Improving the seed yield potential of tetraploid red clover (*Trifolium pratense* L.)

Forbedring av frøsettingspotensialet i tetraploid rødkløver (*Trifolium pratense* L.)

Philosophiae Doctor (PhD) Thesis

Helga Amdahl

Department of Plant Sciences
Faculty of Veterinary Medicine and Biosciences
Norwegian University of Life Sciences

Ås (2016)

The Research Council of Norway

With funding from
The Research Council of Norway

Thesis number 2016:79
ISSN 1894-6402
ISBN 978-82-575-1396-2

Supervisors:

Professor Odd Arne Rognli
Department of Plant Sciences
Norwegian University of Life Sciences
P.O Box 5003, 1432 Ås, Norway
odd-arne.rognli@nmbu.no

Dr. Petter Marum
Graminor AS
NO-2322 Ridabu, Norway
petter.marum@graminor.no

Dr. Muath Alsheikh
Graminor AS/ Department of Plant Sciences,
Norwegian University of Life Sciences
NO-2322 Ridabu, Norway
muath.alsheikh@graminor.no

Dr. Trygve S. Aamlid
Department of Urban Greening and Environmental Engineering, Landvik Research Centre
Norwegian Institute of Bioeconomy Research (NIBIO)
Reddalsveien 215, NO-4886 Grimstad, Norway
trygve.aamlid@nibio.no

Evaluation committee:

Professor Isabel Roldán-Ruiz
Division of Plant Growth and Development
Institute for Agricultural and Fisheries Research, ILVO
B-9090 Melle, Belgium
isabel.roldan-ruiz@ilvo.vlaanderen.be

Dr. Birte Boelt
Department of Agroecology-Crop Health
Aarhus University
Forsøgsvej 1, 4200 Denmark
birte.boelt@agrsci.dk

Dr. Morten Lillemo
Department of Plant Sciences
Norwegian University of Life Sciences
P.O. Box 5003, 1432 Ås, Norway
morten.lillemo@nmbu.no

Table of contents

Acknowledgements

The present work was carried out between 2011 and 2016. It was funded by the plant breeding company Graminor AS and the Industrial Ph.D. Scheme (Nærings-ph.d.) program in the Research Council of Norway (project no. 209702).

The past few years was a great journey. Thank you all at Graminor for seeing potential in me and for giving me this opportunity. I appreciate the encouragement, advices and support that you have given me. I have learned a lot from all of you and I am looking forward to work with you in the future.

I have been so lucky to have an amazing team of supervisors. Thank you professor Odd Arne Rognli for accepting me as a PhD student and guiding me through this journey. Petter Marum; you have been one of a kind supervisor. Your knowledge is admirable, your enthusiasm about plant breeding contagious, discussions with you were valuable. Muath Alsheikh; thank you for your never-ending enthusiasm, positivism, patience and guidance throughout the study. I am deeply grateful to Trygve S. Aamlid for his readiness to help and answering my questions whenever I needed. Thank you for your hospitality, supportive words and encouragement. Thank you for all your support, all discussions and all the laughs, it has been inspiring to work with you.

I would also like to thank co-authors Åshild Ergon and Mallikarjuna Rao Kovi for all discussions and help with writing the manuscripts; Linda Öhlund, the forage breeder at Lantmännen, for coordinating the field trials in Sweden and for answering millions of questions. My special thanks to Mirjana, Kristin, Rune, Jørgen and Elena, the best technicians ever, for all the help with the field trials at Bjørke. Thank you Susanne, my officemate, for all the conversations, discussions and support during the last years. Challenges are so much easier to tackle when you can share them with people who are in a similar situation as yourself.

Family and friends deserve special thanks for being there for me whenever I needed. Thank you Therese and Mathea for demanding a lot of my time in the afternoons and evenings. These breaks were full of precious moments and they gave me energy to continue my work, day after day. You have been amazing Frode. Thank you for believing in me, for all your support (and help with the trials), love and for your patience. I am looking forward to the new chapter in our life.

List of papers

I. Amdahl, H., T. S. Aamlid, Å. Ergon, R. K. Mallikarjuna, P. Marum, M. Alsheikh and O. A. Rognli. 2016. **Seed yield of Norwegian and Swedish tetraploid red clover (*Trifolium pratense* L.) populations.** Crop Science 56:603-612.

II. Amdahl, H, T. S. Aamlid, P. Marum, Å. Ergon, M. Alsheikh and O. A. Rognli. 2016. **Seed yield components in single plants of diverse Scandinavian tetraploid populations (*Trifolium pratense* L.).** Crop Science. Accepted, in press (First Look: doi: 10.2135/cropsci2016.05.0321; Date posted: August 19, 2016)

III. Kovi, M. R., H. Amdahl, M. Alsheikh, and O. A. Rognli. 2016. ***De novo* and reference transcriptome assembly of transcripts expressed during flowering provide insight into the seed setting in tetraploid red clover (*Trifolium pratense* L.)**. (Manuscript).

Summary

Tetraploid red clover has similar forage properties as diploid red clover. However, tetraploid red clover plants are taller, have thicker stems, larger leaves and flower heads, and bigger seeds resulting in higher forage yield than diploid red clover plants. However, the seed yield of tetraploids is significantly lower than of diploids, which is challenging for seed companies. For farmers to be able to benefit from higher forage yield of tetraploids, the seed yield of tetraploid red clover has to be improved. In this project, we studied different aspects of seed yield in tetraploid red clover, focusing on seed yield components. In two consecutive years, twelve Norwegian and Swedish cultivars/breeding lines were studied as spaced plants and in dense canopy trials. Trials with spaced plants were established at one locality while the trials with dense canopies were established at four locations (two in Norway and two in Sweden). Seed yield per flower head was identified as the seed yield component most strongly correlated with the seed yield per plant and with the seed yield per area. Additionally, we found that the seed yield was significantly higher in cultivars developed by crossing of existing tetraploids than in neopolyploids. Our study also aimed to identify putative genes that control seed yield in tetraploid red clover. Transcriptomic analysis was performed on flower buds obtained from two relatively high and two low seed yielding plants, with the aim of identifying transcripts that potentially are involved in determination of seed yield. Genes related to flower development, pollen pistil interactions, photosynthesis and embryo development were differentially expressed in the two genotypes contrasting in seed yield. A significant number of genes related to pollination was overrepresented in the high seed yielding genotypes, which might be a reason for their good seed setting ability. The candidate genes detected in this study might be used to develop molecular tools for breeding tetraploid red clover varieties with improved seed yield potentials.

Sammendrag

Tetraploid rødkløver har samme egenskaper som diploid rødkløver, men den er høyere, har tykkere stengler, større blader, større blomsterhoder og større frø. Som resultat av dette er fôravlingene større sammenlignet med diploid rød kløver. Dessverre gir tetraploid rødkløver betydelig lavere frøavling, noe som byr på utfordringer når frøfirmaene skal produsere frø til sine engfrøblandinger. For at bonden skal fortsette å ha nytte av de gode egenskapene som tetraploid rødkløver har, må frøavlingen forbedres. Vi har studert ulike frøavlingskomponenter som påvirker frøavlspotensialet i tetraploid rødkløver. Målet var å identifisere komponenter som kan benyttes til å øke frøavlingen ved foredling. Tolv norske og svenske sorter/foredlingslinjer ble studert i enkelplanteforsøk og i tett bestand i to påfølgende år. Enkelplanteforsøk ble anlagt på ett sted mens forsøkene i tett bestand ble anlagt på fire steder (to i Norge og to i Sverige). Frøavling per blomsterhode ble identifisert som den frøavlingskomponenten som hadde størst betydning både i enkelplanteforsøk og i tett bestand. I tillegg ble det funnet at tetraploide sorter som var utviklet ved å krysse eksisterende tetraploide planter ga høyere frøavling enn sorter utviklet ved å kromosomfordoble diploider. Et av målene var også å identifisere gener som potensielt påvirker frøavling i tetraploid rødkløver. RNA sekvensering av blomsterknopper fra to planter med relativ høy frøavling og to med relativt lav frøavling hadde som mål å identifisere overuttrykte og underuttrykte transkripter som kunne forklare forskjellene i frøavling mellom disse plantene. Gener relatert til blomster-utvikling, pollen-griffel samspill, fotosyntese og embryo utvikling var forskjellig utrykt i planter med lav og høy frøavling. Et betydelig antall gener relatert til pollinering var overuttrykt i planter med høy frøavling som kan være grunn til dens høy frøavlingsevne. Kandidat gener identifisert i denne studien kan muligens brukes til å utvikle molekylære verktøy for foredling av tetraploid rødkløver med større frøavling.

