 4,833 and 2,644 HTP loci, respectively. The set of 4833 HTP loci located on the chromosomes was used to obtain a measure of marker distribution across the genome in our material. These unique loci correspond to the parts of the red clover genome which were sequenced by the GBS and within which at least two haplotype variants were called and retained after filtering.

Population structure, kinship and LD

The population structure among the 392 accessions were investigated with a principal component analysis (PCA) based on the reference allele frequencies of the SNPs after filtration and imputation. A genomic relationship matrix (GRM) was calculated following the method described by Cericola et al. (2018), without bias correction. The method is based on VanRaden

(2008) and adapted for the use of allele frequency data. The genotype matrix (F_{ij}, with i indexing the samples and j indexing the markers) was centered by the mean allele frequency (average F_j), and the resulting genotype matrix M ($M_j = F_j - \sum_{j=1}^{m}/m$) was used to compute the GRM G, as follows

$$G = \frac{MM'}{\frac{1}{n}\sum_{j=1}^{m} p_j(1-p_j)}$$

where m represents the number of markers, p_j equals the frequency of the j^{th} marker, and n represents the total ploidy number, which is the sum of the ploidy number of the parents used to generate the populations (Ashraf et al., 2014). We used $n = 16$, as suggested by Cericola et al. (2018) for synthetic perennial ryegrass populations. GRMs were calculated with higher n values, and a negligible effect was found in the GWAS results using these different GRMs (data not shown), therefore a value of $n = 16$ was used throughout this study.

LD was determined as the squared partial correlation between the reference allele frequencies of pairs of SNPs ($r^2{}_s$) (Lin et al., 2012; Mangin et al., 2012). This measure removes the bias on the correlation between reference allele frequencies at different loci that is due to kinship. Calculations of $r^2{}_s$values, and LD plots were done in R studio (R Core Team, 2021), the pcor.shrink function of the corpcor package (Schafer et al., 2021) was used for calculation of $r^2{}_s$.

GWAS and candidate gene search

Different GWAS models were performed in a preliminary study with and without correction for kinship and/or population structure with the rrBLUP package (Endelman, 2011). Including the effect of population structure did not improve the GWAS model. Therefore, only kinship, accounted for by the GRMs, was included in the GWAS analyses. However, since none of the analyses performed with the rrBLUP package resulted in significant associations, SNP- and haplotype-based genome-wide association studies (GWAS) for FT were performed with the Multi Locus Mixed Model (MLMM), which incorporates multiple markers identified as significantly associated with the phenotypic trait as covariates simultaneously (Segura et al. 2012). The analyses were performed in R with the mlmm.gwas package (Bonnafous et al., 2019). The model with the lowest eBIC value was selected as the best model, and SNPs and haplotypes with a p-value lower than the Bonferroni threshold were considered as significant.

For each of the significant markers (SNPs and haplotypes) a linear regression between the reference allele frequency and FT was calculated to estimate the allele effect on the

phenotype (slope) and the phenotypic variance explained (R^2). Furthermore, to estimate the effect of kinship on the R^2, two different regression models were fitted, one with and one without kinship (as random genetic effect) with the lmkin and the lm functions of the coxme (Therneau, 2020) and stats (R Core Team, 2021) packages, respectively. The variance explained (R^2) by regression models including all the significant markers detected (by SNPs and HTPs-based GWAS models) simultaneously was also calculated, again both with and without kinship to estimate the kinship effect. All the R^2 values were estimated from the analyses output with the function r.squaredLR of the package MuMin (Barton, 2020). All the statistical analyses were performed in R studio (R Core Team 2021).

Given the limited LD found, candidate genes were identified within regions of +/- 0.5 Kbp flanking the significant SNPs and haplotypes identified by the GWAS in the red clover genome Tp2.0 (De Vega et al., 2015) with the gbrowse function available within the Legume Information System (LIS, legumeinfo.org). Gene-coding sequences (CDS) were used for BLASTn search (blast.ncbi.nlm.nih.gov) in the *A. thaliana* and *M. truncatula* genome to identify the most similar genes in these model species.

Results

Variation in freezing tolerance

LT50 among the 393 phenotyped accessions ranged between -11.47 °C and -5.96 °C with an average of -9.10 °C. The average standard error among the estimated LT50 values was 0.52 °C. The variation in LT50 across different geographical origins is shown in Figure 1. The most freezing tolerant accessions were found among Norwegian and Czech material while the most susceptible accessions were found among the southern European materials (note that higher LT50 corresponds to lower FT).

Marker density, LD and relatedness among accessions

The SNP and HTP density along the chromosomes is shown in Supplementary Figure 5. The average distance between HTP loci that mapped to the 164.2 Mb of the genome assembly assigned to chromosomes was 35 kb. Since several SNP markers are often clustered within the same GBS locus, the number of HTP loci is more representative of the functional marker density than the number of SNPs or haplotype variants. The average LD between SNP markers was very low (0.005-0.01) at distances of 1 Mbp to 1 Kbp and increased to 0.16 at distances shorter than 0.5 kbp (Table 1, Supplementary Figure 6).

The population structure based on the three first principal components of the PCA (Supplementary Figure 7) indicates that the genetic differentiation is associated with the origin of the accessions. Some groups are clearly distinguishable, in particular accessions from Southern Europe, Switzerland, the Nordic countries and to some extent accessions from Asia and Oceania. The rest of the material is admixed and form a relatively united group.

Average kinship between populations was close to zero for both the SNPs and HTP-based GRMs. The highest kinship values between two accessions were 6.94 and 7.38 for the SNPs and HTP-based GRMs, respectively. The mean diagonal value of the two GRMs was 1.27 and 1.31.

SNP- and haplotype-based GWAS

A total of eight and six loci were detected as significantly associated with LT50 by the SNP- and haplotype-based GWAS, respectively (Fig. 2). One significant SNP was located within one significant haplotype, i.e. a total of thirteen significant GBS loci were identified (Table 2**B**). Of these, eight were located on chromosomes while five were located on unanchored

scaffolds. One more haplotype marker on LG1 was just below the Bonferroni threshold of significance, but clearly detached from the expected distribution of p values in the quantile-quantile (QQ) plot (Fig. 2**B**) and was therefore also included among a set of thirteen loci considered in further analyses.

When kinship was accounted for, the R^2 for the GWAS models with the lowest e-BIC values was 0.30 and 0.26 for the SNP and haplotype-based GWAS models, respectively, and when kinship was not accounted for, the respective R^2 was 0.48 and 0.45 (Table 2**A**). These differences in R^2 underline the relatively high effect of kinship among accessions and how this influences the GWAS models.

The R^2 values for the individual significant SNPs, when kinship was and was not accounted for, ranged from 0.013 to 0.054 and from 0.19 to 0.0004, while those for the individual haplotypes ranged from 0.023 to 0.047 and from 0.19 to 0.03 (Table 2**B**). Some of the significant loci were characterized by a relatively strong effect of kinship (Table 2**B**, Supplementary Figure 8, 9). The loci mostly affected by the effect of kinship were also those whose distribution of allele frequencies was clearly associated with the origin of the plant material (Supplementary Figure 8), with at least one group having a marked different allele frequency distribution from the other groups.

Candidate gene search

For nine of the thirteen loci considered, the marker was located within a gene (Table 3), for one locus the marker was located in close proximity (<0.5 Kbp from a gene), while for the remaining three loci the scaffolds were too short for being in close proximity of candidate genes. Their functional annotation and that of the best BLAST hits of their coding sequences to the *A. thaliana* or *M. truncatula* genome are reported in Table 3 and Supplementary Table 2, respectively.

Discussion

Phenotypic variation in FT

The identification of genetic variation and development of genomic tools would promote achievement of improved persistency in future red clover breeding programs (Annicchiarico et al., 2015; Ravagnani et al., 2012). In this study we found considerable phenotypic variation in FT determined as LT50 values among 392 red clover accessions, which was partially associated with the origin of the material. A previous study in red clover of Nordic origin (Zanotto et al., 2021), found that FT (calculated as LT50 with the same method used here) was significantly correlated with winter survival at a continental location characterized by low freezing temperatures in the winter, showing that the LT50 method reveals relevant variation. However, the same study also identified a considerable genotype by environment interaction on winter survival, which underlines the complex nature of this trait. There is a tradeoff in experiments where the need to phenotype a large number of accessions meets the technical limitations of a methodology (LT50 estimated by probit analysis) whose level of error is necessarily dependent on the number of replicates available. Also, plants were not fully cold acclimated, and therefore, the level of FT expressed may not correspond to those that can be reached under natural conditions, and we cannot exclude accessions with different genetic background respond differently to acclimation conditions. Bertrand et al. (2020) found on average lower LT50 values in Canadian red clover acclimated under semi natural conditions over the whole autumn and winter than those found in this and a previous study in Nordic red clover (Zanotto et al., 2021). Despite these limitations, FT determined under controlled conditions was successfully used as a determinant for freezing tolerance and field winter survival in winter wheat (Gusta et al., 2001), perennial ryegrass (Hulke et al., 2008; Waldron et al., 1998) and white clover (Annicchiarico et al., 2001), although a more prolonged test, LD50 (lethal duration time for 50% kill), may be a better proxy for winter survival of winter canola in the field than LT50 (Waalen et al., 2011).

Linkage disequilibrium and relatedness among accessions

The rate of LD decay was very rapid in the studied material, with LD being close to the background level already at 1 kb distance between marker pairs. Rapid LD decay is typical in outbreeding species with a high genetic diversity among and within accessions such as red clover. The very low LD values that we calculated are in line with those found in red clover

(Jones et al., 2020) and perennial ryegrass (Keep et al., 2020) studies considering wide and diversified plant material. Jones et al. (2020) found that the rate of LD decay is proportional to the population size and that was faster among ecotypes than among cultivars. This may explain the much faster LD decay that we found across a large and diverse panel of accessions compared to those reported by De Vega et al. (2015), which only considered a single variety. Considering the average density of one HTP locus every 35 kb and the very low LD that we found, we would need a much higher number of GBS-loci for having a major part of the QTLs associated with FT to be in linkage with at least one of these GBS-loci.

The presence of structure in our material, as revealed by the PCA justifies the use of kinship correction in the GWAS model, which otherwise will likely overestimate the phenotypic effect of QTLs. The expected average diagonal value of a synthetic populations would be 1.0 (Cericola et al. 2018). The higher values that we found may be due to some level of inbreeding within the accessions used in this study and could also be affected by the imputation based on simple mean method, which tends to homogenize the genotypes of accessions around the mean values, therefore increasing the genetic similarity among accessions. On the other hand, the inclusion of kinship in the GWAS model, as explained by Keep at al. (2020) is likely to induce an underestimation of the number of loci and the total phenotypic variance explained.

Furthermore, the filtering strategy that we chose for the SNP and HTP data excludes several alleles with very low MAF (<0.5) that may be private of only few accessions and that may be associated to phenotypic variation with the trait of interest. As pointed out by Keep et al. (2020), rare alleles may have been responsible for a substantial part of the phenotypic variation of many traits and make their genetic determinism difficult to apprehend. However, because of their very low allele frequency, these loci would probably not be detected by the GWAS anyways if not removed. Therefore, the strategy that we choose in term of filtering and correction for relatedness among populations through the kinship effect in the GWAS model is justified, because it ensures that the associations that we found are indeed true with low risk of being false positives.

GWAS and phenotypic variance explained

Population allele frequency data from pools have been successfully used in genome wide association studies and for genomic selection in perennial and Italian ryegrass (Ashraf et al., 2016; Fè et al., 2016, 2015; Guo et al., 2018; Keep et al., 2020; Knorst et al., 2019) and for the

identification of loci associated to persistency and time to stem elongation in red clover (Ergon et al. 2019; Ergon et al., in prep). However, this is to our knowledge the first study that used GBS-generated allele frequency data from population pools to conduct SNP- and haplotype-based GWAS in red clover. Our analyses identified 13 independent GBS loci associated with FT. Only two of these loci were detected in both the SNP- and the haplotype-based GWAS, while the remaining were detected in only one of the analyses, confirming that haplotypes are able to reveal associations that are not detectable by single SNPs and vice versa, as also observed by (Bekele et al., 2018, Ergon et al. in preparetion). Thus, the different marker types can complement each other, revealing a larger total number of significant associations (Hamblin and Jannink, 2011; Lorenz et al., 2010).

However, because of the low marker density that we obtained relative to the low LD observed (see above section on LD) only a limited part of the phenotypic variation available within our material could be explained by the significant markers. Furthermore, the real phenotypic variance explained by the markers detected by the GWAS was much lower after the kinship effect was included in the regression models. Perhaps not surprisingly the markers which were most strongly affected by the kinship (i.e. those with the largest difference in R^2-values estimated in regressions that did and did not account for kinship) were also those which showed the strongest difference in allele frequency (AF) of at least one geographic group of accessions compared to the others. For example, for SNP LG1_11608488 and haplotype LG1:11608425-11608558_03 accessions from southern Europe had a clearly different average AF than those from the other countries, suggesting that the former material has a very low frequency for an allele which likely provide some level of FT. The same was true for SNP LG3_13453807. On the other hand, for SNP LG1_10733810 and HTP LG1:10733751-10733981_31 accessions from Norway, Finland, Serbia, and some from Sweden had a different average allele frequency than the rest of the accessions, suggesting that they indeed are a source for a favorable allele for FT which has low frequency in the other groups. These aspects are to keep in mind in local breeding programs aiming at the improvement of FT, for which it may be useful to introduce material from other countries/areas as source of useful alleles.

Interestingly, both GBS loci which were shared across SNPs- and HTPs-based GWAS (SNP LG1_11608488 and haplotype LG1:11608425-11608558_03, and SNP LG1_10733810 and haplotype LG1:10733751-10733981_31) were among the loci with the highest R^2 from a regression model with LT50, which may suggest that these associations are particularly strong. However, they had also some of the highest difference between R^2 from regressions that did

and did not account for kinship, as explained above, being therefore affected by the relatedness among accessions.

Candidate genes

Because of the very rapid LD decay we considered only genes found within 0.5 kb distance from significant loci as candidate genes, since the likelihood of markers located further away being causative is low. Overall, we found a total of 10 GBS-loci significantly associated with variation if FT that were mapped within gene sequences of the red clover genome, and two more loci were found very close to a gene (< 0.5kb). Markers found within genes or promoters (gene-targeted markers, GTMs) (Poczai et al., 2013) were suggested to be more accurate in describing the potential genetic diversity for specific traits of interest than randomly located markers (van Tienderen et al., 2002). GTMs have a strong potential for use in marker assisted selection for specific traits. Methods to develop GTMs have been successfully applied in red clover (Li et al., 2019), and used to conduct population structure and genetic diversity studies (Osterman et al., 2021). Loci within or in close proximity of gene sequences, such as those that we identified, may be used in the future for the development of markers linked to variation in FT for breeding, genetic diversity studies and conservation purposes specifically targeting this trait. This will need further investigation to confirm the robustness and accuracy of their association with FT, and later their validation across a diversified panel of accessions, before application.