1. Introduction

1.1 Origin and history of red clover

Red clover (*Trifolium pratense,* L.) belongs to the *Fabaceae* family, genus *Trifolium,* and it is considered the oldest cultivated species in this genus. Red clover originates from southeast Eurasia and is indigenous to Europe, the Near East, North Africa and central Asia (Zohary and Heller, 1984; Taylor and Quesenberry, 1996). The first known document mentioning red clover in Europe and probably worldwide was the book "De vegetabilibus" by Albertus Magnus (1193 – 1280) (Taylor and Quesenberry, 1996; Boller et al., 2010). However, the cultivation of red clover as a forage crop was already known in Europe since the 4th century A.D. From the 13th until the 16th century, clover spread from Spain to France, Italy, Belgium, Netherlands and Denmark. By 1663, it was known in the US and by 1776 in Russia (Fig. 1). While farmers in certain areas were reluctant to accept red clover as a new forage crop, Kjærgaard (2003) refers to replacement of fallow with red clover in European crop rotation as the '17th century's green revolution' (Merkenschlager, 1934; Taylor and Quesenberry, 1996; Boller et al., 2010).

The origin of Russian red clover has been debated for a long time. One of the possible sources is thought to be Western Europe where the early type of red clover was grown. Another suggested source is that the local wild red clover, which was of late type and adapted to Russian climate, was domesticated and used in breeding (Merkenschlager, 1934; Wexelsen, 1937). Recent results by Semerikov et al. (2002) are in favour of the Western European origin. In the beginning, red clover was cultivated in the central and western regions of Russia and by the end of the 19th century was also in the Ural region (Semerikov et al., 2002).

Red clover was introduced to the United States in the 17th century. It is believed that the first seeds came from Holland with a ship that was carrying cattle and different seeds to the colonists. Due to the good establishment of this crop and its property as a good fodder crop, it spread to the west and further all over the US. In the beginning of the 20th century red clover production for hay reached its maximum by being grown on approximately 30 mill acres, after which it gradually declined to approximately 15 mill acres in 1957 (Fergus and Hollowell, 1960). Diminishing importance of red clover occurred also in Europe coinciding with the invention (1909) and spread of the Haber-Bosch method for industrial manufacturing of ammonium from atmospheric nitrogen. After the World War II, the access to commercial fertilizer dramatically reduced the importance of clover in agriculture (Kjærgaard, 2003). The industrial production of nitrogen (N) fertilizer expanded and the farmland in Europe and US

changed to monoculture (Kjærgaard, 2003; Taylor, 2008). Decreasing importance of red clover in the US was also due to the shift to the alternative legume species alfalfa (*Medicago sativa* L.) and to the partiall decoupling of livestock production from grazing (Naylor et al., 2005). Over the past couple of decades, concerns have been raised regarding the high-energy cost and harmful environmental effects of industrial N production. Increased focus on sustainable agriculture, organic production and integrated crop-livestock systems have again stimulated the interest in forage legumes (Kjærgaard, 2003; Taylor, 2008).

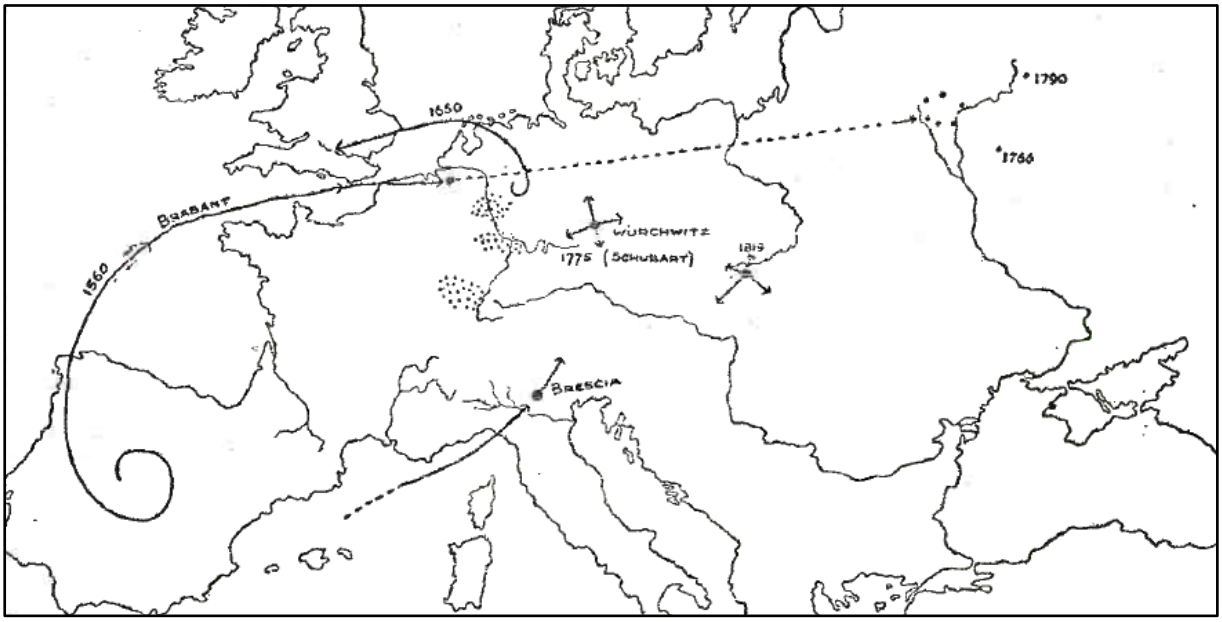


Figure 1. The historical bases of red clover cultivation. Adapted from Merkenschlager (1934).
Large dots represents localities in which temporary trials were made during the crop's early history.
Large dots with arrows attached represents localities from which red clover cultivation spread.
Small dots represent areas in which red clover cultivation ceased because of the Thirty Year's War (1618-1648). Small circles represents Kashira (1766) and Moscow (1790).

1.2 Agronomic importance

Red clover is an important livestock feed grown in mixtures with grasses on approximately four million hectares worldwide (Taylor, 2008; Isobe et al., 2014). Timothy (*Phleum pratense* L.), meadow fescue (*Festuca pratensis* Huds.), tall fescue (*Festuca arundinacea* Schreb) and perennial ryegrass (*Lolium perenne* L.) are the most common companions to red clover. Red clover has a high protein content. This results in a higher content of crude protein and polyunsaturated fatty acids in the final products (milk and meat) when red clover is included in the ruminant's diet (Taylor and Quesenberry, 1996; Lee et al., 2009). In addition, red clover increases the palatability and digestibility of the mixture resulting in higher weight gains and reproduction rates of cattle (Huss-Danell et al., 2007; Taylor, 2008). The importance of red

clover in grass mixtures is also due to its nitrogen (N) fixing ability through symbiosis with the bacteria *Rhizobium leguminosarum* biovar *trifolii*. The amount of N that red clover can fix during one growing season ranges from 40 to 400 kg ha^{-1}, and this reduces the need for N input through fertilization (Taylor and Quesenberry, 1996; Boller et al., 2010; Huss-Danell et al., 2007). Additionally, the high content of polyphenol oxidase (PPO) in red clover slows down the degradation of proteins and thus reduces N losses when making silage (Lee et al., 2004; Sullivan and Hatfield, 2006).

In the 1930s, Wexelsen (1937) estimated the annual value of red clover feed to approximately 40 mill Norwegian kroner (NOK). The current anticipated value of feed in Norwegian milk production is 6.2 billon NOK (H. Volden, the head of TINE Rådgivning, pers. comm., 2016). If the proportion of red clover in the feed is 10 %, the value of red clover feed in Norway is 620 million NOK, or approximately 68 million euros.

1.3 Adaptation and breeding of red clover in Norway

After the introduction of red clover to Norway from the Netherlands and Great Britain in the 18th and 19th century, the cultivation comprised of a number of different types. This resulted in the formation of many locally adapted landraces in different regions. In the beginning, the early type of red clover was grown. However, by the end of 18th century, people were talking about late red clover, which was spreading to countries with short summers and cold winters. It is unclear where this late red clover came from and how it evolved. One hypothesis is that the farmers selected late and winter hardy types from the existing genotypes. However, it is unlikely that sufficiently late and winter hardy types occurred in the early material to such an extent that it could be the basis for the late Norwegian types. Another possible hypothesis is that the early red clover was intercrossed with later wild types when it reached Russia and other Nordic countries; and hence gradually became more winter hardy (Wexelsen, 1937).

During the 19th century, Norway imported red clover seed from several European countries and from the United States. Soon after the first import, the farmers in Norway began to set aside some of their pastures for seed harvest. Seed harvested from these fields proved to be better adapted to the local conditions compared to the imported seed. Around 1850, there were several winter hardy landraces in Norway of which the most known were ‘Molstad’ and ‘Totenkløver’. The first forage yield trials, comparing Norwegian landraces with foreign strains of red clover in mixtures with timothy (*Phleum pratense* L.), were established in 1889. Over the next years, new and improved strains were imported and were gradually included in the trials. However,

the best landraces always proved superior compared to imported material (Wexelsen, 1937). A survey conducted around 1950 showed that about 60 % of Norwegian farmers had been growing the same local cultivar for more than 20 years and some even up to 80 – 90 years (Vestad, 1990). Although 'Molstad', the first Norwegian red clover cultivar, was not officially listed and approved for certified seed production until 1953, the annual Norwegian production of red clover seed was estimated to 400-500 tons already in the 1930s (Vestad, 1990). Since then, most of the red clover seed production has been going on in the eastern and central part of the country.

Dr. W. Christie initiated the Norwegian red clover breeding program in 1924 while the work on developing tetraploid red clover started at the Agricultural University after the Second World War by Prof. Håkon Wexelsen (Wexelsen, 1937; Vestad, 1990). After the retirement of Ass. Prof. Reidar Vestad in 1990s, the red clover germplasm was transferred first to Planteforsk and in 2003 to Graminor. Around 1950, breeding for *Sclerotinia* resistance in red clover was initiated. The first Norwegian tetraploid cultivar 'Tripo' came onto the market in 1964. After the collection of numerous local strains from different latitudes and altitudes in Norway, a new diploid red clover cultivar 'Pradi' was released in the 1981s. This cultivar proved to be better than 'Molstad'. In the 1990's, the most important goals in the red clover breeding program were: disease resistance, winter hardiness, persistence, forage yield, forage quality and seed yield (Vestad, 1990). At present, the focus of the red clover breeding program at Graminor AS is mainly on forage yield, winter hardiness, persistence and seed yield. As of 30 March 2016, there are six diploid and eight tetraploid cultivars on 'The Norwegian official list of varieties'. Among these, three diploids and seven tetraploids are Norwegian cultivars (http://www.plantesortsnemnda.no).

1.4 Morphology of red clover

Red clover is a short-lived perennial species in which individual plants usually survive between two and four years (Fig. 2). It has a taproot, which develops from the primary root and under good growing conditions stretches down to more than 1.5 m depth (Fergus and Hollowell, 1960). The primary shoot of a seedling is vegetative. Stems elongate from the crown, which is more or less at the soil level, and terminate with a flower head. Stems are with and without non-glandular trichomes (hairs). The number of stems and nodes per plant is highly variable. While there are reports on spaced plants having between ten and sixty stems, plants seeded in a dense canopy have between four and nine (Fergus and Hollowell, 1960; Van Minnebruggen

Figure 2. Morphology of a red clover plant (*Trifolium pratense* L.). Source: https://no.pinterest.com/kathyjosand/botanicals

et al., 2013). Stems have a limited number of internodes, 4–6, (Taylor and Quesenberry, 1996; Boller et al., 2010). All leaves are alternate and trifoliar with the exception to the first true leaf, which is unifoliate. The colour of the leaves is green, often with a typical light or pale green spot (chevron) in each leaflet. Axillary buds, one at each node, develop into branches with or without flower heads. Each flower head consists of up to 300 florets (Fergus and Hollowell, 1960). A floret is composed of a calyx with five lobes, a corolla with five petals, and ten stamens, which encircle the pistil. In general, there is two ovules per floret/ovary, however, in most cases only one fertilized ovule per floret develops into seed (Fergus and Hollowell, 1960). Seeds are kidney shaped, in diploids about 1.5 to 2.2 × 1 mm in size (Taylor and Quesenberry, 1996).

1.5 Phenology and flower induction in red clover

The developmental stage of forage grasses and legumes is closely related to forage quality, and therefore used for estimation of the optimal harvest time (Sanderson and Wedin, 1989; Bakken et al., 2005). Bakken et al. (2005) divided the phenological development of red clover into 4 main stages: vegetative, stem elongation, reproductive and seed development, each with several sub-stages (Table 1). Red clover does not have a vernalization requirements. It is a long day plant, and northern ecotypes usually require longer photoperiods than southern ones in order to flower (Lunnan, 1989).

Leaves and stems have the highest digestibility in the vegetative and stem elongation stages. In the reproductive stage, the concentration of the neutral detergent fiber (NDF) increases (Åman and Nordkvist, 1983; Sanderson and Wedin, 1989). During senescence, the content of crude protein also decreases. According to Åman and Nordkvist (1983), digestibility remained constant during the first four harvests after which it decreased. Higher temperature stimulates faster phenological development of red clover. Sanderson and Wedin (1989) reported 94 % of the variation in herbage NDF concentration to be explained by the phenological stage. The

concentration of the micronutrients (Co, Fe, Mn and Ni) was quite stable during different phenological stages while the concentration of the Cu, Mo and Zn decreased with plant age (Lindström et al., 2013).

Table 1. Phenological stages and sub-stages during development of red clover (Bakken et al., 2005).

Developmental stage	Sub-stage	Description
Vegetative phase	V0	First leaf visible
	V1	First leaf fully developed
	V2	Second leaf fully developed
	V3	Third leaf fully developed
Stem elongation	E0	Beginning of elongation
	E1	First internode visible
	E2	Second internode visible
Reproductive phase	R0	The bud can be noticed
	R1	The first bud is visible
	R2	The first flower stem is fully developed
	R3	Visible anthers on the first flower
	R4	First flower begins to wilt
	R5	Other flowers begin to wilt
Seed development	S0	

1.6 Genetics of red clover

Red clover is a diploid species (2n=2x=14) with a basic chromosome number n=7 (Zohary and Heller, 1984). The chromosomes are small, between 1.5 and 3.0 µ (Ellerström and Sjödin, 1966) and the estimated genome size is ~440 Mb, which is similar to the genome size of *Lotus japonicus* L. and *Oryza sativa* L. (Sato et al., 2005; Kataoka et al., 2012). There are three red clover genetic linkage maps available today. Isobe et al. (2003) developed the first map, based on Restriction Fragment Length Polymorphism (RFLP) markers. Later on, Sato et al. (2005) constructed a map consisting of microsatellite (SSR) markers while Herrmann et al. (2006) developed a linkage map based on Amplified Fragment Length Polymorphism (AFLP) markers. These linkage maps were constructed based on mapping families generated by crossing heterozygous parents (two-way pseudo F2 testcross). Results from further analysis of RFLP and microsatellite markers by Kölliker et al. (2006) and by Sato et al. (2005) indicated transferability of these markers among red clover germplasm (Isobe et al., 2014). Isobe et al.

(2009) developed a consensus genetic linkage map using six mapping populations and found that the locus order from previous linkage maps was highly conserved on their consensus map. This enabled the use of the consensus linkage map as a reference for further genetic analysis of red clover genotypes. Kataoka et al. (2012) tried to integrate linkage and chromosome maps of red clover and found chromosomal collinearity among red clover varieties. By using next generation sequencing (NGS) technology, Yates et al. (2014) provided the first *de novo* assembly of the red clover transcriptome based on RNA-Sequencing data. They tried to identify genes and markers related to drought tolerance.

Alfalfa (*Medicago sativa* L.) is a model species for forage legumes, and its transcriptome assembly provides a good source for studying other forage legumes. Genome sequences of other model legumes such as *Medicago truncatula* (Gaertn.), with its draft genome published in 2011, and *Lotus japonicus* L. are also widely used in comparative legume genome studies (Young et al., 2003; Isobe et al., 2012). Ištvánek et al. (2014) published recently a draft assembly of 16 red clover genotypes. De Vega et al. (2015) went further by constructing a chromosome-scale reference draft genome, which is the second genome assembly of a forage legume.