Among the genes tagged by significant SNPs and HTPs, four are coding for proteins with transporter function (TP_8902, Tp_36921, Tp_32086, Tp_2102) which may be involved in biochemical mechanisms requiring the transport of metabolites at the cellular levels linked to the stress response, such as that provoked by low freezing temperatures. However, one of them (Tp_8902) is coding for Membrane transporter D1 n protein in red clover while the best blast hit in both *A. thaliana* and *M. truncatula* encodes an inositol transporter 1 (INT1) protein. This protein was reported as being the only known protein that facilitates myo-inositol import from the vacuole into the cytoplasm in *A. thaliana* (Strobl et al., 2018). Myo-inositol is a cyclic sugar alcohol that is accumulated in different plants in response to abiotic stress (Sengupta et al., 2008). Inositol transporters were also reported to play important roles in transitional metabolism and various signaling pathways in plants (Zhou et al., 2021). Two other candidate genes (Tp_10429 and Tp_10399) codes for proteins of the peroxidase superfamily, which are involved in oxidative stress response. Plants exposed to low temperatures combined with light

are subjected to oxidative stress because of an energy imbalance in the photosystems (Huner et al., 1993). Gene TP_1704 is coding for a transducing/WD40 repeat-like superfamily protein, a vast class of proteins involved in several functions, among these protein–protein interactions, modulating a variety of cellular processes such as plant stress and hormone response (Xu et al., 2019). The best hits for gene TP_1704 in the *Arabidopsis thaliana* is coding for the WD40 Repeat Protein NEDD1, known as being involved in microtubule organization during cell division in this species (Zeng et al., 2009). Two genes (Tp_28991 and Tp_24108) are coding for proteins involved in DNA processing. Finally, genes Tp_34014 and Tp_2834 are coding for proteins with a hydrolytic function. The first, belongs to the alpha/beta-Hydrolases superfamily having a vast number of functions such as hydrolases, lyases, transferases, hormone precursors or transporters, chaperones or routers of other proteins (Lenfant et al., 2013). Interestingly, the best hit of the CDS of this gene in the *M. truncatula* genome is coding for an enzyme, caffeoylshikimate esterase, involved in the shikimate biosynthetic pathway and suggested as being essential for lignification in all plant species (Ha et al., 2016). It's worth nothing that other secondary metabolites of the flavonoids class synthesized through the shikimate pathway were reported as being involved in cold acclimation in Arabidopsis and red clover (Schulz et al., 2016 Zanotto et al. in press), therefore we can speculate that this gene may be involved in the synthesis of these secondary metabolites in consequence of low temperatures.

Acknowledgments

This research was funded by EUCLEG (EU Project No. 727312). We acknowledge the technical assistance from Øyvind Jørgensen, Liv Berge, Philippe Barre, Steven Maenhout, Lea Frey and all the other members of the EUCLEG consortium that helped with productive discussions and suggestions at the several meetings and workshops that were organized.

Author contributions

CG and RK were responsible for collecting and distributing red clover populations. LS was responsible for DNA extractions and genotyping. LS and TR designed the genotyping strategy. TR performed SNP and haplotype calling. ÅE designed the phenotyping strategy and SZ performed the phenotyping. SZ performed all analyses of phenotypic and marker data with supervision from MP and ÅE. SZ wrote the paper with supervision from ÅE. All authors critically revised the manuscript and approved the final version.

References

Abberton, M.T., Marshall, A.H., 2005. Progress in breeding perennial clovers for temperate agriculture. The Journal of Agricultural Science 143, 117–135. https://doi.org/10.1017/S0021859605005101

Annicchiarico, P., Collins, R.P., Fornasier, F., Rhodes, I., 2001. Variation in cold tolerance and spring growth among Italian white clover populations. Euphytica 122, 407–416. https://doi.org/10.1023/A:1012918100232

Annicchiarico, P., Barrett, B., Brummer, E.C., Julier, B., Marshall, A.H., 2015. Achievements and Challenges in Improving Temperate Perennial Forage Legumes. Critical Reviews in Plant Sciences 34, 327–380. https://doi.org/10.1080/07352689.2014.898462

Annicchiarico, P., Pagnotta, M.A., 2012. Agronomic value and adaptation across climatically contrasting environments of Italian red clover landraces and natural populations. Grass and Forage Science 67, 597–605. https://doi.org/10.1111/j.1365-2494.2012.00887.x

Ashraf, B.H., Jensen, J., Asp, T., Janss, L.L., 2014. Association studies using family pools of outcrossing crops based on allele-frequency estimates from DNA sequencing. Theoretical and Applied Genetics 127, 1331–1341. https://doi.org/10.1007/s00122-014-2300-4

Barton, K. (2020). MuMIn: Multi-Model Inference. R package version 1.43.17. https://CRAN.R-project.org/package=MuMIn

Bekele, W.A., Wight, C.P., Chao, S., Howarth, C.J., Tinker, N.A., 2018. Haplotype-based genotyping-by-sequencing in oat genome research. Plant Biotechnology Journal 16, 1452–1463. https://doi.org/10.1111/pbi.12888

Bélanger, G., Castonguay, Y., Bertrand, A., Dhont, C., Rochette, P., Couture, L., Drapeau, R., Mongrain, D., Chalifour, F.-P., Michaud, R., 2006. Winter damage to perennial forage crops in eastern Canada: Causes, mitigation, and prediction. Canadian Journal of Plant Science 86, 33–47. https://doi.org/10.4141/P04-171

Bertrand, A., Bipfubusa, M., Castonguay, Y., Rocher, S., Szopinska-Morawska, A., Papadopoulos, Y., Renaut, J., 2016. A proteome analysis of freezing tolerance in red clover (Trifolium pratense L.). BMC Plant Biology 16, 65. https://doi.org/10.1186/s12870-016-0751-2

Bertrand, A., Rocher, S., Claessens, A., Bipfubusa, M., Papadopoulos, Y., Castonguay, Y., 2020. Biochemical and molecular responses during overwintering of red clover populations recurrently selected for improved freezing tolerance. Plant Science 292, 110388. https://doi.org/10.1016/j.plantsci.2019.110388

Biazzi, E., Nazzicari, N., Pecetti, L., Brummer, E.C., Palmonari, A., Tava, A., Annicchiarico, P., 2017. Genome-Wide Association Mapping and Genomic Selection for Alfalfa (Medicago sativa) Forage Quality Traits. PLOS ONE 12, e0169234. https://doi.org/10.1371/journal.pone.0169234

Bonnafous, F., Duhnen, A., Gody, L., Guillaume, O., Mangin, B., Pegot-Espagnet, P., Segura, V., Vilhjalmsson, B.J., Mabire, C. and Flutre, T., (2019). mlmm.gwas: Pipeline for GWAS Using MLMM. R package version 1.0.6. https://CRAN.R-project.org/package=mlmm.gwas

Byrne, S., Czaban, A., Studer, B., Panitz, F., Bendixen, C., Asp, T., 2013. Genome Wide Allele Frequency Fingerprints (GWAFFs) of Populations via Genotyping by Sequencing. PLOS ONE 8, e57438. https://doi.org/10.1371/journal.pone.0057438

Cericola, F., Lenk, I., Fè, D., Byrne, S., Jensen, C.S., Pedersen, M.G., Asp, T., Jensen, J., Janss, L., 2018. Optimized Use of Low-Depth Genotyping-by-Sequencing for Genomic Prediction Among Multi-Parental Family Pools and Single Plants in Perennial Ryegrass (Lolium perenne L.). Front. Plant Sci. 9. https://doi.org/10.3389/fpls.2018.00369

De Vega, J.J., Ayling, S., Hegarty, M., Kudrna, D., Goicoechea, J.L., Ergon, Å., Rognli, O.A., Jones, C., Swain, M., Geurts, R., Lang, C., Mayer, K.F.X., Rössner, S., Yates, S., Webb, K.J., Donnison, I.S., Oldroyd, G.E.D., Wing, R.A., Caccamo, M., Powell, W., Abberton, M.T., Skøt, L., 2015. Red clover (*Trifolium pratense* L.) draft genome provides a platform for trait improvement. Scientific Reports 5, 17394. https://doi.org/10.1038/srep17394

Dias, P.M.B., Julier, B., Sampoux, J.-P., Barre, P., Dall'Agnol, M., 2008. Genetic diversity in red clover (Trifolium pratense L.) revealed by morphological and microsatellite (SSR) markers. Euphytica 160, 189–205. https://doi.org/10.1007/s10681-007-9534-z

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., Mitchell, S.E., 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PLoS One 6. https://doi.org/10.1371/journal.pone.0019379

Endelman, J.B., 2011. Ridge Regression and Other Kernels for Genomic Selection with R Package rrBLUP. The Plant Genome 4, 250–255. https://doi.org/10.3835/plantgenome2011.08.0024

Ergon, Å., Skøt, L., Sæther, V.E., Rognli, O.A., 2019. Allele Frequency Changes Provide Evidence for Selection and Identification of Candidate Loci for Survival in Red Clover (Trifolium pratense L.). Front. Plant Sci. 10. https://doi.org/10.3389/fpls.2019.00718

Fè, D., Ashraf, B.H., Pedersen, M.G., Janss, L., Byrne, S., Roulund, N., Lenk, I., Didion, T., Asp, T., Jensen, C.S., Jensen, J., 2016. Accuracy of Genomic Prediction in a Commercial Perennial Ryegrass Breeding Program. The Plant Genome 9, plantgenome2015.11.0110. https://doi.org/10.3835/plantgenome2015.11.0110

Fè, D., Cericola, F., Byrne, S., Lenk, I., Ashraf, B.H., Pedersen, M.G., Roulund, N., Asp, T., Janss, L., Jensen, C.S., Jensen, J., 2015. Genomic dissection and prediction of heading date in perennial ryegrass. BMC Genomics 16, 921. https://doi.org/10.1186/s12864-015-2163-3

Frame, J., Charlton, J.F.L., Laidlaw, A.S., 1997. Temperate forage Legumes.

Greene, S.L., Gritsenko, M., Vandemark, G., 2004. Relating Morphologic and RAPD Marker Varlation to Collection Site Environment in wild Populations of Red Clover (Trifolium Pratense L.). Genetic Resources and Crop Evolution 51, 643–653. https://doi.org/10.1023/B:GRES.0000024655.48989.ab

Greenwell, B. M., & Schubert Kabban, C. M. (2014). investr: An R Package for Inverse Estimation. R Journal, 6(1), 90–100. https://doi.org/10.32614/RJ-2014-009

Guo, X., Cericola, F., Fè, D., Pedersen, M.G., Lenk, I., Jensen, C.S., Jensen, J., Janss, L.L., 2018. Genomic Prediction in Tetraploid Ryegrass Using Allele Frequencies Based on Genotyping by Sequencing. Front. Plant Sci. 9. https://doi.org/10.3389/fpls.2018.01165

Gusta, L.V., O'Connor, B.J., Gao, Y.-P., Jana, S., 2001. A re-evaluation of controlled freeze-tests and controlled environment hardening conditions to estimate the winter survival potential of hardy winter wheats. Can. J. Plant Sci. 81, 241–246. https://doi.org/10.4141/P00-068

Ha, C.M., Escamilla-Trevino, L., Yarce, J.C.S., Kim, H., Ralph, J., Chen, F., Dixon, R.A., 2016. An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in

Medicago truncatula. The Plant Journal 86, 363–375. https://doi.org/10.1111/tpj.13177
Hamblin, M.T., Jannink, J.-L., 2011. Factors Affecting the Power of Haplotype Markers in Association Studies. The Plant Genome 4. https://doi.org/10.3835/plantgenome2011.03.0008
Helgadóttir, Á., Marum, P., Dalmannsdóttir, S., Daugstad, K., Kristjánsdóttir, T.A., Lunnan, T., 2008. Combining Winter Hardiness and Forage Yield in White Clover (Trifolium repens) Cultivated in Northern Environments. Ann Bot 102, 825–834. https://doi.org/10.1093/aob/mcn159
Helgadóttir Á., Frankow-Lindberg B. E., Seppänen M. M., Søegaard K. & Østrem L. (2014). European grasslands overview: Nordic region. In Hopkins et al (Eds.), EGF at 50: The future of European grasslands. Grassland Science in Europe. 19, 15–28.
Herrmann, D., Boller, B., Studer, B., Widmer, F., Kölliker, R., 2008. Improving Persistence in Red Clover: Insights from QTL Analysis and Comparative Phenotypic Evaluation. Crop Science 48, 269–277. https://doi.org/10.2135/cropsci2007.03.0143
Hulke, B.S., Watkins, E., Wyse, D.L., Ehlke, N.J., 2008. Freezing tolerance of selected perennial ryegrass (Lolium perenne L.) accessions and its association with field winterhardiness and turf traits. Euphytica 163, 131–141. https://doi.org/10.1007/s10681-007-9631-z
Huner, N.P.A., Öquist, G., Hurry, V.M., Krol, M., Falk, S., Griffith, M., 1993. Photosynthesis, photoinhibition and low temperature acclimation in cold tolerant plants. Photosynth Res 37, 19–39. https://doi.org/10.1007/BF02185436
Inostroza, L., Bhakta, M., Acuña, H., Vásquez, C., Ibáñez, J., Tapia, G., Mei, W., Kirst, M., Resende, M., Munoz, P., 2018. Understanding the Complexity of Cold Tolerance in White Clover using Temperature Gradient Locations and a GWAS Approach. The Plant Genome 11, 0. https://doi.org/10.3835/plantgenome2017.11.0096
Jensen, E.S., Peoples, M.B., Boddey, R.M., Gresshoff, P.M., Hauggaard-Nielsen, H., J.R. Alves, B., Morrison, M.J., 2012. Legumes for mitigation of climate change and the provision of feedstock for biofuels and biorefineries. A review. Agron. Sustain. Dev. 32, 329–364. https://doi.org/10.1007/s13593-011-0056-7
Jones, C., De Vega, J., Lloyd, D., Hegarty, M., Ayling, S., Powell, W., Skøt, L., 2020. Population structure and genetic diversity in red clover (*Trifolium pratense* L.) germplasm. Scientific Reports 10, 8364. https://doi.org/10.1038/s41598-020-64989-z
Keep, T., Sampoux, J.-P., Blanco-Pastor, J.L., Dehmer, K.J., Hegarty, M.J., Ledauphin, T., Litrico, I., Muylle, H., Roldán-Ruiz, I., Roschanski, A.M., Ruttink, T., Surault, F., Willner, E., Barre, P., 2020. High-Throughput Genome-Wide Genotyping To Optimize the Use of Natural Genetic Resources in the Grassland Species Perennial Ryegrass (Lolium perenne L.). G3: Genes, Genomes, Genetics 10, 3347–3364. https://doi.org/10.1534/g3.120.401491
Klimenko, I., Razgulayeva, N., Gau, M., Okumura, K., Nakaya, A., Tabata, S., Kozlov, N.N., Isobe, S., 2010. Mapping candidate QTLs related to plant persistency in red clover. Theor Appl Genet 120, 1253–1263. https://doi.org/10.1007/s00122-009-1253-5
Knorst, V., Byrne, S., Yates, S., Asp, T., Widmer, F., Studer, B., Kölliker, R., 2019. Pooled DNA sequencing to identify SNPs associated with a major QTL for bacterial wilt resistance in Italian ryegrass (Lolium multiflorum Lam.). Theoretical and Applied Genetics 132, 947–958. https://doi.org/10.1007/s00122-018-3250-z
Korte, A., Farlow, A., 2013. The advantages and limitations of trait analysis with GWAS: a review. Plant Methods 9, 29. https://doi.org/10.1186/1746-4811-9-29
Kouamé, C.N., Quesenberry, K.H., 1993. Cluster analysis of a world collection of red clover germplasm. Genet Resour Crop Evol 40, 39–47. https://doi.org/10.1007/BF00053463

Lenfant, N., Hotelier, T., Bourne, Y., Marchot, P., Chatonnet, A., 2013. Proteins with an alpha/beta hydrolase fold: Relationships between subfamilies in an ever-growing superfamily. Chemico-Biological Interactions, 11th International Meeting on Cholinesterases 203, 266–268. https://doi.org/10.1016/j.cbi.2012.09.003