1.7 Tetraploid red clover

1.7.1 Background and history

The first tetraploid red clover plants were produced in 1939 after the development of the colchicine method for producing polyploids (Sjödin and Ellerström, 1986). A common routine was to apply colchicine solution to germinating seeds, to young seedlings or to apical meristems of growing plants (Taylor and Quesenberry 1996; Boller et al., 2010). Another method for producing tetraploid red clover plants is to treat pollinated flower heads with nitrous oxide (N_2O). After germination of seeds from these flower heads, the plants are allowed to flower and their pollen is then analysed for ploidy level. Haploid pollen has tetraeder shape while the diploid pollen is an octaeder (Taylor and Quesenberry 1996; Boller et al., 2010). The third method of producing new tetraploids is by gametic non-reduction where one diploid and one tetraploid plant are intercrossed (Taylor and Quesenberry 1996; Meglic and Smith, 1992). The first tetraploid plants had thicker and longer stems, broader leaves and bigger flower heads compared to their diploid progenitors. These traits resulted in higher forage yield of tetraploids compared to diploids. Further studies showed that tetraploids were more persistent due to higher resistance to *Sclerotinia trifoliorum* (Julén, 1975). Crude and digestible protein and

crude fiber content was the same in diploid and tetraploid red clover while the percent of dry matter was lower in the tetraploids (Levan, 1945). Tetraploid red clover has never gained a widespread use in US (Taylor and Quesenberry, 1996), whereas in Europe, tetraploids are still of interest in countries such as Sweden, Germany, Switzerland, Belgium, Estonia and Norway.

Plant breeders soon realized that the greatest disadvantage of tetraploid red clover was its low seed yield. It became evident that breeding for improved seed set would be necessary for tetraploid cultivars to make it to the market, and efforts to improve the seed yield therefore started (Julén, 1975). Today, 77 years after the first tetraploids were produced; their lower seed yield compared to diploids is still a relevant issue. Several explanations for the lower seed yield have been suggested over the years but unfortunately, the problem is still not solved.

The results from early work in Sweden indicated that the seed yield improved after several years of seed multiplication (Sjödin and Ellerström, 1986). The explanation for this was that irregularities during meiosis would disappear due to natural selection, thus leaving the selected population with improved fertility and good or acceptable seed yield. It was also suggested that tetraploids would perform better if the diploid starting material had a high seed yield potential. Therefore, it would be favourable to select for higher seed yield in the diploid material before chromosome doubling (Povilaitis and Boyes, 1959; Wit, 1961; Picard and Berthaut, 1966; Julén, 1975). It was also proposed that the use of diploid plants with a wide genetic background would increase the chances for development of improved tetraploids. In Switzerland, colchicine is still used for doubling of chromosomes while NO_2 is used in Sweden (pers. comm. B. Boller, 2013 and L. Öhlund, 2015). In Norway, chromosome doubling stopped in the late 1990's (P. Marum, pers. comm.).

In Norway, colchicine was used for chromosome doubling of red clover for the first time in 1947 at the Agricultural University of Norway at Ås (Vestad, 1990). Numerous local diploid populations were treated with colchicine however, the background of the first tetraploid red clover cultivar 'Tripo', which was released in 1964, is unknown. Norwegian breeders at that time were aware of the challenges regarding the seed yield and an intensive selection for improved seed yield in tetraploids started in 1979. 'Kolpo', the second Norwegian tetraploid cultivar that came on the official variety list in 1989, had been subjected to six to seven cycles of natural selection for improved seed yield before it was released (Vestad, 1990). Selection for higher seed yield has improved the seed yield potential of tetraploids and several tetraploid cultivars have been released in Norway. However, the seed yield is still not in the range of the

seed yields of the diploid cultivars. There are three herbage seed companies in Norway but only one is producing seed of tetraploid red clover cultivars.

1.7.2 Self-incompatibility

In order to create and sustain genetic diversity within a species, plants have evolved various mechanisms that prevent self-fertilization (Takayama and Isogai, 2005; Ridout et al., 2005). One of such mechanisms is gametophytic self-incompatibility (GSI), which is present in red clover (Taylor and Smith, 1979). Self-incompatibility (SI) genes control the recognition between pollen and stigma, thus ensuring that ovules are fertilized with pollen from another plant (Leduc et al.; 1990; Charlesworth, 2002). This mechanism prevents inbreeding, which otherwise will result in a strong reduction in plant vigor (Leffel, 1963).

A single multi-allelic locus, called *S* locus, determines the compatibility relationships in red clover (Fergus and Hollowell, 1960; Lawrence, 1996; Ridout et al., 2005; Riday and Krohn, 2010). A basic characteristic of S alleles in red clover is dominance, meaning that only one allele is active and the other is inert (recessive) (Pandey, 1955). Diploid red clover is highly self-incompatible and rarely produces any seed after self-pollination. In tetraploid red clover, however, some allelic combinations can significantly reduce GSI (Pandey, 1956). While this is not enough to ensure regular fertilization, Pandey (1956) stated that self-compatibility due to ploidy might become a huge handicap as tetraploid populations may degenerate due to inbreeding. For that reason, he suggested that plants selected in the first few generations should be tested for self-compatibility.

The estimated number of S alleles in red clover is 200. This is a higher number than in white clover (*Trifolium repens* L.) and alsike clover (*Trifolium hybridum* L.) implying that the probability of pollen having the same allele as the stigma is very small (in a population of 200 individuals, 1 %) (Lawrence, 1996). This is important for red clover breeding when combining different number of parents to make a synthetic population.

1. 8 Red clover seed production – worldwide and in Scandinavia

Seed crops of red clover in Canada and northern Europe are usually established with a cover crop of spring wheat, spring barley or oats, and the seeding rate of red clover is usually 2-3 kg ha^{-1} (Aamlid, 2015; Boelt et al., 2015). In the US, Central Europe and New Zealand, it is more common to establish red clover in pure stands (without a cover crop), and the seeding rate is usually higher, 9 kg ha^{-1} (Boelt et al., 2015).

In the US and Central Europe, it is common with one cut of red clover for forage in the spring of the seed harvest year. The spring harvest stimulate and synchronize flowering, disturb the life cycle of harmful insects, and provide additional income to the farmer (Boelt et al., 2015). In Norway, it has been typically considered that defoliation in spring will delay flowering and seed maturation too much, given the short growing season (Aamlid et al. 2010). On the contrary, many Swedish seed growers have started to defoliate in spring to control *Tripleurospermum inodorum* and other weeds, especially in organic seed crops. In the US, it is common to harvest red clover seed for two consecutive years (Oliva et al., 1994; Steiner et al., 1997), while it is more common with only one harvest year in Scandinavia.

Water stress during flowering reduces the number of seeds per floret in red clover. Irrigation shortly before or during flowering is therefore advisable in certain areas (Oliva et al., 1994). Lately, an increasing number of red clover seed growers in Oregon have installed irrigation, and this practice has enabled seed yields over 1,000 kg ha^{-1} (N. Anderson, pers. comm., 2016). In Norway, the natural rainfall during the growing season is usually higher, and irrigation is not a common practice.

Various diseases and insect pest can significantly affect the seed yields of red clover. *Hyperia nigrirostis* and *Apion* seed weevils are the most common insect pests in the seed production fields in Norway. The occurrence of these weevils depends on weather conditions, varies from year to year, and even more between seed production districts. In Southern Sweden, seed yields can be significantly reduced due to *Apions* in some years (Lundin et al., 2012). In Norway, however, the occurrence is usually considered too sporadic to recommend prophylactic applications of insecticides to all red clover seed crops (Aamlid et al., 2009). In Canada, the caterpillar *Coleophora daeuratella* causes severe damages especially in the second year of red clover seed fields (C. Yoder, pers. comm., 2016). Among diseases, the most common in Norway is *Sclerotinia trifoliorum*. This fungus attacks the red clovers taproot, thus reducing its persistency (Aamlid, 2015). Steiner et al. (1997) concluded that neither root rot (*Fusarium*

oxysporum Schlect.) nor root borer (*Hylastinus obscurus* Marsham) are responsible for the declining seed yields throughout the growing period as previously suggested by Smith (1994) and Oliva et al. (1994).

While the average seed yield in USA has increased in the recent years (Anderson, pers. comm., 2016), seed yields in Denmark, Sweden and Norway, have decreased. The average seed yield in Norway and Sweden between 2002 and 2012 was 247 and 300 kg ha^{-1} for diploids, and 164 and 225 kg ha^{-1} for tetraploids, respectively (Norwegian seed companies, pers. comm., 2015; I. Andersson, pers. comm., 2015). For comparison, the average seed yield in Oregon, which has 90 % of the US production, is about 800 kg ha^{-1}.

1.9 Seed yield of tetraploid vs. diploid red clover

Forage grasses and legumes are in general low seed yielders (Sleper and Poehlman, 2006). Low seed yield in red clover was documented already in the beginning of the 20th century (Martin, 1913; Wexelsen, 1937). The negative correlation between forage and seed yield in red clover challenges the development of high seed yielding cultivars (Boller et al., 2010; Sleper and Poehlman, 2006). In addition, tetraploid plants of red clover usually produce even less seed than diploid plants. At the very beginning of the work on improving the seed yield, suggested reasons for this phenomenon were: 1) inefficient/inadequate pollination, 2) irregularities in the development of pollen and ovules, and during pollen germination, and 3) post fertilization irregularities (Ellerström and Sjödin, 1966; Bragdø-Aas, 1970; Julén, 1975; Büyükkartal, 2003). However, since the researchers thought that seed yield was genetically controlled they tried to improve the seed yield potential of tetraploids by selecting for the most influential seed yield components. Dennis (1975) performed a thorough study comparing diploid and tetraploid red clover. Despite of considerable effort to increase the seed yield between the 1940's and the 1980's, tetraploid cultivars still produce, on average, 34 and 25 % lower seed yields than diploid cultivars in the Norwegian and Swedish seed production, respectively. On the other hand, at some locations and in some years, tetraploids may produce equal or higher seed yields compared to diploids (B. Andersson, personal communication, 2010).

1.10 Reasons for lower seed yields in tetraploids

1.10.1 Inefficient/inadequate pollination

Pollination is a crucial factor for seed production in red clover. Honeybees (*Apies mellifera* L.) and bumblebees (*Bombus* ssp,*)* are the main pollinators of red clover. There are different estimates on the optimal number of *Bombus* species for maximum seed set, from 2,000 to 4,000 ha^{-1} (Rao and Stephen, 2009). It seems that honeybees are the primary pollinators of clover at southern while bumblebees are more important at northern latitudes (Julén, 1954; Valle, 1961). Some other insects may also visit red clover flowers, but are usually of less importance for pollination (Starling et al., 1950). The pollen carried by the pollinator is transferred to the stigma, which the pollinator usually hits when landing on the flower (Fig. 3A). Each floret can

Figure 3. (A) Bumble bee (*Bombus terrestris* L.) visiting red clover flower. (B) Red clover flower with visible holes in the base of the floret (indicated by arrows), made by bumblebees. Photos by Helga Amdahl.

be pollinated several times as long as the stigma is receptive (Taylor and Smith, 1979). The pollinator sticks its proboscis into the corolla tube and sucks the nectar. Each flower head remains open for 6 to 8 days but each floret must be pollinated within 2 to 4 days after opening (Rao and Stephen 2009). In this way, synchrony in flowering and honey- and bumblebee foraging is crucial for optimal pollination and seed production (Rao and Stephen, 2009). Some pollinators are 'robbing' the nectar by making a hole at the base of the corolla tube (Fig. 3B) (Starling et al., 1950). This phenomenon is thought to reduce the efficiency of pollination. However, the pollinator does not have to penetrate the corolla tube to pollinate. It is enough that it hits the stigma when landing on the flower.

Darwin (1888) stated that bumblebees are ideal pollinators of red clover. However, their number in nature is usually inadequate for optimal pollination of red clover seed crops (Starling

et al., 1950; Dunham, 1939). This was confirmed by the significant seed yield increase following the introduction of bumblebee colonies in New Zealand (Plath, 1925). Later, Wilsie and Gilbert (1940) concluded that honeybees are more important than bumblebees for red clover pollination due to their higher abundance. A number of studies on the importance of pollinators for red clover seed yield at the beginning of the 20th century found seed production of diploid red clover challenging already at that time (Wexelsen, 1937).

Honeybees generally prefer white clover, *Brassicas* and other crops to red clover, and within red clover, they also forage more on diploid than on tetraploid plants (Julén, 1954). Sjödin and Ellerström (1986) measured the time honeybees spent on diploid and tetraploid flowers during pollination and concluded that honeybees needed 40 % longer time to pollinate tetraploid flowers. Bumblebees, on the other hand, visited more flowers per time unit; they worked longer during a day and under more various weather conditions (Rao and Stephen, 2009).

Tetraploid red clover plants have, on average, 1 mm longer corolla tubes compared to diploid clover (Vleugels et al., 2015a; Amdahl et al., 2016). Average corolla tube length in diploid red clover is around 9.5 mm (Starling et al., 1950). It is well known that the length of the corolla tube varies between different cultivars. In addition, measurements on the same strain differ between years, locations and at different time of the year (Starling et al., 1950). Bumblebees, with their on average 11 mm long proboscis, are considered to be more efficient pollinators of tetraploid red clover than honeybees, which have on average 6.5 mm long proboscis (Starling et al., 1950; Julén, 1954; Julén, 1975; Taylor and Smith, 1979; Rao and Stephen, 2009).

There are more than 200 different bumblebee species in the world (Ødegaard et al., 2010). Each region has its own native species, whose importance for red clover seed yield depends on their abundance during the flowering season (Rao and Stephen, 2009). The abundancy varies from year to year but in general, the abundancy of several species have decreased during the last 100 years due to restructuring of agricultural landscapes all over the world (Ødegaard et al., 2010). *Bombus hortorum*, the long tongued bumblebee, was the most abundant in central Finland in the 1960s (Valle, 1961). In the Willamette Valley (Oregon, US), which is the world's largest red clover seed production area, *B. vosnesenskii* is the most important species. In Norway, the most abundant species are *B. lucorum, B. terrestris and B. pratorum* (Ødegaard et al., 2010). Long tongued *B. subterraneus* was considered eradicated from Norway, however a finding of a queen in 2010 and 2011 gives hope that a small population of *B. subterraneus* is again to be found in South-East Norway (Aase et al., 2011; http://forskning.no/insekter-zoologi-okologi/2011/10/slattehumla-surrer).

Julén (1954) found a higher frequency of bumblebees but a lower degree of pollination in tetraploids compared to diploids. A high abundancy of bumblebees was less important for pollination in diploids (Julén, 1954). It was reported that bumblebees preferred diploid flowers with shorter corolla tubes when having a free choice, however, Vleugels et al. (2016) was not able to confirm this preference.

Julén (1954) and Pammel and King (ref. in Rao and Stephen, 2009) suggested breeding for shorter corolla tube in tetraploid clover as a method to increase their attractiveness for pollinators. However, Julén (1971) later found that selection for shorter corolla tube resulted in reduced seed yield. Julén (1954) and Bragdø-Aas (1970) obtained the same low seed yield with hand pollination as with open pollination. Therefore, and based on the recent results published by Vleugels et al. (2015a,b), the length of the corolla tube is most probably not the main reason for lower seed yield in tetraploid compared to diploid red clover.

1.10.2 Irregularities in the development of pollen and ovules and during pollen germination

During the development of male and female gametophyte (pollen grain and ovule) in tetraploids, irregularities in chromosome pairing occur. Ellerström and Sjödin (1966) observed formation of uni-, tri- and quadrivalents during meiosis. Uneven distribution of chromosomes during the meiotic division can inhibit further development of micro- and megaspores (abortion of pollen and ovule) or result in the development of gametes with aneuploid chromosome number. During fertilization, the union of gametes with aneuploid chromosome numbers might result in nonfunctional zygotes, thus reducing seed set (Ellerström and Sjödin, 1966).

Irregularities during meiosis are thought to be more frequent in the "new" tetraploids, i.e. in the first generations after chromosome doubling. It has been proposed that the frequency of irregularities decreases due to natural selection of normal/functional gametes in the following generations (Ramsey and Schemske, 2002). Povilaitis and Boyes (1956) studied meiosis in the C1, C2 and C6 generation of tetraploid red clover. They estimated that one third of the microspores and one third of macrospores had irregularities during chromosome pairing. However, the irregularities in microsporogenesis could not account for the high frequency of pollen abortion. Instead, they proposed that pollen abortion is genetically controlled. In contrast, a high frequency of meiotic abnormalities in the macrospores accounted for significant reduction in seed yield due to a limited number of normal ovules per ovary, i.e. less than two (Povilaitis and Boyes, 1959). Environmental conditions had a minor effect on pollen

abortion of autotetraploid red clover, while it had a significant effect on embryo-sac production in diploid red clover (Povilaitis and Boyes, 1959). The conclusion from this study was that it is the development of ovules that is affected by chromosome doubling and not the development of pollen grains.