Li, W., Riday, H., Riehle, C., Edwards, A., Dinkins, R., 2019. Identification of Single Nucleotide Polymorphism in Red Clover (Trifolium pratense L.) Using Targeted Genomic Amplicon Sequencing and RNA-seq. Front Plant Sci 10, 1257. https://doi.org/10.3389/fpls.2019.01257

Lin, C.-Y., Xing, G., Xing, C., 2012. Measuring linkage disequilibrium by the partial correlation coefficient. Heredity 109, 401–402. https://doi.org/10.1038/hdy.2012.54

Liu, X., Huang, M., Fan, B., Buckler, E.S., Zhang, Z., 2016. Iterative Usage of Fixed and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies. PLOS Genetics 12, e1005767. https://doi.org/10.1371/journal.pgen.1005767

Lorenz, A.J., Hamblin, M.T., Jannink, J.-L., 2010. Performance of Single Nucleotide Polymorphisms versus Haplotypes for Genome-Wide Association Analysis in Barley. PLOS ONE 5, e14079. https://doi.org/10.1371/journal.pone.0014079

Lüscher, A., Mueller-Harvey, I., Soussana, J.F., Rees, R.M., Peyraud, J.L., 2014. Potential of legume-based grassland–livestock systems in Europe: a review. Grass and Forage Science 69, 206–228. https://doi.org/10.1111/gfs.12124

Mangin, B., Siberchicot, A., Nicolas, S., Doligez, A., This, P., Cierco-Ayrolles, C., 2012. Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. Heredity 108, 285–291. https://doi.org/10.1038/hdy.2011.73

Marshall, A.H., Collins, R.P., Vale, J., Lowe, M., 2017. Improved persistence of red clover (Trifolium pratense L.) increases the protein supplied by red clover/grass swards grown over four harvest years. European Journal of Agronomy 89, 38–45. https://doi.org/10.1016/j.eja.2017.06.006

Osterman, J., Hammenhag, C., Ortiz, R., Geleta, M., 2021. Insights Into the Genetic Diversity of Nordic Red Clover (Trifolium pratense) Revealed by SeqSNP-Based Genic Markers. Front Plant Sci 12, 748750. https://doi.org/10.3389/fpls.2021.748750

Poczai, P., Varga, I., Laos, M., Cseh, A., Bell, N., Valkonen, J.P., Hyvönen, J., 2013. Advances in plant gene-targeted and functional markers: a review. Plant Methods 9, 6. https://doi.org/10.1186/1746-4811-9-6

Raineri, E., Ferretti, L., Esteve-Codina, A., Nevado, B., Heath, S., Pérez-Enciso, M., 2012. SNP calling by sequencing pooled samples. BMC Bioinformatics 13, 239. https://doi.org/10.1186/1471-2105-13-239

Ravagnani, A., Abberton, M.T., Skøt, L., 2012. Development of Genomic Resources in the Species of Trifolium L. and Its Application in Forage Legume Breeding. Agronomy 2, 116–131. https://doi.org/10.3390/agronomy2020116

Reckling, M., Bergkvist, G., Watson, C.A., Stoddard, F.L., Zander, P.M., Walker, R.L., Pristeri, A., Toncea, I., Bachinger, J., 2016. Trade-Offs between Economic and Environmental Impacts of Introducing Legumes into Cropping Systems. Front. Plant Sci. 7. https://doi.org/10.3389/fpls.2016.00669

Sato, S., Isobe, S., Asamizu, E., Ohmido, N., Kataoka, R., Nakamura, Y., Kaneko, T., Sakurai, N., Okumura, K., Klimenko, I., Sasamoto, S., Wada, T., Watanabe, A., Kohara, M., Fujishiro, T., Tabata, S., 2005. Comprehensive Structural Analysis of the Genome of Red Clover (Trifolium pratense L.). DNA Research 12, 301–364. https://doi.org/10.1093/dnares/dsi018

Schulz, E., Tohge, T., Zuther, E., Fernie, A.R., Hincha, D.K., 2016. Flavonoids are determinants of freezing tolerance and cold acclimation in Arabidopsis thaliana. Sci Rep 6. https://doi.org/10.1038/srep34027

Segura, V., Vilhjálmsson, B., Platt, A. et al. An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet 44, 825–830 (2012). https://doi.org/10.1038/ng.2314

Sengupta, S., Mukherjee, S., Basak, P., Majumder, A.L., 2015. Significance of galactinol and raffinose family oligosaccharide synthesis in plants. Front. Plant Sci. 6. https://doi.org/10.3389/fpls.2015.00656

Strobl, S.M., Kischka, D., Heilmann, I., Mouille, G., Schneider, S., 2018. The Tonoplastic Inositol Transporter INT1 From Arabidopsis thaliana Impacts Cell Elongation in a Sucrose-Dependent Way. Frontiers in Plant Science 9.

Therneau, T. M., (2020). coxme: Mixed Effects Cox Models. R package version 2.2-16. https://CRAN.R-project.org/package=coxme

Juliane Schafer, J., Opgen-Rhein, R., Zuber, V., Ahdesmaki, M., Silva, P.D. and Strimmer, K. (2021). corpcor: Efficient Estimation of Covariance and (Partial) Correlation. R package version 1.6.10. https://CRAN.R-project.org/package=corpcor

Schaumont, D., Veeckman, E., Jeugt, F.V. der, Haegeman, A., Glabeke, S. van, Bawin, Y., Lukasiewicz, J., Blugeon, S., Barre, P., Leyva-Pérez, M. de la O., Byrne, S., Dawyndt, P., Ruttink, T., 2022. Stack Mapping Anchor Points (SMAP): a versatile suite of tools for read-backed haplotyping. https://doi.org/10.1101/2022.03.10.483555

VanRaden, P.M., 2008. Efficient Methods to Compute Genomic Predictions. Journal of Dairy Science 91, 4414–4423. https://doi.org/10.3168/jds.2007-0980

van Tienderen, P.H., de Haan, A.A., van der Linden, C.G., Vosman, B., 2002. Biodiversity assessment using markers for ecologically important traits. Trends in Ecology & Evolution 17, 577–582. https://doi.org/10.1016/S0169-5347(02)02624-1

Visser, C.L.M. de, Schreuder, R., Stoddard, F., 2014. The EU's dependency on soya bean import for the animal feed industry and potential for EU produced alternatives. OCL 21, D407. https://doi.org/10.1051/ocl/2014021

Voisin, A.-S., Guéguen, J., Huyghe, C., Jeuffroy, M.-H., Magrini, M.-B., Meynard, J.-M., Mougel, C., Pellerin, S., Pelzer, E., 2014. Legumes for feed, food, biomaterials and bioenergy in Europe: a review. Agron. Sustain. Dev. 34, 361–380. https://doi.org/10.1007/s13593-013-0189-y

Waldron, B.L., Ehlke, N.J., Vellekson, D.J., White, D.B., 1998. Controlled Freezing as a Indirect Selection Method for Field Winterhardiness in Turf-Type Perennial Ryegrass. Crop Science 38, cropsci1998.0011183X003800030032x. https://doi.org/10.2135/cropsci1998.0011183X003800030032x

Wang, W.Y.S., Barratt, B.J., Clayton, D.G., Todd, J.A., 2005. Genome-wide association studies: theoretical and practical concerns. Nature Reviews Genetics 6, 109–118. https://doi.org/10.1038/nrg1522

Xu, X., Wan, W., Jiang, G., Xi, Y., Huang, H., Cai, J., Chang, Y., Duan, C.-G., Mangrauthia, S.K., Peng, X., Zhu, J.-K., Zhu, G., 2019. Nucleocytoplasmic Trafficking of the Arabidopsis WD40 Repeat Protein XIW1 Regulates ABI5 Stability and Abscisic Acid Responses. Molecular Plant 12, 1598–1611. https://doi.org/10.1016/j.molp.2019.07.001

Zanotto, S., Palmé, A., Helgadóttir, Á., Daugstad, K., Isolahti, M., Öhlund, L., Marum, P., Moen, M.A., Veteläinen, M., Rognli, O.A., Ergon, Å., n.d. Trait characterization of genetic resources reveals useful variation for the improvement of cultivated Nordic red clover. Journal of Agronomy and Crop Science 2021. https://doi.org/10.1111/jac.12487

Zeng, C.J.T., Lee, Y.-R.J., Liu, B., 2009. The WD40 Repeat Protein NEDD1 Functions in Microtubule Organization during Cell Division in Arabidopsis thaliana. The Plant Cell 21, 1129–1140. https://doi.org/10.1105/tpc.109.065953

Zhou, Y., Foka, I.C.K., Li, H., 2021. Tonoplast Inositol Transporters: Roles in Plant Abiotic Stress Response and Crosstalk with Other Signals. https://doi.org/10.20944/preprints202106.0515.v1

Table 1. Linkage disequilibrium (LD) at different distances along the seven red clover chromosomes. LD was calculated as the average squared partial correlations ($r^2{}_s$) between allele frequencies of SNP pairs at given distances ± 0.5 Kbp. The number of pairs for each chromosome and distance is given in parentheses.

Chromosome	1Mb	0.5Mb	10Kb	5Kb	2.5Kb	1kb	0.5kb*	Mean
Chr1	0.0049 (318)	0.0054 (258)	0.0081 (393)	0.0089 (129)	0.0103 (222)	0.0098 (261)	0.2074 (4446)	0.0048
Chr2	0.0047 (165)	0.0043 (377)	0.0068 (274)	0.0108 (102)	0.0155 (116)	0.0125 (240)	0.1680 (4144)	0.0048
Chr3	0.0043 (333)	0.0049 (251)	0.0070 (351)	0.0077 (287)	0.0087 (310)	0.0171 (278)	0.2018 (4661)	0.0048
Chr4	0.0045 (178)	0.0053 (310)	0.0068 (213)	0.0063 (115)	0.0119 (169)	0.0090 (261)	0.1637 (2810)	0.0050
Chr5	0.0055 (112)	0.0067 (163)	0.0058 (169)	0.0051 (77)	0.0062 (100)	0.0081 (334)	0.1089 (2455)	0.0063
Chr6	0.0048 (166)	0.0054 (113)	0.0061 (283)	0.0073 (154)	0.0100 (142)	0.0103 (249)	0.1385 (2407)	0.0052
Chr7	0.0048 (117)	0.0052 (185)	0.0081 (212)	0.0071 (205)	0.0078 (165)	0.0092 (310)	0.1662 (3161)	0.0049
Mean	0.0048 (198)	0.0053 (236)	0.0070 (270)	0.0076 (152)	0.0101(175)	0.0109 (276)	0.1649 (3436)	0.0048

*Average LD in the interval 0-0.5 kb distance between pairs of SNP markers

Table 2. Phenotypic variance in LT50 explained by significant markers identified in GWAS. Regression analysis was performed, using allele frequencies of significant markers as predictors of LT50, both with and without kinship accounted for. **(A)** R^2 for the regression models including all the significant SNPs or haplotype variants simultaneously. **(B)** R^2 for the regression model using single significant SNPs and haplotype variants.

A

Model	$\boldsymbol{R^2}$with kinship	$\boldsymbol{R^2}$no kinship	Number of loci
SNP	0.30	0.48	8
Haplotype	0.26	0.45	7[1]

B

SNP/haplotype	$\boldsymbol{R^2}$with kinship	$\boldsymbol{R^2}$no kinship
LG1_10733810	0.054	0.093
LG1_11608488	0.047	0.19
scaf_836_20601	0.046	0.10
scaf_1414_9867	0.043	0.01
LG1_20019082	0.020	0.08
scaf_153_142327	0.016	0.00057
LG6_12095124	0.014	0.000443
LG3_13453807	0.013	0.12
LG1:11608425-11608558_3[1,2]	0.047	0.19
LG1:10733751-10733981_31	0.046	0.09
LG2:26124549-26124674_01	0.036	0.15
LG1:4339801-4340029_24	0.036	0.03
LG3:13143867-13144034_06	0.035	0.12
scaf_670:26744-26998_19	0.031	0.08
scaf_28637:389-568_03	0.023	0.03

[1]One haplotype (in *italics*) had a p-value below the Bonferroni threshold of significance in the GWAS but were included in further analyses because was very close to being significant. [2]Numbers at the end of each haplotype correspond to the variant number within HTP loci.

Table 3. Genes found within ±0.5 kb flanking significant SNPs and haplotypes identified by the GWAS. Gene number, position and putative annotation are presented. For SNPs and haplotypes sharing the same positions, only the SNP name is indicated. Three of the scaffolds containing significant markers were very short and contained no genes.

SNP/haplotype[1]	Gene Number[2]	Position Mb (start..end)	Annotation[3]
1_10733810	34014	10733892..10735416	alpha/beta-Hydrolases superfamily protein ; ***-; AT2G47630.1
1_11608488	8902	11607937..11610123	Membrane transporter D1 n
1_20019082	36921	20017489..20020298	cationic amino acid transporter 5; IPR002293 (Amino acid/polyamine transporter I);
6_12095124	32086	12094236..12095454	sucrose transporter 4 ; *-*-; AT1G09960.1
3_13453807	2834	13447395..13454431	3-hydroxyisobutyryl-CoA hydrolase-like protein; IPR001753 (Crotonase superfamily); GO:0003824
2_26124549-26124674_01	2102	26122598..26132038	Chloride channel C; IPR001807 (Chloride channel, voltage gated), IPR002902 (Gnk2-homologous domain); GO:0005216 (ion channel activity), GO:0005247 (voltage-gated chloride channel activity), GO:0006821 (chloride transport), GO:0016020 (membrane), GO:0030554 (adenyl nucleotide binding), GO:0055085 (transmembrane transport)
1_4339801-4340029_24	10429	4337894..4340197	Peroxidase superfamily protein; IPR010255 (Haem peroxidase); GO:0004601 (peroxidase activity),

			GO:0006979 (response to oxidative stress), GO:0020037 (heme binding), GO:0055114 (oxidation-reduction process)
	10399	4340314..4342504	Peroxidase superfamily protein; IPR010255 (Haem peroxidase); GO:0004601 (peroxidase activity), GO:0006979 (response to oxidative stress), GO:0020037 (heme binding), GO:0055114 (oxidation-reduction process);
3_13143867-13144034_06	1704	13138241..13146815	Transducin/WD40 repeat-like superfamily protein; IPR015943 (WD40/YVTN repeat-like-containing domain); GO:0005515 (protein binding)
Scaffold_836_20601	28991	19246..26121	TP/DNA-binding protein; IPR003594 (Histidine kinase-like ATPase, ATP-binding domain); GO:0005524 (ATP binding);****;
Scaffold_153_142327	26084	139266..146660	Kinase interacting (KIP1-like) family protein; IPR011684 (KIP1-like)

[1]In bold: marker found within the gene sequence, in Italics: marker had p value slightly below the Bonferroni threshold of significance.[2]Gene number in the Legume information system. [3]Annotation from the Legume Information System (legumeinfo.org).