1.10.2.1 Pollen germination

The germination of a red clover pollen grain depends on hydration and that no inhibitory substances are produced by the stigma (Martin, 1913). Recent studies showed that pollen grain has to adhere physically to the papillae cells of the stigma in order to germinate. The hydration water usually comes from the stigma (Dresselhaus and Franklin-Tong, 2013). The time required for germination of pollen grains was 8 -10 min *in vitro* and at room temperature (Martin, 1913). Qin et al. (2011) found that germination in vitro was significantly delayed compared to *in vivo*, which indicates that this process is even faster in the field. Lankinen and Öhlund (2013) found that germination of pollen grains was significantly better in diploid compared to tetraploid red clover, which could be due to the differences in pollen quality/fertility between these two ploidy levels. Pollen quality explained 17–32 % of the variation in tetraploid red clover seed set (Bragdø-Aas, 1970).

1.10.2.2 Pollen tube growth

After germination, the growth of the pollen tube (Fig.4) through the style is facilitated by a male-female interaction (Cheung et al., 2010). Julén (1950, 1954) and Pandey (1955) reported haploid pollen tubes to have a higher growth rate compared to diploid pollen tubes, while Lankinen and Öhlund (2013) did not detect any significant differences. Haploid pollen tubes reached the embryo sac after four hours while diploid pollen tubes needed much longer time (Julén, 1950). Compared with haploid pollen tubes, the growth rate of diploid pollen tubes was also more delayed by low temperature suggesting that this could be due to a shorter lifetime of the synergids, which usually guide the pollen tube to the embryo sac and stimulate pollen bursting (Higashiyama et al., 2001). Low temperatures could have shortened their lifetime resulting in insufficient stimulation of pollen bursting, i.e. rupture of the pollen tube and deposition of its cytoplasm and two sperm cells into the female gametophyte (Pandey, 1955).

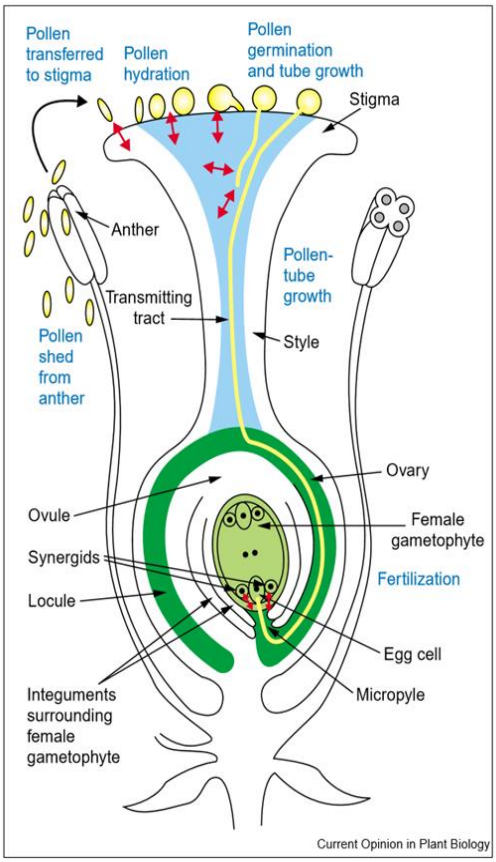


Figure 4. Flower structure and main events involved in pollination. Source: Franklin-Tong, 2002.

1.10.3 Post fertilization irregularities

After fertilization, the zygote is growing by cell division. At the same time, the fertilized polar nuclei develop into the endosperm whose enlargement is also ensured by cell division. The endosperm provides nutrition for the developing embryo. Pandey (1955) found that the development of embryo and endosperm was faster in tetraploids compared to diploids. Additionally, he observed that sometimes the cell division of the endosperm in tetraploid red clover is slower than that of the embryo, which results in seed shriveling and abortion. Buyukkartal (2008) observed that only 5.8 % of ovules in tetraploid clover formed seed. Reasons for this are not known.

If at least one of the two gametes are aneuploid, the zygote will also become aneuploid. A proportion of these zygotes aborts during embryo development, contributing to the reduction

in seed set (Ellerström and Sjödin, 1966). The frequency of aneuploid zygotes in tetraploid red clover was estimated to 16 % by Ellerström and Sjödin (1966) and 8-72 % by Julén (1954).

1.11 Seed yield components

The seed yield of a crop is a product of its individual seed yield components. In red clover, these components are (1) the number of flower heads per unit area (sometimes recorded as number of flower heads per plant or per shoot), (2) the number of florets per flower head, (3) the number of seeds per floret (seed set), and (4) the weight per seed. In addition, some combined seed yield components are also widely used such as the number of seeds per flower head and seed weight per flower head. Seed yield components are commonly used to measure, compare and explain differences in seed yield between diploid and tetraploid red clover and between various populations and cultivars. The realized seed yield is a result of the interaction between physiological processes involved in the formation of different seed yield components, environment conditions, management practices and pests (http://cropandsoil.oregonstate.edu). These dynamic processes influence assimilate partitioning and thus the contribution of each seed yield component to the total seed yield (Fairey and Hampton, 1997). Increasing the total seed yield is most efficiently accomplished by increasing the number of seeds. Changes in one or more components can be compensated for by changes in other components.

In order to identify the seed yield component that most efficiently could be used in the selection for higher seed yield, observations have been done on both single plants and dense canopies. The number of inflorescence per plant was the most important seed yield component in single plants of birdsfoot trefoil (*Lotus corniculatus* L.) (Li and Hill, 1989) and white clover (*Trifolium repens* L.) (Clifford and Baird, 1993). Herrmann et al. (2006) and Malengier and Baert (2007) identified the number of seeds per plant and number of flower heads per plant to be highly correlated with seed yield per plant in trials with single plants of diploid and tetraploid red clover, respectively. There are, however, indications that spaced plants do not always parallel dense canopies agronomically (Hampton and Fairey, 1997), and the seed number per plant is therefore unlikely to be an efficient trait for selection for higher seed yield.

In a dense canopy of diploid red clover, Taylor and Quesenberry (1996) reported number of flower heads per unit area to have the strongest influence on seed yield. This result was also confirmed by path coefficient analysis (Montardo et al., 2003). In contrast, Oliva et al. (1994) found number of seeds per flower head to have the strongest direct effect on seed yield in a study with diploid red clover.

2. The Thesis

2.1 Background and objectives

Red clover (*Trifolium pratense* L.) is one of the species that is bred by the Norwegian plant breeding company Graminor AS (www.graminor.no). However, the Norwegian seed companies are modestly interested in seed production of the tetraploid cultivars developed by Graminor. The reason for this is the low and variable seed yields of tetraploids, which makes it economically risky for farmers to multiply seed of tetraploid cultivars on contract. On this background, Graminor initiated the present PhD project with the aim to study the problems related to the seed yield potential of tetraploid red clover. As described in detail in the introduction, there are several possible explanations for the low and variable seed yields of tetraploid red clover; however, in this project two aspects have been addressed:

1. Identification and description of seed yield components in diverse germplasm of tetraploid red clover. It was expected that the outcomes from these investigations would identify the most efficient seed yield components for selection of higher seed yields, and thus improve seed yield of new tetraploid red clover cultivars.

2. Acquisition of genetic and molecular knowledge of seed yield traits in red clover. A long-term goal is to identify molecular tools for screening of high seed yield tetraploid red clover. In this study, we generated and associated transcriptomic profiles with the seed yield of selected diverse plants and correlated changes in gene expression and networks (accumulation) in response to seed yield in red clover.

This project was funded by the Research Council of Norway and Graminor Breeding AS, and conducted as an industrial PhD project.

2.2 Material and methods

Seven Norwegian and five Swedish cultivars and populations were used to study the seed yield components in single plant and dense canopies in tetraploid red clover. In 2011, dense canopy trials were established at four locations: Bjørke (61.22^{o} N, 20.42^{o} E; 147 m.a.s.l) and Landvik (58.34° N, 08.52° E; 10 m.a.s.l.) in Norway, and Svalöf (55.56° N, 13.06° E; 70 m.a.s.l) and Lännäs (63.09° N, 17.39° E; 26 m.a.s.l) in Sweden. These trials were harvested in 2012. In 2012-2013, dense canopy trials with slightly altered material were repeated at the Norwegian locations: Bjørke and Landvik. A description of the traits that were measured in dense canopy trials is presented in Table 2. Additionally, single plant trials were established in 2011 and 2012 at Bjørke. Traits that were measured on the single plants are described in Table 3. Description and background of each cultivar and its affiliation to the particular trial is given in Table 4. Analysis of the results from the field trials were conducted using Analysis of variance (ANOVA), Pearson correlation analysis, orthogonal contrasts, genotype × environment interaction (G × E) and path coefficient analysis.