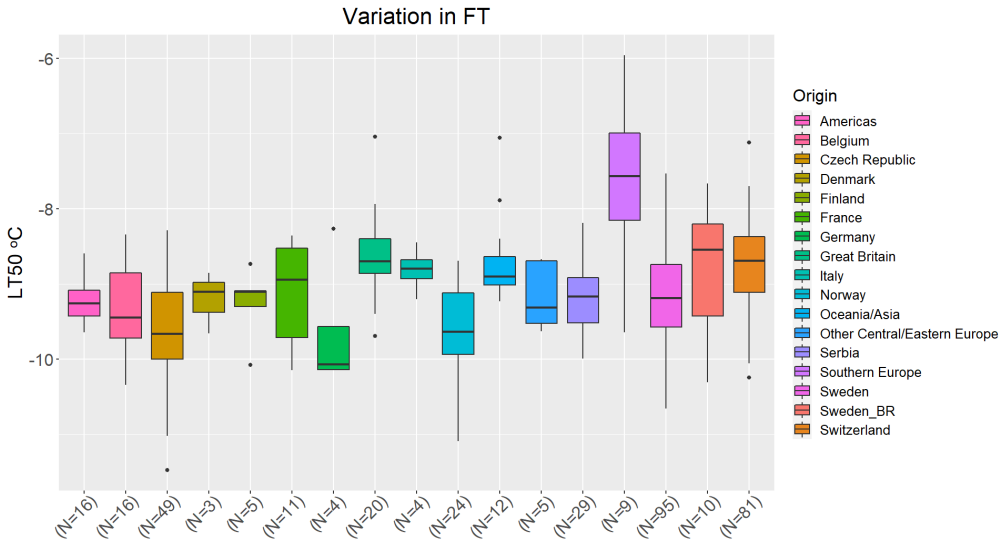
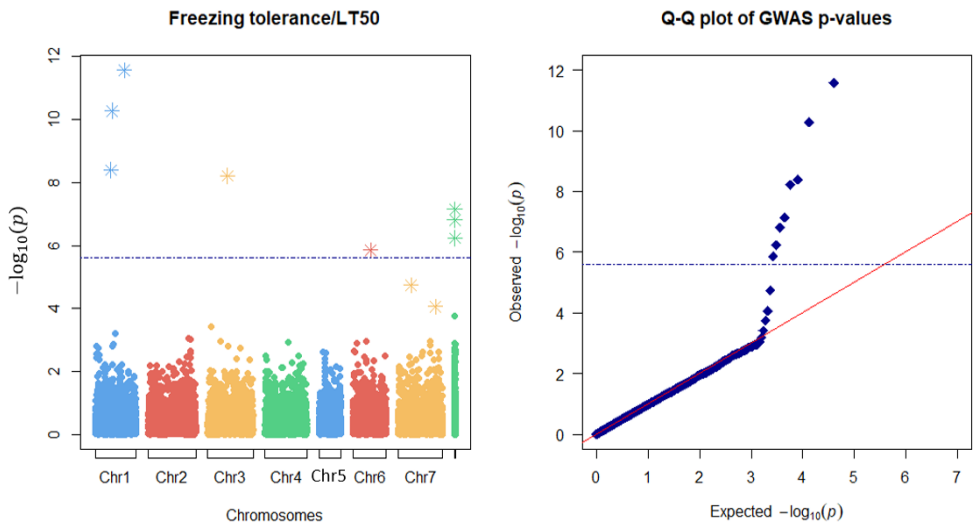


Figure 1. Boxplots showing the variation in freezing tolerance (LT50) among 393 red clover accessions from different countries or regions. F2 breeding populations (BR) from Sweden were kept as a separate group. N; number of accessions.

A



B

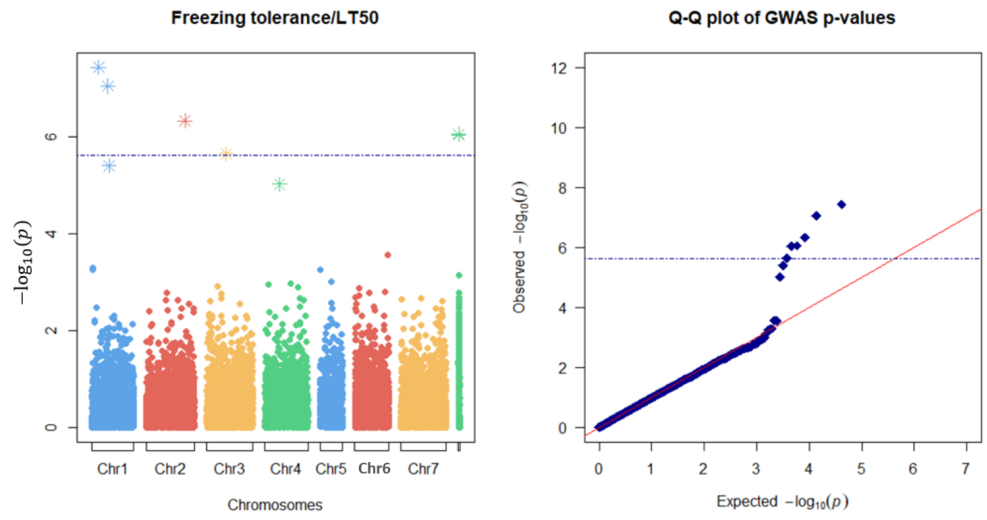


Figure 2. Manhattan and quantile-quantile (QQ) plots of p-values of the SNPs (**A**) and haplotypes (**B**) in the GWAS models with the lowest e-BIC values. Blue broken horizontal lines represent the Bonferroni threshold while asterisks represent the SNPs used as cofactors in the MLMM analyses.

Supplementary material

Supplementary Table 1: Information about the red clover accessions phenotyped for freezing tolerance in this study. For each accession the ID from the EUCLEG project, the name, the providing institution, the origin and the type of material is indicated.

Accession ID	Accession Name	Institution	Origin	Type
EUC_TP_001	Dimanche	INRA, FR	FRA	Cultivar
EUC_TP_002	Discovery	INRA, FR	FRA	Cultivar
EUC_TP_003	Formica	Agroscope, CH	CHE	Cultivar
EUC_TP_004	Milvus	Agroscope, CH	CHE	Cultivar
EUC_TP_005	Pavo	Agroscope, CH	CHE	Cultivar
EUC_TP_006	S586 AberClaret	IBERS, UK	GBR	Cultivar
EUC_TP_007	S592 AberChianti	IBERS, UK	GBR	Cultivar
EUC_TP_008	Gandalf	Graminor, NO	NOR	Cultivar
EUC_TP_009	Lea	Graminor, NO	NOR	Cultivar
EUC_TP_010	K 17	IKBKS, RS	SRB	Cultivar
EUC_TP_011	K 38	IKBKS, RS	SRB	Cultivar
EUC_TP_012	SW Ares	Lantmännen, SE	SWE	Cultivar
EUC_TP_013	Cyllene	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_014	Himalia	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_015	Metis	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_016	SANGRIA	RAGT2n, FR	FRA	Cultivar
EUC_TP_017	Saija	Boreal, FI	FIN	Cultivar
EUC_TP_018	Global	ILVO, BE	BEL	Cultivar
EUC_TP_019	Merviot	ILVO, BE	BEL	Cultivar
EUC_TP_020	Bonus	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_021	NGB1132	NordGen, SE	FIN	Landrace
EUC_TP_022	NGB1133	NordGen, SE	FIN	Landrace
EUC_TP_023	NGB1142	NordGen, SE	FIN	Landrace
EUC_TP_024	NGB14322	NordGen, SE	FIN	Wild
EUC_TP_025	NGB1730	NordGen, SE	DNK	Cultivar
EUC_TP_026	NGB2161	NordGen, SE	NOR	Landrace
EUC_TP_027	NGB2391	NordGen, SE	SWE	Landrace
EUC_TP_028	NGB2392	NordGen, SE	SWE	Landrace
EUC_TP_029	NGB2458	NordGen, SE	SWE	Landrace
EUC_TP_030	NGB2461	NordGen, SE	SWE	Landrace
EUC_TP_031	NGB2487	NordGen, SE	SWE	Landrace
EUC_TP_032	NGB2490	NordGen, SE	SWE	Landrace
EUC_TP_033	NGB2492	NordGen, SE	SWE	Landrace
EUC_TP_034	NGB4089	NordGen, SE	SWE	Landrace
EUC_TP_035	Grasslands Colenso	PGG Wrightson, NZ	NZL	Cultivar
EUC_TP_036	Sensation	PGG Wrightson, NZ	NZL	Cultivar
EUC_TP_037	Affoltern i.E._328	Agroscope, CH	CHE	Landrace
EUC_TP_038	Affoltern i.E._6	Agroscope, CH	CHE	Landrace
EUC_TP_039	Belpberg_225	Agroscope, CH	CHE	Landrace
EUC_TP_040	Belpberg_226	Agroscope, CH	CHE	Landrace
EUC_TP_041	Belpberg_229	Agroscope, CH	CHE	Landrace
EUC_TP_042	Bern_78	Agroscope, CH	CHE	Landrace
EUC_TP_043	Bowil_119	Agroscope, CH	CHE	Landrace
EUC_TP_044	Bruetten_3	Agroscope, CH	CHE	Landrace
EUC_TP_045	Bubikon_8	Agroscope, CH	CHE	Landrace
EUC_TP_046	Burgistein_300	Agroscope, CH	CHE	Landrace
EUC_TP_047	Columba	Agroscope, CH	CHE	Cultivar
EUC_TP_048	Corvus	Agroscope, CH	CHE	Cultivar

EUC_TP_049	Dafila	Agroscope, CH	CHE	Cultivar
EUC_TP_050	Frauenkappelen_86	Agroscope, CH	CHE	Landrace
EUC_TP_051	Goldbach i.E._167	Agroscope, CH	CHE	Landrace
EUC_TP_052	Huttwil_50	Agroscope, CH	CHE	Landrace
EUC_TP_053	Huttwil_60	Agroscope, CH	CHE	Landrace
EUC_TP_054	Koeniz_231	Agroscope, CH	CHE	Landrace
EUC_TP_055	Koeniz_279	Agroscope, CH	CHE	Landrace
EUC_TP_056	Krauchthal_176	Agroscope, CH	CHE	Landrace
EUC_TP_057	Lanzenhaeusern_291	Agroscope, CH	CHE	Landrace
EUC_TP_058	Lestris	Agroscope, CH	CHE	Cultivar
EUC_TP_059	Merula	Agroscope, CH	CHE	Cultivar
EUC_TP_060	Milonia	Agroscope, CH	CHE	Cultivar
EUC_TP_061	Monaco	Agroscope, CH	CHE	Cultivar
EUC_TP_062	Niederwangen_262	Agroscope, CH	CHE	Landrace
EUC_TP_063	Niederwangen_75	Agroscope, CH	CHE	Landrace
EUC_TP_064	Oberthal_121	Agroscope, CH	CHE	Landrace
EUC_TP_065	Oberuzwil_280	Agroscope, CH	CHE	Landrace
EUC_TP_066	Pastor	Agroscope, CH	CHE	Cultivar
EUC_TP_067	Riedbach_88	Agroscope, CH	CHE	Landrace
EUC_TP_068	Rueegsau_160	Agroscope, CH	CHE	Landrace
EUC_TP_069	Rueti_314	Agroscope, CH	CHE	Landrace
EUC_TP_070	Schmidigen_336	Agroscope, CH	CHE	Landrace
EUC_TP_071	Semperina	Agroscope, CH	CHE	Cultivar
EUC_TP_072	Signau_140	Agroscope, CH	CHE	Landrace
EUC_TP_073	Sumiswald_189	Agroscope, CH	CHE	Landrace
EUC_TP_074	Ueberstorf_294	Agroscope, CH	CHE	Landrace
EUC_TP_075	Ueberstorf_346	Agroscope, CH	CHE	Landrace
EUC_TP_076	Ufhusen_52	Agroscope, CH	CHE	Landrace
EUC_TP_077	Uttigen_2	Agroscope, CH	CHE	Landrace
EUC_TP_078	Wynigen_335	Agroscope, CH	CHE	Landrace
EUC_TP_079	Zaeziwil_125	Agroscope, CH	CHE	Landrace
EUC_TP_080	Zaeziwil_127	Agroscope, CH	CHE	Landrace
EUC_TP_081	Aa 3100	IBERS, UK	GBR	Old cultivar
EUC_TP_082	Aa 3148	IBERS, UK	GBR	Old cultivar
EUC_TP_083	Aa 3149	IBERS, UK	GBR	Old cultivar
EUC_TP_084	AA 32	IBERS, UK	GBR	Breeder's Line
EUC_TP_085	Aa 3459	IBERS, UK	GBR	Old cultivar
EUC_TP_086	Aa 4190	IBERS, UK	POL	Ecotype
EUC_TP_087	Aa 4292	IBERS, UK	CZE	Ecotype
EUC_TP_088	Aa 4298	IBERS, UK	SVK	Ecotype
EUC_TP_089	Aa 4351	IBERS, UK	BGR	Ecotype
EUC_TP_090	Aa 4379 Britta	IBERS, UK	GBR	Cultivar
EUC_TP_091	Aa 4390	IBERS, UK	PRT	Ecotype
EUC_TP_092	Aa 4400	IBERS, UK	GBR	Ecotype
EUC_TP_093	Aa 4444	IBERS, UK	ITA	Ecotype
EUC_TP_094	Aa 4516	IBERS, UK	ESP	Ecotype
EUC_TP_095	Aa 4519	IBERS, UK	ESP	Ecotype
EUC_TP_097	Aa 4528	IBERS, UK	ESP	Ecotype
EUC_TP_098	Aa 4939	IBERS, UK	NZL	Cultivar
EUC_TP_099	Aa 5674	IBERS, UK	ARG	Cultivar
EUC_TP_100	Aa 5675	IBERS, UK	ARG	Cultivar

EUC_TP_101	Aa 5676	IBERS, UK	ARG	Cultivar
EUC_TP_102	Aa 5677	IBERS, UK	ARG	Cultivar
EUC_TP_103	Aa 5678	IBERS, UK	ARG	Cultivar
EUC_TP_104	Aa 5746 Harmonie	IBERS, UK	GBR	Cultivar
EUC_TP_105	Aa4380 Altaswede	IBERS, UK	CAN	Cultivar
EUC_TP_106	S543 AberRuby	IBERS, UK	GBR	Cultivar
EUC_TP_107	Aa 4940 Broadway	IBERS, UK	NZL	Cultivar
EUC_TP_108	TP9525	Agroscope, CH	CHE	Breeding material
EUC_TP_109	TP9645	Agroscope, CH	CHE	Breeding material
EUC_TP_110	TP9445	Agroscope, CH	CHE	Breeding material
EUC_TP_111	TP9735	Agroscope, CH	CHE	Breeding material
EUC_TP_112	TP9315	Agroscope, CH	CHE	Breeding material
EUC_TP_113	Gumpensteiner Rotklee	HBLFA, AT	AUT	Old cultivar
EUC_TP_114	Cinnamon Plus	USDA, USA	USA	Cultivar
EUC_TP_115	DFRC11	USDA, USA	USA	Breeding material
EUC_TP_116	DFRC12	USDA, USA	USA	Breeding material
EUC_TP_117	DFRC13	USDA, USA	USA	Breeding material
EUC_TP_118	DFRC14	USDA, USA	USA	Breeding material
EUC_TP_119	DFRC15	USDA, USA	USA	Breeding material
EUC_TP_120	FF 9615	USDA, USA	USA	Cultivar
EUC_TP_121	Marathon	USDA, USA	USA	Cultivar
EUC_TP_122	Starfire I	USDA, USA	USA	Cultivar
EUC_TP_123	Starfire II	USDA, USA	USA	Cultivar
EUC_TP_124	GnRk0729	Graminor, NO	NOR	Breeding material
EUC_TP_125	GnRk0747	Graminor, NO	NOR	Breeding material
EUC_TP_126	KvRk0201	Graminor, NO	NOR	Breeding material
EUC_TP_127	LGRk8801	Graminor, NO	NOR	Breeding material
EUC_TP_128	LGRk9415	Graminor, NO	NOR	Breeding material
EUC_TP_129	Linus	Graminor, NO	NOR	Cultivar
EUC_TP_130	LøRk0286	Graminor, NO	NOR	Breeding material
EUC_TP_131	LøRk0287	Graminor, NO	NOR	Breeding material
EUC_TP_132	Linn	Graminor, NO	NOR	Cultivar
EUC_TP_134	LøRk9207	Graminor, NO	NOR	Cultivar
EUC_TP_135	LøRk9625	Graminor, NO	NOR	Breeding material
EUC_TP_136	LøRk9627	Graminor, NO	NOR	Breeding material
EUC_TP_137	LøRk9628	Graminor, NO	NOR	Breeding material
EUC_TP_138	LøRk9753	Graminor, NO	NOR	Breeding material
EUC_TP_139	VåRk0401	Graminor, NO	NOR	Breeding material
EUC_TP_140	VåRk0510	Graminor, NO	NOR	Breeding material
EUC_TP_141	VåRk0512	Graminor, NO	NOR	Breeding material
EUC_TP_142	VåRk0513	Graminor, NO	NOR	Breeding material
EUC_TP_143	VåRk0624	Graminor, NO	NOR	Breeding material
EUC_TP_144	VåRk0625	Graminor, NO	NOR	Breeding material
EUC_TP_145	K 39	IKBKS, RS	SRB	Cultivar
EUC_TP_146	Diplomat	DSV, DE	DEU	Cultivar
EUC_TP_147	SW 1479004	Lantmännen, SE	SWE	Breeding material
EUC_TP_148	SW 1578301	Lantmännen, SE	SWE	Breeding material
EUC_TP_149	SW 1678001	Lantmännen, SE	SWE	Breeding material
EUC_TP_150	SW RK1092	Lantmännen, SE	SWE	Breeding material
EUC_TP_151	SW RK1117	Lantmännen, SE	SWE	Breeding material
EUC_TP_152	SW RK1118	Lantmännen, SE	SWE	Breeding material