Two 'Lasang' plants from the first single plant experiment, with the highest scores for most of the traits, and two 'Tripo' plants with the lowest scores for most of the traits, were chosen for further analysis. Flower buds in early stage of development were sampled from these four plants at three different stages (early, middle and late) during the flowering period. Single-end reads of 50 bp length were generated using the Illumina sequencing platform (HISEQ 2000). The *de novo* assembly was performed in a similar manner as described by Kovi et al. (2016). Further, differentially expressed genes (DEGs) between different flowering time points were identified. The GO classification of DEGs in the two genotypes were generated using the WEGO (Ye et al., 2006) program. Finally, all the clean reads from the two genotypes ('Tripo' and 'Lasang') at the three flowering time point (early, middle and late) were mapped to the red clover reference genome (De Vega et al., 2015) using the ultrafast universal RNA-seq. aligner program STAR (Dobin et al., 2013).

Table 2. Traits observed in the 2012 and 2013 dense canopy experiments.

Type of trait	Trait	Description
Seed yield and seed yield components	Seed yield (SY)	All the plots at one location were harvested at the same time, directly with combiners when the seeds were considered ripe. After harvest the seeds were dehulled and cleaned roughly at each location before being sent to Landvik for final cleaning and laboratory analysis. Seed yields were corrected to 100% purity and 12% seed moisture content.
	Number of flower heads at maximum flowering (NoMax)	In the spring of the harvest year a 0.75 × 0.75 m subplot within each plot was designated. Flower heads were counted inside each subplot and the number was converted to the number of flower heads per m^2.
	Seed yield per flower head (SYFH)	50 randomly chosen ripe flower heads were collected from each plot just before harvest. Seeds were threshed by hand, cleaned, weighed and the seed yield per flower head calculated.
	Thousand seed weight (TSW)	100 seeds were randomly selected from the pure seed yield of each of the four blocks per populations. The seeds were weighed and the thousand seed weight calculated.
Developmental earliness	Date for the first five open flower heads (Date5)	Registered as the number of days after 31 May. The plots were observed every second to third day.
	Maturity (Mat)	Percent of black flower heads with the upper peduncle completely dry prior to harvest. Approximately 100 random flower heads in each plot were observed and classified as black/ripen, green/withered or red/still flowering.
Seed quality	Normal seedlings (NS)	100 randomly chosen seeds from each variety were incubated on a germination table at 20°C. Normal seedlings were counted after 7 and 14 days.
	Hard seeds (HS)	Percent of hard seeds was calculated according to International Rules for Seed Testing (ISTA 1999) after counting the normal seedlings.

Table 3. Description of traits measured and calculated in the two experiments with single plants of red clover

Trait	Description	Experiment
Number of flower heads per 19 July (FH197)	Newly opened flower heads were labeled and counted every third day from the start of flowering until 19 July in order to estimate earliness.	1, 2
Counted number of flower heads per plant (CountFH)	In 2011, the counting and labeling of newly opened flower heads every third day continued until 3 August on nine random plants per population.	1
Calculated number of flower heads (CalFH)	Number of flower heads on the nine chosen plants per population was calculated by dividing the seed yield per plant (SYP) with the average seed yield per flower head (SYFH).	1, 2
Number of florets per flower head (FFH)	Every third day during the flowering period one flower head from each of the nine plants per population were cut and the number of florets counted.	1, 2
Length of the corolla tube (Corolla)	After counting the number of florets per flower head, the corolla tube length of twenty-five randomly chosen florets was measured.	1, 2
Seed number per flower head (SFH)	At maturity in September, one flower head from each labelling date was harvested separately from each of the nine plants per population. The seeds were threshed by hand, cleaned and counted.	
Seed yield per flower head (SYFH)	Weight of the seeds that had been counted to determine SFH.	1, 2
Fertility (FERT)	Maximum fertility is assumed as 2 since the maximum number of ovule per floret is 2. The current fertility was then calculated by dividing the number of seeds per flower head (SFH) with number of florets per flower head (FFH) and expressed in percent.	1, 2
Seed yield per plant (SYP)	Each plant was cut at 1 October, dried, threshed, cleaned and the seed yield determined.	1, 2

Table 4. Description of tetraploid red clover populations used in the seed production experiments.

Country of origin	Population	Approval year	Description of populations	Background of the populations	Dense Canopy†	Single plant‡	Breeding method§
Norway	’Tripo’	1964	First Norwegian 4x variety. No longer in production (Vestad, 1990).	No available information.	1, 2	1,2	D
	‘Reipo’	2002	Low winter hardiness, high dry matter yield, middle late flowering, god tolerance to some diseases (Norwegian variety testing, 2002)	Outcome after one selection cycle (SC) for seed yield and five SC for forage yield on 3600 single tetraploid plants that belonged to nine populations (no available information on which 9 populations) (Vestad, 1990).	1, 2		C
	‘Lavine’	2007	Winter hardy variety that never came to the market because of the low seed yield (Norwegian variety testing, 2007).	Outcome from the doubling of ‘Kongsvoll-2x’ (collection of several diploid local populations from mountain area in Norway).	1, 2		D
	‘L-4374’	Breeding population	Result after the work on improving the seed yield of tetraploids by selection for seed yield per plant.	Outcome after two SC for vegetative characters, four SC for high forage yield, one SC for nematode resistance and two SC for high seed yield on 3600 single tetraploid plants that belonged to nine populations (no available information on which 9 populations) (Vestad, 1990; Aamlid and Marum, unpublished data, 1998).	2	2	C
	‘Lars’	2012	More winter hardy than ‘Lasang’, middle early flowering, high dry matter yield, good tolerance to diseases. This is an upcoming variety on the Norwegian market (Norwegian variety testing, 2012).	A cross between 6 plants of the tetraploid variety ‘Lone’ and 6 plants of the tetraploid variety ‘Betty’. These 12 plants were selected based on their good competitive ability when grown together with timothy.	1		C
	‘Lasang	2013	Winter hardy and persistent variety with medium earliness, high dry matter yield and good tolerance to diseases. Not in a production because the seed companies have chosen to focus on Lars (Norwegian variety testing, 2013).	Outcome from ‘L-4374’ after two cycles of selection for high number of seeds per flower head (Aamlid and Marum, unpublished data, 1998).	1, 2	1,2	C
	‘LøRk0733’	Breeding population	Result from one cycle of natural selection in Lasang.	One cycle of natural selection on ‘Lasang’.	2	2	C

Table 4. continued

Country of origin	Population	Approval year	Description of populations	Background of the populations	Dense Canopy†	Single plant‡	Breeding method§
Sweden	'Betty'	1992	Good winter hardiness and persistence, early flowering, used in northern Sweden, moderate dry matter yield (Norwegian variety testing, 200?).	Outcome of doubling the diploid variety 'Bjursele'. There was no particular selection for higher seed yield in either 'Bjursele' or 'Betty' (L. Öhlund, personal communication, 2015).	1, 2	2	D
	'SW Nancy'	2001	Middle late, used in south Sweden, good persistence (L. Öhlund, personal communication, 2015).	Crossing of 7 single plants of the tetraploid variety 'Fanny'. These plants were selected for high seed yield combined with good vegetative characters. The variety 'Fanny' was made by crossing a certain number of plants that were selected for high nematode resistance and high seed yield (L. Öhlund, personal communication, 2015).	1, 2	2	C
	'SW Torun'	2002	Used in north Sweden, considerably later flowering than 'Betty', higher dry matter yield than 'Betty' (L. Öhlund, personal communication, 2015).	Outcome of doubling the diploid variety 'Jesper'. There was no particular selection for higher seed yield in either 'Jesper' or 'SW Torun' (L. Öhlund, personal communication, 2015).	1, 2	2	D
	'Vicky'	2009	Better persistence than other varieties on the official variety list in Sweden (L. Öhlund, personal communication, 2015).	Outcome of doubling the diploid variety 'Eva'. 'Eva' was selected for improved persistency (L. Öhlund, personal communication, 2015).	1		C
	'SW RK1111'	Breeding population	No available information.	Plants from the breeding population 'SW RK1051' were established as a polycross. Seeds harvested from surviving plants after the second harvest year were put together to form 'SW RK1111' (L. Öhlund, personal communication, 2015).	1		C

† Population included in experiment 1 and/or 2 in the dense canopy trials.
‡ Population included in experiment 1 and/or 2 in the single plant trials.
§ Tetraploid populations developed by chromosome doubling of diploid plants

2.3 Main results and discussion

2.3.1 Seed yield components for selection of higher seed yield in tetraploid red clover

Seed yield per flower head was the seed yield component most strongly correlated with both seed yield per area (Paper 1) and seed yield per plant (Paper 2). Based on experience and the results from our trials, we suggest seed weight per flower head, among all seed yield components, to be used in selection for higher seed yields in tetraploid red clover. By selecting for higher seed yield per flower head, selection includes more than one seed yield component since the seed yield per flower head is comprised of number of florets per flower head, fertility (number of seeds per floret) and seed weight. This is in accordance with the results of Oliva et al. (1994) who reported that fertility was the seed yield component with the highest positive correlation with seed yield per plant. However, we consider seed yield per flower head to be more efficient and less time consuming to screen than fertility.