EUC_TP_153	SW RK1119	Lantmännen, SE	SWE	Breeding material
EUC_TP_154	SW RK1120	Lantmännen, SE	SWE	Breeding material
EUC_TP_155	SW RK1121	Lantmännen, SE	SWE	Breeding material
EUC_TP_156	SW RK1122	Lantmännen, SE	SWE	Breeding material
EUC_TP_157	SW RK1123	Lantmännen, SE	SWE	Breeding material
EUC_TP_158	SW RK1124	Lantmännen, SE	SWE	Breeding material
EUC_TP_159	SW RK1125	Lantmännen, SE	SWE	Breeding material
EUC_TP_160	SW RK1131	Lantmännen, SE	SWE	Breeding material
EUC_TP_161	SW RK1132	Lantmännen, SE	SWE	Breeding material
EUC_TP_162	SW RK1133	Lantmännen, SE	SWE	Breeding material
EUC_TP_163	SW RK1134	Lantmännen, SE	SWE	Breeding material
EUC_TP_164	SW Yngve	Lantmännen, SE	SWE	Cultivar
EUC_TP_165	SWÅ RK09093	Lantmännen, SE	SWE	Breeding material
EUC_TP_166	0780MP2	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_167	08102MP2	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_168	08102MP4	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_169	Callisto	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_170	Elara	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_171	Ganymed	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_172	Hegemon	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_173	Helike	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_174	HŽ 2004 80 – 01	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_175	JL 2n 07 80 MP 1	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_176	Kalyke	DLF Seeds, CZ	CZE	Cultivar
EUC_TP_177	TPD-05-11-18002	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_178	TPD-05-11-3087	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_179	TPD-05-11-3088	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_180	TPD-05-13-3080	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_181	TPD-05-13-3085	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_182	TPD-05-13-3091	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_183	TPD-05-14-1011	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_184	TPD-05-15-3127	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_185	TPD-05-15-3128	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_186	TPD-05-15-3129	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_187	TPD-05-16-3076	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_188	TPD-05-16-3146	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_189	TPD-05-16-3177	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_190	KARIM	RAGT2n, FR	FRA	Cultivar
EUC_TP_191	MISTRAL	RAGT2n, FR	FRA	Cultivar
EUC_TP_192	RAVVI	RAGT2n, FR	FRA	Cultivar
EUC_TP_193	TREVVIO	RAGT2n, FR	FRA	Cultivar
EUC_TP_194	Grasslands Hamua	AgResearch, NZ	NZL	Cultivar
EUC_TP_195	Grasslands Turoa	AgResearch, NZ	NZL	Cultivar
EUC_TP_196	Relish	AgResearch, NZ	NZL	Cultivar
EUC_TP_197	Ruby/Enterprise	AgResearch, NZ	NZL	Cultivar
EUC_TP_198	Natsuyu	Hokkaido Ag. Res., JP	JPN	Cultivar
EUC_TP_199	Ryokuyu	Hokkaido Ag. Res., JP	JPN	Cultivar
EUC_TP_200	NS-Mlava	IFVCNS, RS	SRB	Cultivar
EUC_TP_201	NS-Petnica	IFVCNS, RS	SRB	Cultivar
EUC_TP_202	NS-Sana	IFVCNS, RS	SRB	Cultivar
EUC_TP_203	Una(NS)	IFVCNS, RS	SRB	Cultivar

EUC_TP_204	Zoja (NS)	IFVCNS, RS	SRB	Cultivar
EUC_TP_205	Avisto	ILVO, BE	BEL	Cultivar
EUC_TP_206	Crossway	PGG Wrightson, NZ	NZL	Cultivar
EUC_TP_207	Lemmon	ILVO, BE	BEL	Cultivar
EUC_TP_208	Merkemse	ILVO, BE	BEL	Landrace
EUC_TP_209	Tp.12.12	ILVO, BE	BEL	Breeding material
EUC_TP_210	Tp.14.7	ILVO, BE	BEL	Breeding material
EUC_TP_211	Tandy	ILVO, BE	BEL	Cultivar
EUC_TP_212	Agil	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_214	Brisk	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_215	Chlumecký	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_218	Feng	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_219	Garant	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_223	Respect	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_224	Slavín	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_225	Slavoj	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_227	Spurt	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_228	Start	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_229	Suez	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_231	Trubadur	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_232	Van	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_233	Vendelín	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_234	Vltavín	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_236	Zefyr	Agricultural Res. Ltd., CZ	CZE	Cultivar
EUC_TP_237	Harmonie	NPZ, DE	DEU	Cultivar
EUC_TP_238	Regent	NPZ, DE	DEU	Cultivar
EUC_TP_239	NGB1736	NordGen, SE	DNK	Cultivar
EUC_TP_240	NGB2347	NordGen, SE	SWE	Cultivar
EUC_TP_241	NGB2349	NordGen, SE	SWE	Cultivar
EUC_TP_242	NGB2395	NordGen, SE	SWE	Landrace
EUC_TP_243	NGB2452	NordGen, SE	SWE	Landrace
EUC_TP_244	NGB2453	NordGen, SE	SWE	Landrace
EUC_TP_245	NGB2464	NordGen, SE	SWE	Landrace
EUC_TP_246	NGB2465	NordGen, SE	SWE	Landrace
EUC_TP_247	NGB2466	NordGen, SE	SWE	Landrace
EUC_TP_248	NGB2468	NordGen, SE	SWE	Landrace
EUC_TP_249	NGB2469	NordGen, SE	SWE	Landrace
EUC_TP_250	NGB2471	NordGen, SE	SWE	Landrace
EUC_TP_251	NGB2472	NordGen, SE	SWE	Landrace
EUC_TP_252	NGB2473	NordGen, SE	SWE	Landrace
EUC_TP_253	NGB2474	NordGen, SE	SWE	Landrace
EUC_TP_254	NGB2475	NordGen, SE	SWE	Landrace
EUC_TP_255	NGB2476	NordGen, SE	SWE	Landrace
EUC_TP_256	NGB2477	NordGen, SE	SWE	Landrace
EUC_TP_257	NGB2481	NordGen, SE	SWE	Landrace
EUC_TP_258	NGB2482	NordGen, SE	SWE	Landrace
EUC_TP_259	NGB2494	NordGen, SE	SWE	Landrace
EUC_TP_260	NGB2495	NordGen, SE	SWE	Landrace
EUC_TP_261	NGB2569	NordGen, SE	SWE	Landrace
EUC_TP_262	NGB2598	NordGen, SE	SWE	Landrace
EUC_TP_263	NGB2599	NordGen, SE	SWE	Landrace

EUC_TP_264	NGB2600	NordGen, SE	SWE	Landrace
EUC_TP_265	NGB2739	NordGen, SE	SWE	Cultivar
EUC_TP_266	NGB2740	NordGen, SE	SWE	Cultivar
EUC_TP_267	NGB2742	NordGen, SE	SWE	Cultivar
EUC_TP_268	NGB2745	NordGen, SE	SWE	Cultivar
EUC_TP_269	NGB2746	NordGen, SE	SWE	Cultivar
EUC_TP_270	NGB2747	NordGen, SE	SWE	Cultivar
EUC_TP_271	NGB2748	NordGen, SE	SWE	Cultivar
EUC_TP_272	NGB2749	NordGen, SE	SWE	Cultivar
EUC_TP_273	NGB2750	NordGen, SE	SWE	Cultivar
EUC_TP_274	NGB2751	NordGen, SE	SWE	Cultivar
EUC_TP_275	NGB4126	NordGen, SE	DNK	Cultivar
EUC_TP_277	NGB7510	NordGen, SE	SWE	Cultivar
EUC_TP_278	NGB9966	NordGen, SE	SWE	Landrace
EUC_TP_279	Diadem	INRA, FR	FRA	Cultivar
EUC_TP_280	Diper	INRA, FR	FRA	Old cultivar
EUC_TP_281	Diplo	INRA, FR	FRA	Cultivar
EUC_TP_282	Kindia	INRA, FR	FRA	Cultivar
EUC_TP_283	Affoltern i.E._325	Agroscope, CH	CHE	Landrace
EUC_TP_284	Bern_76	Agroscope, CH	CHE	Landrace
EUC_TP_285	Bigenthal_163	Agroscope, CH	CHE	Landrace
EUC_TP_286	Biglen_352	Agroscope, CH	CHE	Landrace
EUC_TP_287	Englisberg_249	Agroscope, CH	CHE	Landrace
EUC_TP_288	Grossdietwil_21	Agroscope, CH	CHE	Landrace
EUC_TP_289	Haeusernmoos_333	Agroscope, CH	CHE	Landrace
EUC_TP_290	Koeniz_239	Agroscope, CH	CHE	Landrace
EUC_TP_291	Koeniz_247	Agroscope, CH	CHE	Landrace
EUC_TP_292	Lauperswil_138	Agroscope, CH	CHE	Landrace
EUC_TP_293	MontCalme	Agroscope, CH	CHE	Old cultivar
EUC_TP_294	Neuenegg_340	Agroscope, CH	CHE	Landrace
EUC_TP_295	Niederscherli_273	Agroscope, CH	CHE	Landrace
EUC_TP_296	Oberbottigen_7	Agroscope, CH	CHE	Landrace
EUC_TP_297	Oberoenz_321	Agroscope, CH	CHE	Landrace
EUC_TP_298	Oeschenbach_330	Agroscope, CH	CHE	Landrace
EUC_TP_299	Renova	Agroscope, CH	CHE	Old cultivar
EUC_TP_300	Riggisberg_318	Agroscope, CH	CHE	Landrace
EUC_TP_301	Rueedisbach_332	Agroscope, CH	CHE	Landrace
EUC_TP_302	Rüttinova	Agroscope, CH	CHE	Old cultivar
EUC_TP_303	Schmitten_5	Agroscope, CH	CHE	Landrace
EUC_TP_304	Signau_154	Agroscope, CH	CHE	Landrace
EUC_TP_305	Wasen i.E._199	Agroscope, CH	CHE	Landrace
EUC_TP_306	Weier i.E._327	Agroscope, CH	CHE	Landrace
EUC_TP_307	Aa 3090	IBERS, UK	GBR	Old cultivar
EUC_TP_308	Aa 3108	IBERS, UK	GBR	Old cultivar
EUC_TP_310	Aa 4189	IBERS, UK	POL	Ecotype
EUC_TP_311	Aa 4297	IBERS, UK	CZE	Ecotype
EUC_TP_312	Aa 4398	IBERS, UK	GBR	Ecotype
EUC_TP_313	Aa 4403	IBERS, UK	GBR	Ecotype
EUC_TP_315	Aa 4445	IBERS, UK	ITA	Ecotype
EUC_TP_316	Aa 4448	IBERS, UK	ITA	Ecotype
EUC_TP_317	Aa 4456	IBERS, UK	ITA	Ecotype

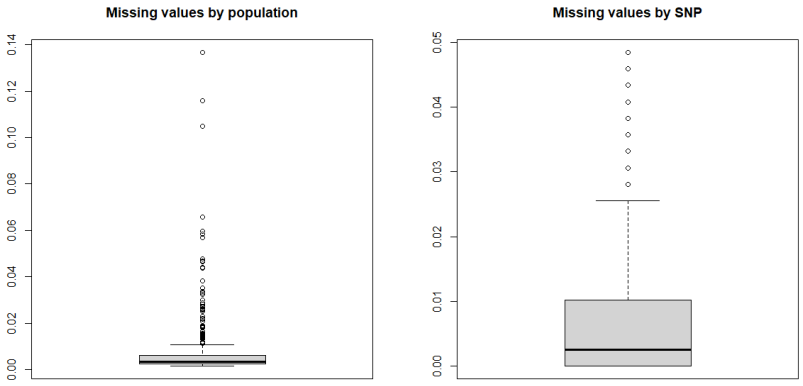
EUC_TP_318	Aa 4515	IBERS, UK	ESP	Ecotype
EUC_TP_319	Aa 4520	IBERS, UK	ESP	Ecotype
EUC_TP_320	Aa 4525	IBERS, UK	ESP	Ecotype
EUC_TP_321	Aa 4527	IBERS, UK	ESP	Ecotype
EUC_TP_322	Aa 4529	IBERS, UK	ESP	Ecotype
EUC_TP_324	Aa 4593	IBERS, UK	GBR	Cultivar (Landrace)
EUC_TP_325	Aa 4934	IBERS, UK	GBR	Breeder's Line
EUC_TP_326	Aa 4936	IBERS, UK	GBR	Breeder's Line
EUC_TP_327	Aa 4937	IBERS, UK	GBR	Breeder's Line
EUC_TP_328	Aa 5417	IBERS, UK	GBR	Breeder's Line
EUC_TP_329	SW 1479001	Lantmännen, SE	SWE	Breeding material
EUC_TP_330	SW 1479002	Lantmännen, SE	SWE	Breeding material
EUC_TP_331	SW 1479003	Lantmännen, SE	SWE	Breeding material
EUC_TP_332	SW 1678002	Lantmännen, SE	SWE	Breeding material
EUC_TP_333	SW 1678003	Lantmännen, SE	SWE	Breeding material
EUC_TP_334	SW 1678004	Lantmännen, SE	SWE	Breeding material
EUC_TP_335	SW 1678401	Lantmännen, SE	SWE	Breeding material
EUC_TP_336	SW 1678402	Lantmännen, SE	SWE	Breeding material
EUC_TP_337	SW 1678403	Lantmännen, SE	SWE	Breeding material
EUC_TP_338	SW RK1095	Lantmännen, SE	SWE	Breeding material
EUC_TP_339	SW RK1096	Lantmännen, SE	SWE	Breeding material
EUC_TP_340	SW RK1097	Lantmännen, SE	SWE	Breeding material
EUC_TP_341	SW RK1102	Lantmännen, SE	SWE	Breeding material
EUC_TP_342	SW RK1160	Lantmännen, SE	SWE	Breeding material
EUC_TP_343	SW RK1161	Lantmännen, SE	SWE	Breeding material
EUC_TP_344	SW RK1162	Lantmännen, SE	SWE	Breeding material
EUC_TP_345	SW RK1164	Lantmännen, SE	SWE	Breeding material
EUC_TP_346	SWA 1376104	Lantmännen, SE	SWE	Breeding material
EUC_TP_347	SWA 1376105	Lantmännen, SE	SWE	Breeding material
EUC_TP_348	SWA 1476014	Lantmännen, SE	SWE	Breeding material
EUC_TP_349	SWA 1476016	Lantmännen, SE	SWE	Breeding material
EUC_TP_350	SWA 1476017	Lantmännen, SE	SWE	Breeding material
EUC_TP_351	SWA 1476019	Lantmännen, SE	SWE	Breeding material
EUC_TP_352	SWA 1575301	Lantmännen, SE	SWE	Breeding material
EUC_TP_353	SWA 1575302	Lantmännen, SE	SWE	Breeding material
EUC_TP_354	SWA 1575304	Lantmännen, SE	SWE	Breeding material
EUC_TP_355	SWA 1575305	Lantmännen, SE	SWE	Breeding material
EUC_TP_356	SWA 1575306	Lantmännen, SE	SWE	Breeding material
EUC_TP_357	SWA 1575307	Lantmännen, SE	SWE	Breeding material
EUC_TP_358	SWA 1575308	Lantmännen, SE	SWE	Breeding material
EUC_TP_359	SWA 1575309	Lantmännen, SE	SWE	Breeding material
EUC_TP_360	SWA 1575312	Lantmännen, SE	SWE	Breeding material
EUC_TP_361	SWA 1576005	Lantmännen, SE	SWE	Breeding material
EUC_TP_362	SWA 1675205	Lantmännen, SE	SWE	Breeding material
EUC_TP_363	SWA 1675206	Lantmännen, SE	SWE	Breeding material
EUC_TP_364	SWA 1675207	Lantmännen, SE	SWE	Breeding material
EUC_TP_365	SWA 1675208	Lantmännen, SE	SWE	Breeding material
EUC_TP_366	SWA 1675212	Lantmännen, SE	SWE	Breeding material
EUC_TP_367	TPD-05-04-3000	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_368	TPD-05-11-3007	DLF Seeds, CZ	CZE	Breeding material
EUC_TP_369	TPD-05-12-3018	DLF Seeds, CZ	CZE	Breeding material

EUC_TP_370	Avala (NS)	IFVCNS, RS	SRB	Cultivar
EUC_TP_371	BL-1-Banja Luka	IFVCNS, RS	SRB	Breeding material
EUC_TP_372	BL-3-Banja Luka	IFVCNS, RS	SRB	Breeding material
EUC_TP_373	BL-4-Banja Luka	IFVCNS, RS	SRB	Breeding material
EUC_TP_374	BL-5-Banja Luka	IFVCNS, RS	SRB	Breeding material
EUC_TP_375	D-1	IFVCNS, RS	SRB	Breeding material
EUC_TP_376	D-10	IFVCNS, RS	SRB	Breeding material
EUC_TP_377	D-2	IFVCNS, RS	SRB	Breeding material
EUC_TP_378	D-3	IFVCNS, RS	SRB	Breeding material
EUC_TP_379	D-4	IFVCNS, RS	SRB	Breeding material
EUC_TP_380	D-5	IFVCNS, RS	SRB	Breeding material
EUC_TP_381	D-6	IFVCNS, RS	SRB	Breeding material
EUC_TP_382	D-7	IFVCNS, RS	SRB	Breeding material
EUC_TP_383	D-8	IFVCNS, RS	SRB	Breeding material
EUC_TP_384	D-9	IFVCNS, RS	SRB	Breeding material
EUC_TP_385	M10-Kopaonik	IFVCNS, RS	SRB	Wild
EUC_TP_386	M11-Kopaonik	IFVCNS, RS	SRB	Wild
EUC_TP_387	M12-Kopaonik	IFVCNS, RS	SRB	Wild
EUC_TP_388	M13-Kopaonik	IFVCNS, RS	SRB	Wild
EUC_TP_389	M14-Kopaonik	IFVCNS, RS	SRB	Wild
EUC_TP_390	NS-Ravanica	IFVCNS, RS	SRB	Cultivar
EUC_TP_391	Broadway	PGG Wrightson, NZ	NZL	Cultivar
EUC_TP_392	Kontiki	DSV, DE	DEU	Cultivar
EUC_TP_393	Mercury	ILVO, BE	BEL	Cultivar
EUC_TP_394	Merian	ILVO, BE	BEL	Cultivar
EUC_TP_395	Oudenaerdse	ILVO, BE	BEL	Landrace
EUC_TP_396	Primus	ILVO, BE	BEL	Landrace
EUC_TP_397	Tp.08.4	ILVO, BE	BEL	Breeding material
EUC_TP_398	Tp.08.5	ILVO, BE	BEL	Breeding material
EUC_TP_399	Violetta	ILVO, BE	BEL	Cultivar
EUC_TP_400	Waesse	ILVO, BE	BEL	Landrace
EUC_TP_446	Affoltern i.E._186	Agroscope, CH	CHE	Landrace
EUC_TP_447	Arni b.Biglen_351	Agroscope, CH	CHE	Landrace
EUC_TP_449	Horgen_1	Agroscope, CH	CHE	Landrace
EUC_TP_454	Lanzenhaeusern_292	Agroscope, CH	CHE	Landrace
EUC_TP_456	Riggisberg_311	Agroscope, CH	CHE	Landrace
EUC_TP_660	LøRk0498	Graminor, NO	NOR	Breeding material
EUC_TP_661	SWA 1575303	Lantmännen, SE	SWE	Breeding material
EUC_TP_662	SWA 1576001	Lantmännen, SE	SWE	Breeding material

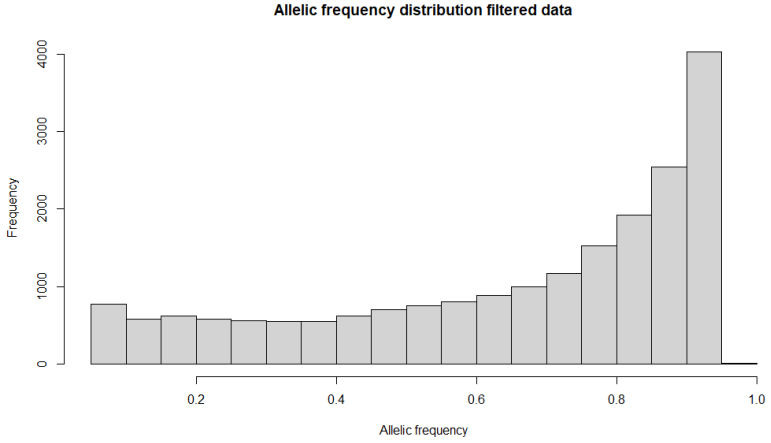
Supplementary Table 2. Best BLASTn-hits in *Arabidopsis thaliana* and *Medicago truncatula* (e-value<1 e-20) of the flanking gene positioned closest to markers that were significant in the GWAS. In parentheses: coverage (%) / identity (%).

Locus	Gene ID	Dist.[2]	Best BLASTn hit in *A. thaliana*	Best BLASTn hit in *M. truncatula*
LG1_10733810	tripr.gene34014	80	alpha/beta-Hydrolases superfamily protein (AT3G62860) (57/75)	caffeoylshikimate esterase (LOC11422042), mRNA (74/88)
LG1_11608488	tripr.gene8902	0	inositol transporter 1 (INT1), mRNA (71/70)	inositol transporter 1 (LOC11438245), mRNA (83/92)
LG1_20019082	tripr.gene36921	0	amino acid transporter 1 (AAT1), mRNA (70/66)	cationic amino acid transporter 1 (LOC25482484), mRNA (83/89)
LG6_12095124	tripr.gene32086	0	.	sucrose transport protein SUC8 (LOC11443134), mRNA (18/74)
LG3_13453807	tripr.gene2834	0	beta-hydroxyisobutyryl-CoA hydrolase 1 (CHY1), mRNA (85/71)	3-hydroxyisobutyryl-CoA hydrolase 1 (LOC25500243), mRNA(100/91)
LG2:26124549-26124674/+01	tripr.gene2102	0	chloride channel C (CLC-C), mRNA (80/73)	chloride channel protein CLC-c (LOC11433847), mRNA (82/91)
LG1:4339801-4340029/+24	tripr.gene10429	0	Peroxidase superfamily protein (PRX52), mRNA (26/70)	cationic peroxidase 1 (LOC11428337), mRNA (92/85)
LG1:4339801-4340029/+24	tripr.gene10399	410	.	cationic peroxidase 1 (LOC11428337), mRNA (89/83)
LG3:13143867-13144034/+06	tripr.gene1704	0	Transducin/WD40 repeat-like superfamily protein (NEDD1), mRNA (44/74	protein NEDD1 (LOC11411297), mRNA (100/91)
scaf_836_20601	tripr.gene28991	0	ATP/DNA binding protein (AT3G48770), mRNA (36/67)	uncharacterized LOC11421221 (LOC11421221), mRNA (97/90)
scaf_153_142327	tripr.gene26084	0	kinase interacting (KIP1-like) family protein (NET1D), mRNA (8/80)	protein NETWORKED 1A (LOC25485160), transcript variant X2, mRNA (97/90)

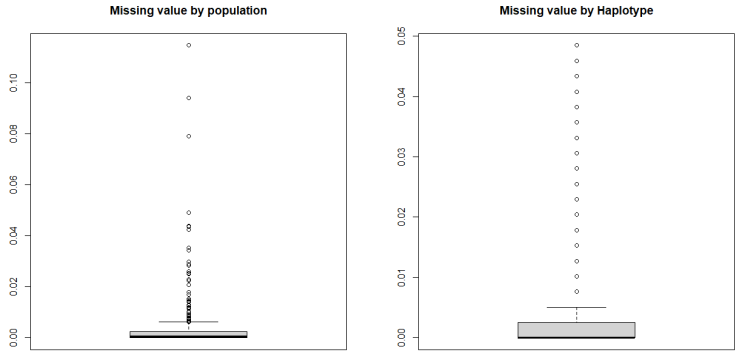
[1]Tp2.0, Legume Information System (legumeinfo.org). [2]Distance from significant marker (bp).



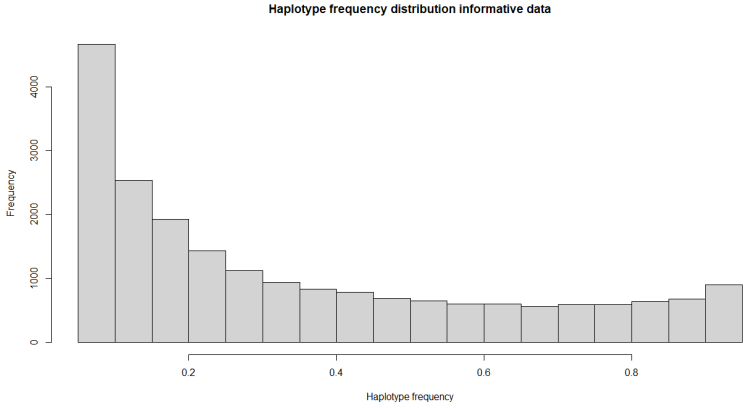
Supplementary figure 1: Boxplots showing the proportion of missing values by population and SNP, respectively, for the 392 red clover accessions having less than 80% missing values after filtering SNPs for completeness at 95% (max 5% missing values) and MAF>=0.05.



Supplementary figure 2: Histogram showing the reference allele frequency (RAF) distribution after filtering SNPs for completeness at 95% (max 5% missing values) and MAF>=0.05.

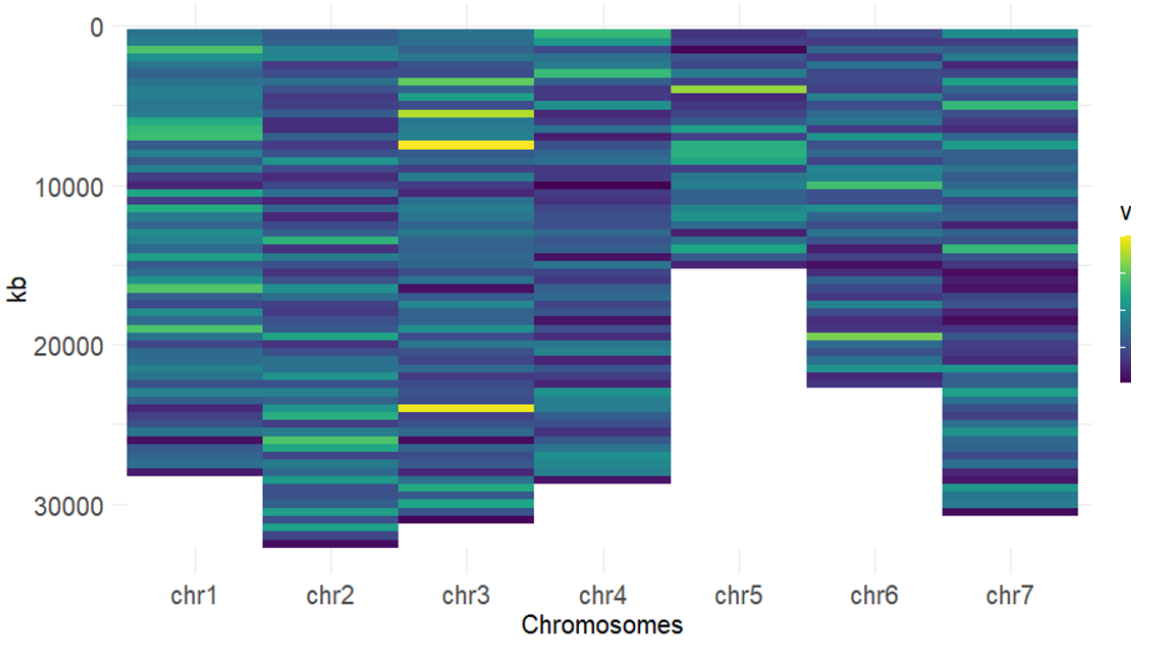


Supplementary figure 3: Boxplots showing the proportion of missing values by population and haplotype, respectively, for the 392 red clover accessions after filtering the haplotypes for completeness at 95% (max 5% missing values) and MAF>=0.05.

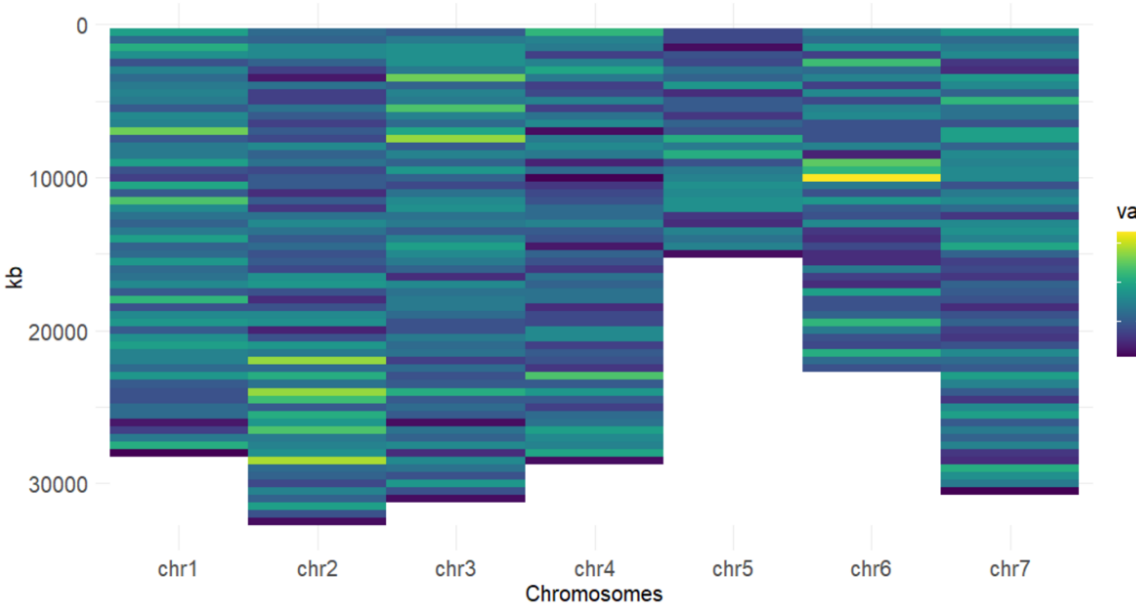


Supplementary figure 4: Histogram showing the haplotype variant frequency distribution after filtering for completeness at 95% (max 5% missing values) and MAF>=0.05.

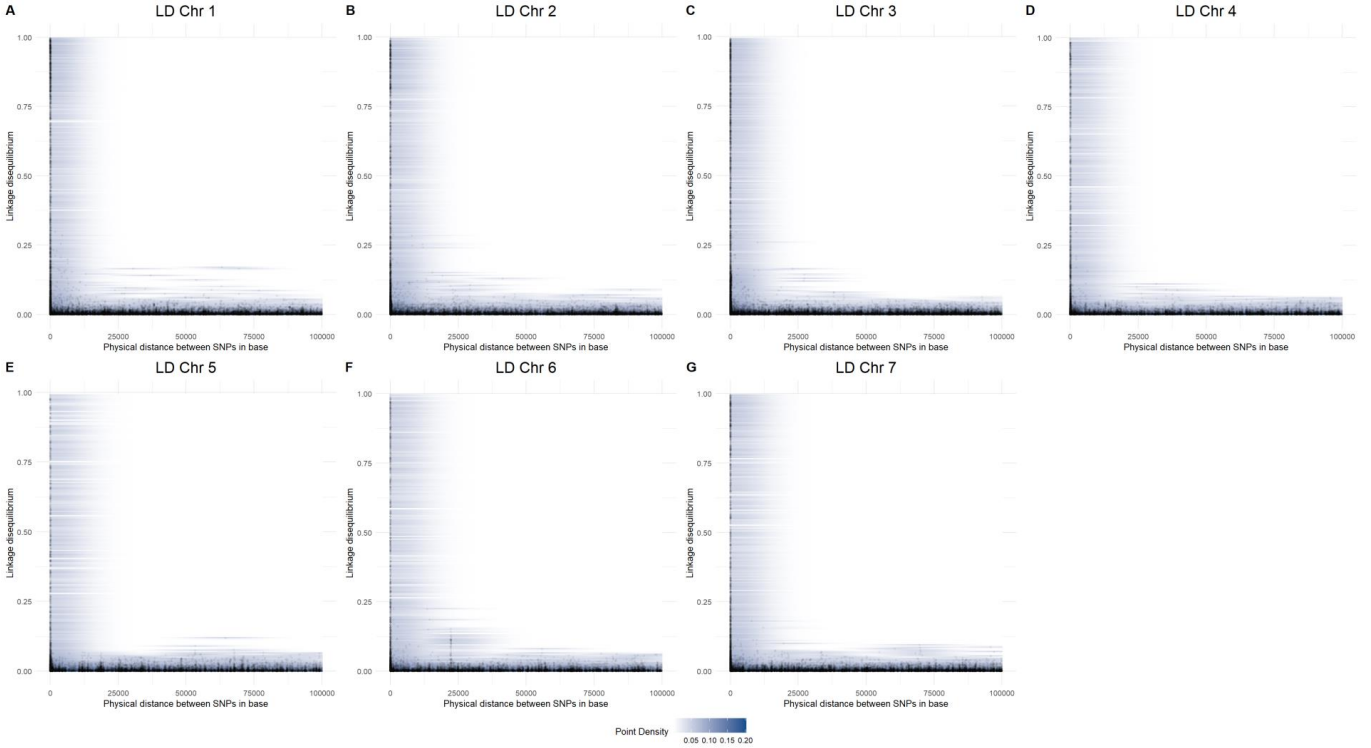
A


0
10000
20000
30000
kb
chr1
chr2
chr3
chr4
chr5
chr6
chr7
Chromosomes

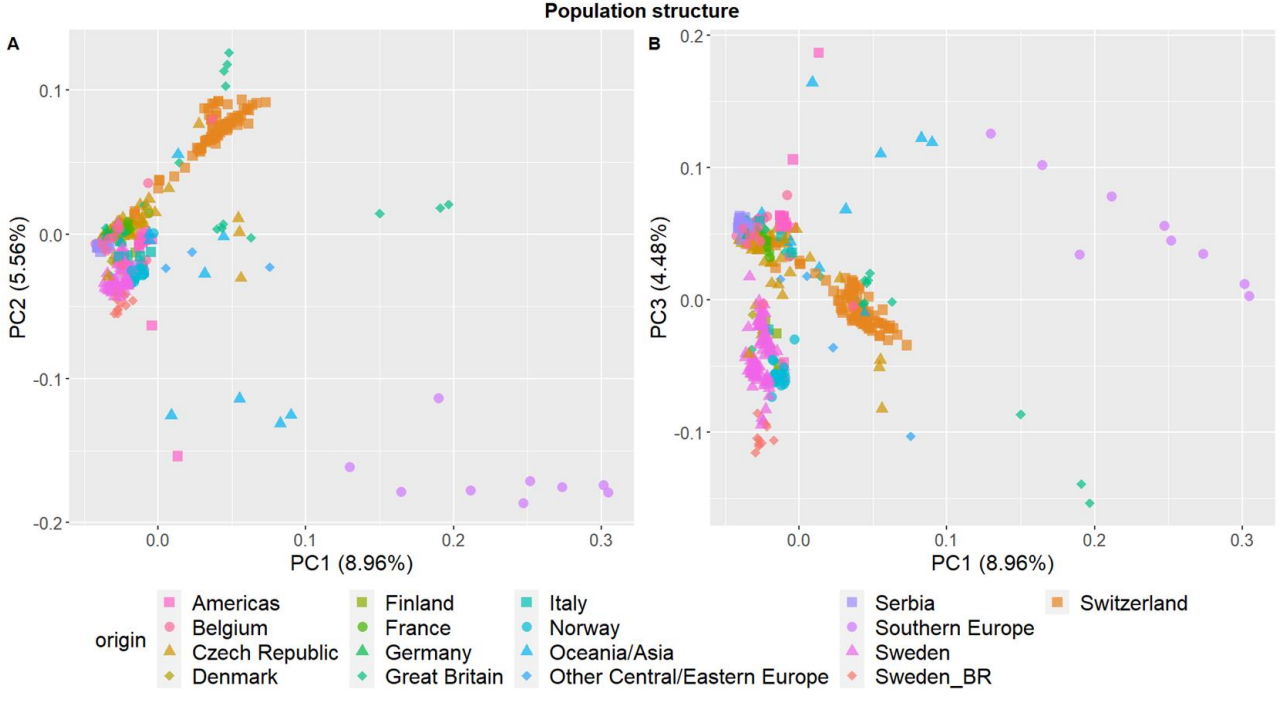
B



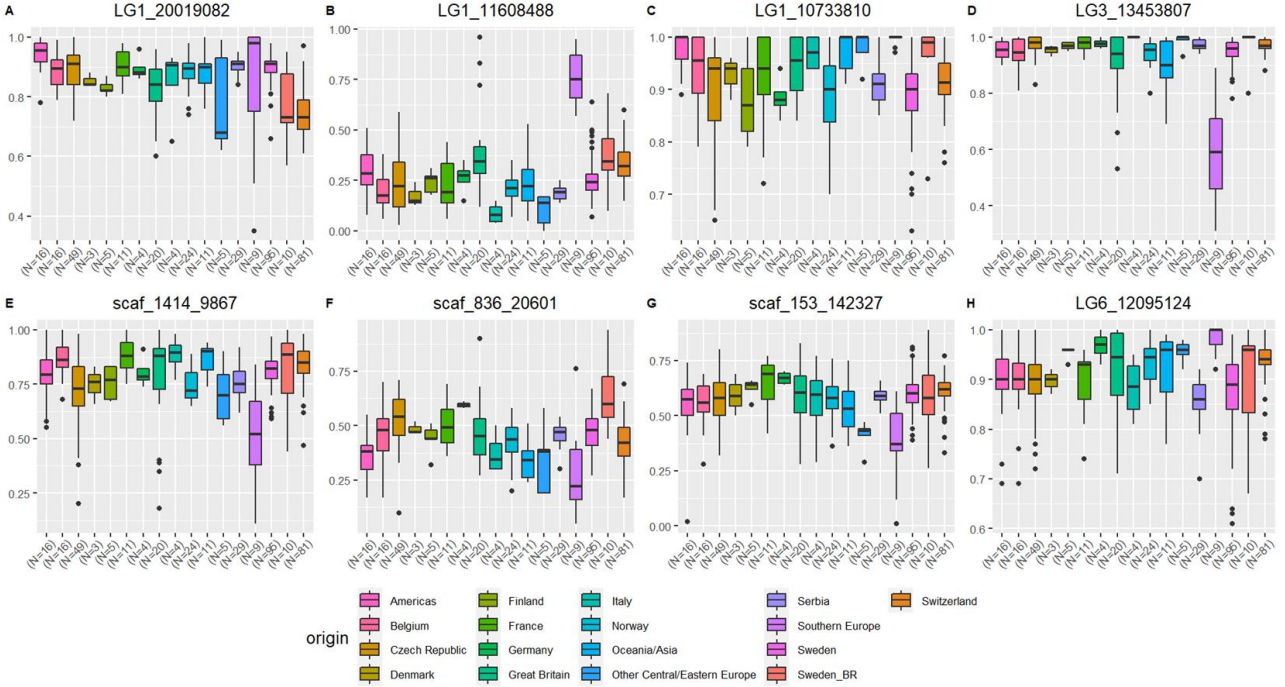
Supplementary Figure 5: Heat maps of each chromosome representing number of SNPs (**A**) and HTPs (**B**) within windows of 500 Kbp.

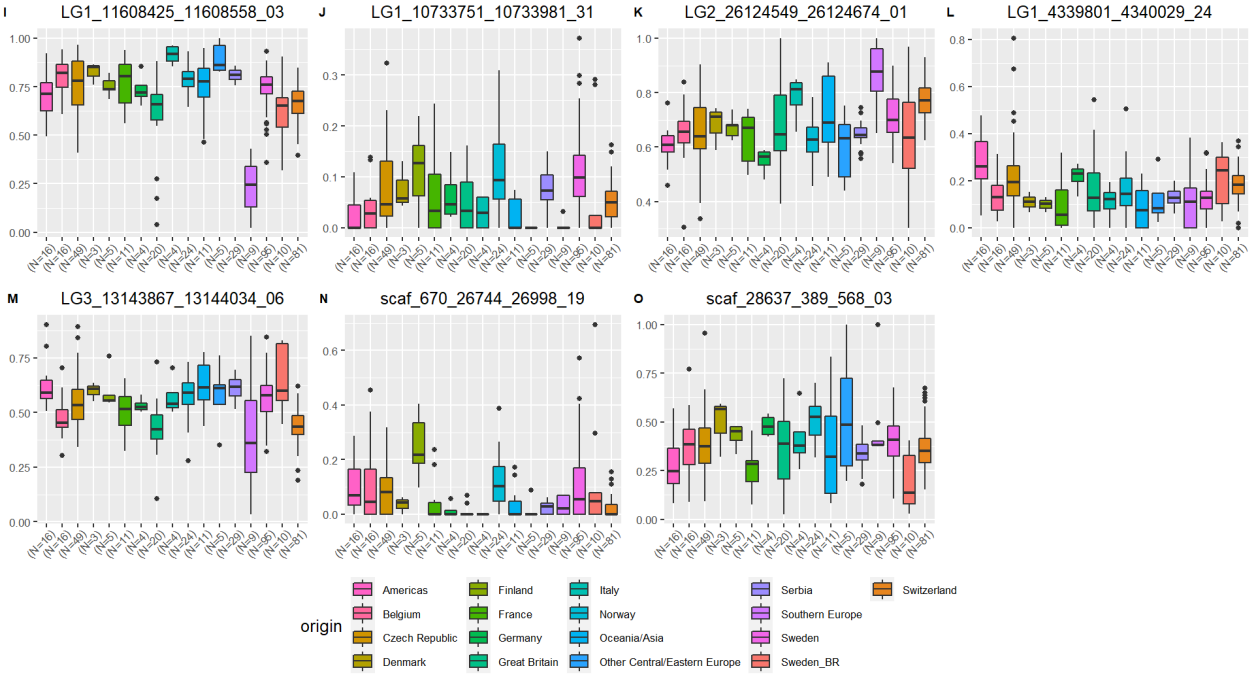


Supplementary Figure 6. LD in the seven chromosomes (**A-G**) as scatter plot of squared partial correlations between reference allele frequencies of pairs of SNP markers and genomic distance between SNP markers in bp.



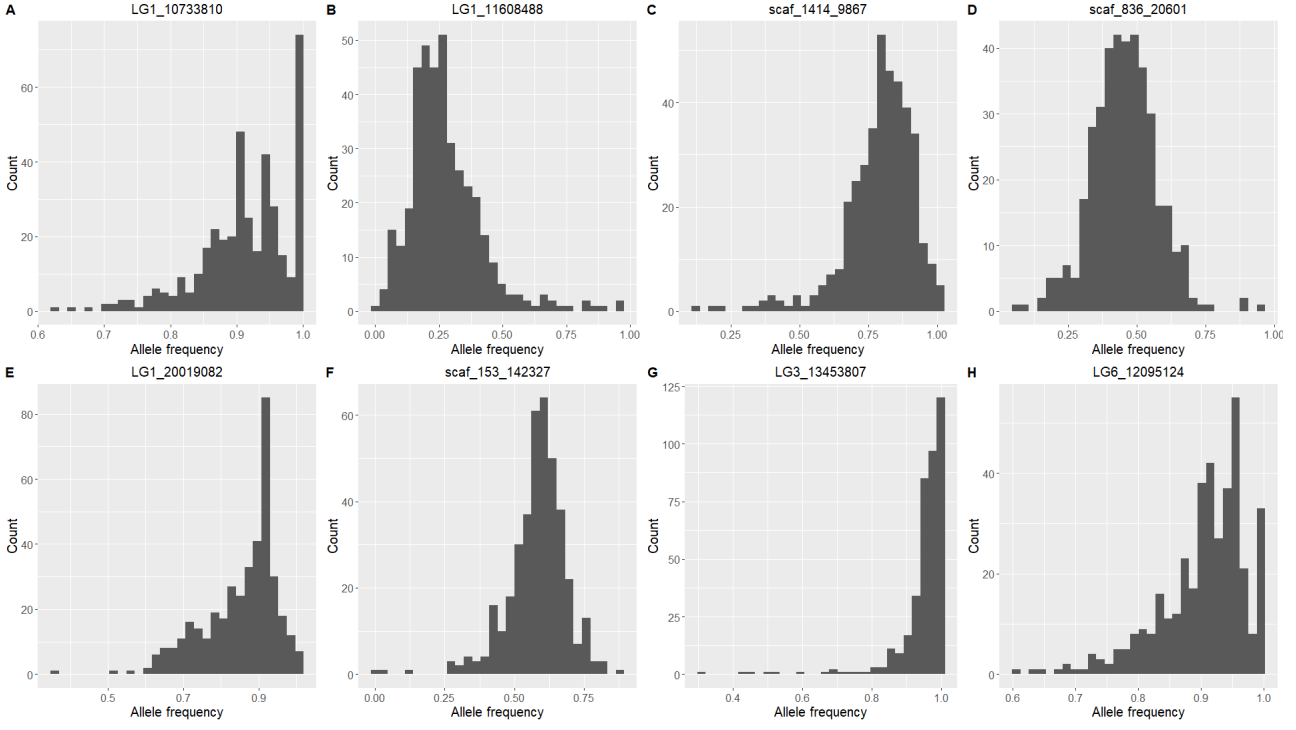
Supplementary figure 7: PCA showing the population structure among the 392 accessions based on filtered and imputed SNPs. Accessions are grouped by country/area of origin.

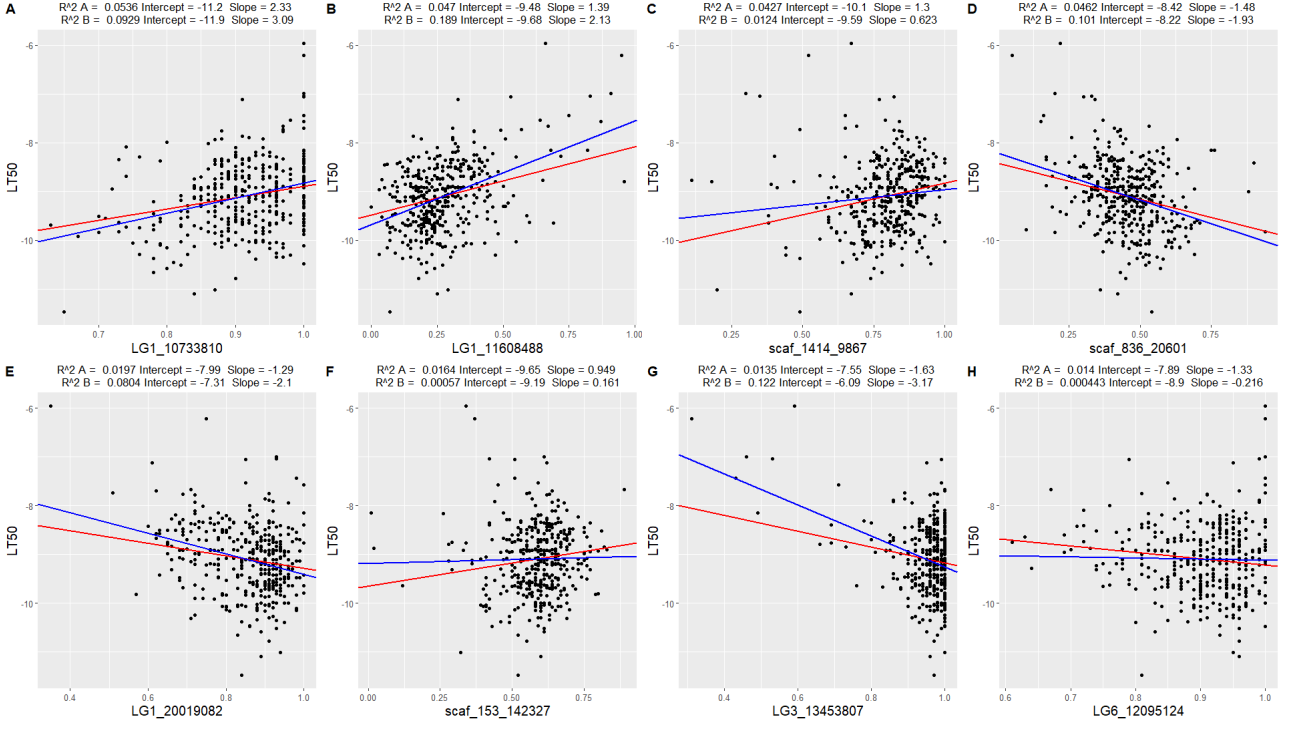

A
LG1_20019082
B
LG1_11608488
C
LG1_10733810
D
LG3_13453807
E
scaf_1414_9867
F
scaf_836_20601
G
scaf_153_142327
H
LG6_12095124
origin
Americas
Belgium
Czech Republic
Denmark
Finland
France
Germany
Great Britain
Italy
Norway
Oceania/Asia
Other Central/Eastern Europe
Serbia
Southern Europe
Sweden
Sweden_BR
Switzerland



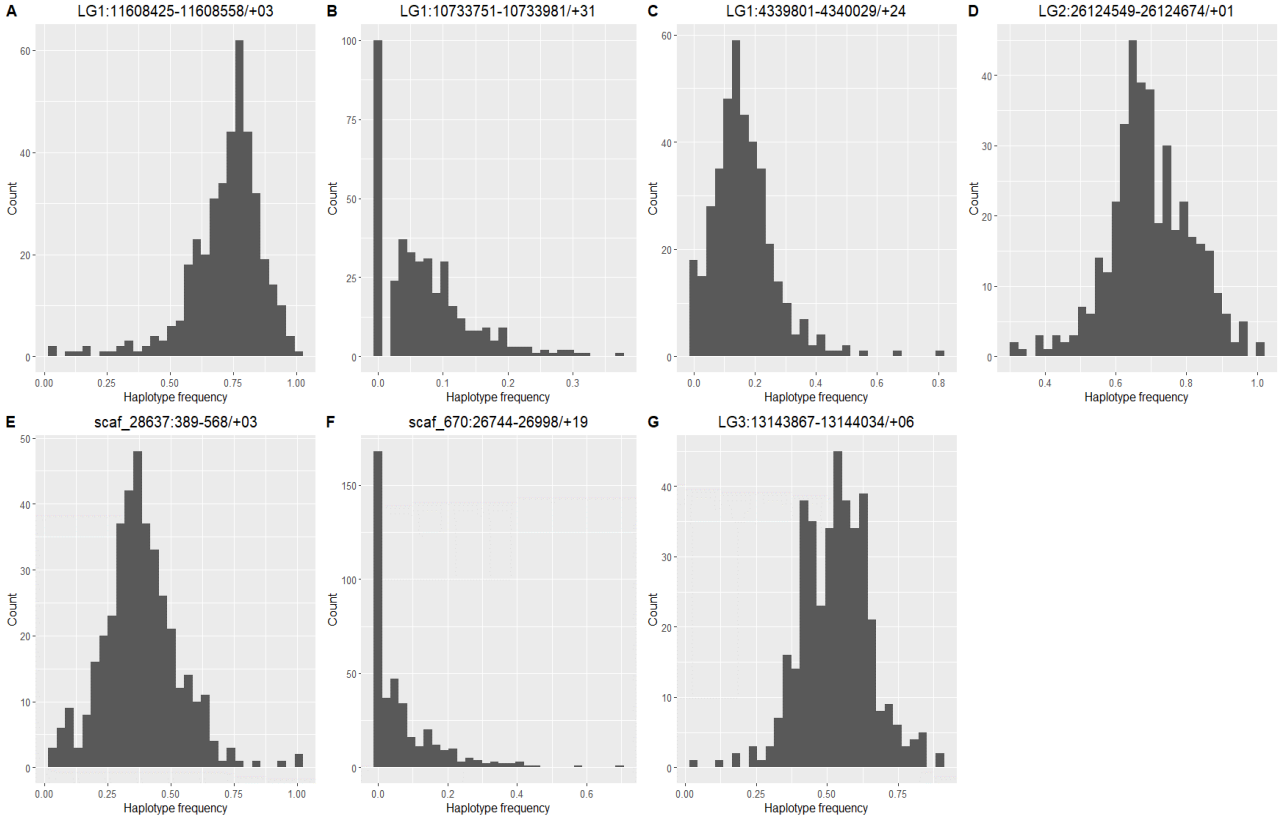
Supplementary Figure 8. Boxplots showing the variation/distribution of allele frequencies for the SNPs (**A**-**H**) and HTPs (**I**-**O**) significantly associated with FT for red clover accessions grouped by country/area of origin.

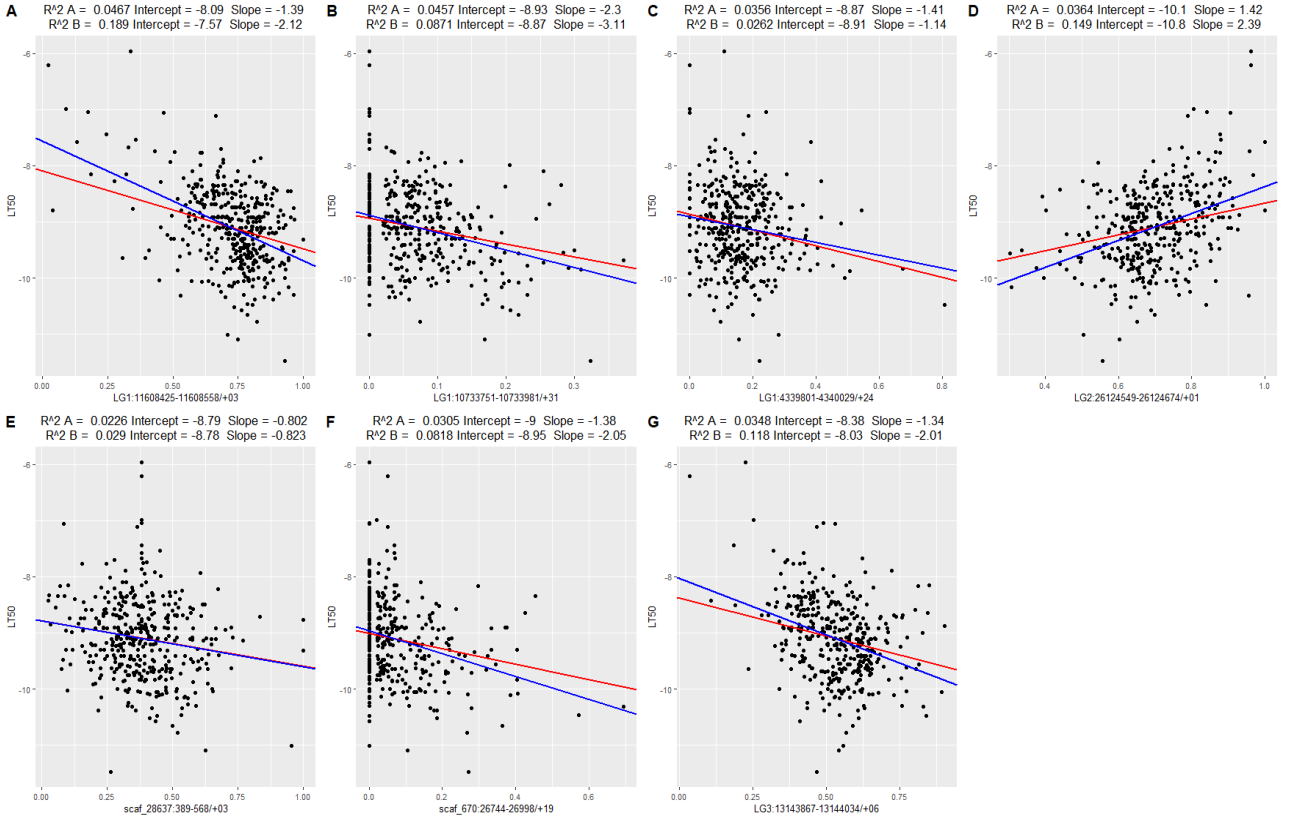
X


A
LG1_10733810
B
LG1_11608488
C
scaf_1414_9867
D
scaf_836_20601
E
LG1_20019082
F
scaf_153_142327
G
LG3_13453807
H
LG6_12095124
Count
Allele frequency


A
R^2 A = 0.0536 Intercept = -11.2 Slope = 2.33
R^2 B = 0.0929 Intercept = -11.9 Slope = 3.09
LT50
LG1_10733810
B
R^2 A = 0.047 Intercept = -9.48 Slope = 1.39
R^2 B = 0.189 Intercept = -9.68 Slope = 2.13
LT50
LG1_11608488
C
R^2 A = 0.0427 Intercept = -10.1 Slope = 1.3
R^2 B = 0.0124 Intercept = -9.59 Slope = 0.623
LT50
scaf_1414_9867
D
R^2 A = 0.0462 Intercept = -8.42 Slope = -1.48
R^2 B = 0.101 Intercept = -8.22 Slope = -1.93
LT50
scaf_836_20601
E
R^2 A = 0.0197 Intercept = -7.99 Slope = -1.29
R^2 B = 0.0804 Intercept = -7.31 Slope = -2.1
LT50
LG1_20019082
F
R^2 A = 0.0164 Intercept = -9.65 Slope = 0.949
R^2 B = 0.00057 Intercept = -9.19 Slope = 0.161
LT50
scaf_153_142327
G
R^2 A = 0.0135 Intercept = -7.55 Slope = -1.63
R^2 B = 0.122 Intercept = -6.09 Slope = -3.17
LT50
LG3_13453807
H
R^2 A = 0.014 Intercept = -7.89 Slope = -1.33
R^2 B = 0.000443 Intercept = -8.9 Slope = -0.216
LT50
LG6_12095124

Y


A
LG1:11608425-11608558/+03
B
LG1:10733751-10733981/+31
C
LG1:4339801-4340029/+24
D
LG2:26124549-26124674/+01
E
scaf_28637:389-568/+03
F
scaf_670:26744-26998/+19
G
LG3:13143867-13144034/+06
Count
Haplotype frequency



Supplementary figure 9: Distribution across 392 red clover accessions and linear regression with LT50 of (**X**) reference allele frequency of individual SNPs; and (**Y**) haplotype allele frequencies; detected as significantly associated with LT50 (in °C, y-axis values) by the MLMM association analyses. Regression lines are plotted both for models accounting for kinship (red lines) and not accounting for kinship (blue lines). $R^2$and slope values are also given in each plot for both models with and without kinship ("R^2 A" and "R^2B" at the top of each regression plots respectively).

Errata

Errata list

Page	Line	Original text	Corrected text
ii	7	*Prof. Em. Beat Boller*	Dr. Beat Boller
9	6	2n=2n=14	2n=2x=14
10	8	«most of the year»	«the period of the year»
10	9	Inclusions	Inclusion
13	14	forms	characteristics
13	21	Some details	details
13	27-31	"This method does not allow for".. progeny..	"Because of difficulties related to conduct".. progeny
16	13	has	have
17	23	because	because of
18	28	is not	it is not
21	26	levels.,	levels,
22	9-8	differ in these responses to this treatment	differ in response of these traits to CA
22	24	Plant materials	Plant material
23	30	the three experiments	the repeated experiments
24	29	Waters ACQUITY	(Waters ACQUITY)
24	32-33	The LC system was a Dionex 3000 LC equipped..	The LC system (Dionex 3000 LC) was equipped..
24	34	high resolution Thermo Fisher Q-Exactive	high resolution Q-Exactive (Thermo Fisher)
25	18	data set	data sets
26	15-17	FT is indeed not always correlated with winter survival, because other stresses than freezing temperatures may be predominant, and because resistance mechanisms towards the various stress factors are multiple and likely to be, to some extent, genetically independent	At these locations other stresses than freezing temperatures may be predominant, and resistance mechanisms towards the various stress factors are likely multiple and, to some extent, genetically independent,
27	9	"It may be possible that we here.."	"We may here.."
27	10	'stress tolerators	'stress tolerators'
27	20	"but we cannot exclude that more freezing tolerant accessions are able "	"but more freezing tolerant accessions are better able"

31	8	confirms	confirm
31	11	the two analyses found	the two analyses (SNP and short read haplotype-based GWAS) found
31	23	in term	In terms
31	26	exclusions	exclusion
31	30	relative	relatively
32	16	(INT1) protein. A protein that facilitates	(INT1) protein which facilitates
33	10-11	by (Blanco-Pastor et al. (2020)	by Blanco-Pastor et al. (2020)

ISBN: 978-82-575-1909-4
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no