Several studies previously concluded that the number of flower heads per plant was the component with the highest positive correlation with seed yield per area (Taylor and Quesenberry, 1996) and with seed yield per plant (Montardo et al., 2003; Malengier and Baert, 2007; Vleugels et al., 2016). Even though the number of flower heads was the second most strongly correlated component with seed yield per area (Paper 1) and seed yield per plant (Paper 2) in our trials, these components are not optimal for use in selection for higher seed yields in tetraploids. One reason is that counting of flower heads per plant was very time consuming during these investigations. Some of the single plants had ~1,200 flower heads prior to harvest, which is such a high number that is unaffordable to screen efficiently on a large number of plants. Another reason is that spaced plants develop into larger plants with considerably more flower heads compared to plants growing in a dense canopy. As mentioned above, the number of flower heads was up to 1,200 in some single plants, while the number of flower heads in the dense canopies was not more than 850 m^{-2}. Thus, single plants standing alone do not resemble single plants in a dense canopy when it comes to number of flower heads per plant. Seed yield per flower head was however, more similar between single plants (Paper 2) and dense canopies (Paper 1). Because of this, we believe that selection for higher seed yield per flower head on spaced plants, rather than number of flower heads per plant, could be used more efficiently to increase the seed yield in a dense canopy/population. To confirm this, further studies with direct comparisons of seed yield components of the same genotypes growing as spaced plants and in dense canopies will be needed.

2.3.2 Making neopolyploids vs. crossing of existing tetraploid plants

Doubling of the chromosome number of superior diploid red clover plants to develop new autotetraploid plants (neopolyploids) is still conducted by breeders in Switzerland, Belgium and Sweden, and perhaps in some other European countries, however, not in Norway (Boller et al., 2010; T. Vleugels, pers. comm., 2016; L. Öhlund, pers. comm., 2015). The average seed yield of autotetraploid red clover cultivars over the past ten years in Sweden was 27 % higher compared to Norway, and the difference in seed yield between diploid and tetraploid cultivars was approximately 10 % less in Sweden than in Norway (Felleskjøpet Agri and Strand Unikorn, Norwegian seed companies, pers. comm.; I. Andersson, pers. comm.). Knowing this, the question arose if this better performance of Swedish cultivars could be due to the continuous enrichment of the genetic base of the Swedish tetraploid red clover germplasm by neopolyploids. Comparison of seven Norwegian and five Swedish tetraploid cultivars/populations (Paper 1) revealed no significant differences in seed yield between Swedish and Norwegian populations. We were not able to detect any reason for why, in practical growing conditions, Swedish tetraploid red clover produce more seeds in Sweden than Norwegian tetraploid varieties in Norway. There is a common belief that environmental conditions in the red clover seed producing area in Sweden are more optimal for seed production compared to the Norwegian seed producing area. Our dense canopy trials, carried out in only one year, could not either confirm or reject this belief. On the other hand, significantly higher seed yields were obtained from cultivars/populations that were developed by crossing of existing tetraploid plants than from cultivars that were generated using neopolyploids. Lower fertility of neopolyploids compared to their diploid ancestor is a commonly accepted feature (Ramsey and Schemske, 2002; Gottschalk, 1978) as well as that tetraploid red clover cultivars are, in general, lower seed yielding compared to diploid cultivars. Gene doubling, loss of genes, altered gene dosage and gene expression, neofunctionalization and genomic rearrangements are all consequences of polyploidization (Levy and Feldman, 2004 in Adams and Wendel, 2005) and it might be that lower seed yield of tetraploid red clover is affected by some of these events. It might also be that gene(s) controlling fertility in red clover are silenced in the first generations after polyploidization, which result in substantially reduced seed yield compared to the diploid progenitor (Osborne et al., 2003; Liu and Wendel, 2002). Fortunately, after several cycles of natural and breeders' selection, the fertility can be improved, however, in most cases; there is an upper limit for fertility (Ramsey and Schemske, 2002). Ramsey and Schemske (2002) revealed that the evolution of pollen viability and seed

fertility is relatively rapid. They calculated the genetic gain after selection for higher fertility to be 14 % for pollen viability and 39.7 % for seed fertility per generation, concluding that reduced fertility might be temporary. Cultivars developed by crossing of existing tetraploid plants seems to have had longer time to improve their fertility through natural and breeders' selection compared to cultivars developed from neopolyploids. Whether improvement of seed yield of tetraploid red clover is possible by chromosome doubling of high seed yielding diploid plants, as recommended already in 1959 by Povilaitis and Boyes (1959), Wit (1961) and Picard and Berthaut (1966), remains to be explored. As far as we know, this is the first study comparing these two methods for breeding of tetraploid red clover and the results can be useful for further breeding programs.

2.3.3 Molecular responses underlying seed development and seed yield in tetraploid red clover

The seed yield of red clover, particularly of tetraploid cultivars, has not been significantly improved in Scandinavia for a long time. One of the reasons for this could be the negative correlation between forage and seed yield found in many investigations (Clifford and Baird, 1993; Steiner et al., 1997; Vasiljević et al., 2000; Herrmann et al., 2006; Sleper and Poehlman, 2006). Since high forage yield is the primary goal of most red clover breeding programs, simultaneous breeding for high seed yield might be challenging. The outcrossing nature of red clover, inbreeding depression, self-incompatibility system and long selection cycles are additional obstacles for faster genetic gains regarding seed yield (Sleper and Poehlman, 2006; Annicchiarico et al., 2015). The highly heterozygous and heterogeneous genetic background of individual plants belonging to single families have also been limiting progress in developing molecular tools for use in breeding.

Today, the recent availability of red clover genome sequences (Ištvánek et al., 2014; De Vega et al., 2015), transcriptome assemblies (Yates et al., 2014) and advancements in next generation sequencing (NGS) technologies enable us to address the seed yield challenges in red clover with new insights. In order to understand the molecular background of seed yield in red clover, we analyzed two contrasting genotypes regarding the seed yield and performed transcriptome analysis to identify genes that potentially are involved in the regulation of processes involved in development of the seed and determination of the seed yield potential (paper III). *De novo* and reference based transcriptome assembly was conducted to study the global transcriptome changes in flowering buds at three flowering stages (early, middle and late). *De novo* and reference based gene expression data indicated that EF (early flowering) genes in 'Lasang' and

LF (late flowering) genes in 'Tripo' might play a key role for the differences in seed yield in these cultivars. Gene expression profiles, gene ontology enrichment, and KEGG pathway analysis indicated that genes related to flower development, pollen pistil interactions, photosynthesis, and embryo development are differentially expressed in the 'Tripo' and 'Lasang' genotypes, possibly affecting the difference in seed yield observed between these two genotypes. A significant number of genes related to pollination was overrepresented in 'Lasang', which might be a reason for its good seed setting ability.

Further, to validate the DEGs (differentially expressed genes), we developed a physical map based on the DEGs using the reference genome sequence and tried to co-locate DEGs with published seed yield related QTLs (Herrmann et al., 2006) by comparative mapping (paper III). Six SSR markers, which were in the QTL regions for the seed yield components, co-located with six DEGs on four linkage groups. Thus, the candidate genes detected in this study can help in future research for breeding improved tetraploid red clover varieties with enhanced seed yield.

2.4 Main conclusions and future perspectives

Success in breeding for higher seed yields in red clover relies on knowledge and availability of efficient tools to use in selection. The present work has revealed that seed yield per flower head would be the most efficient seed yield component to use in selection for improved seed yield in tetraploid red clover. The same result was obtained in spaced plant and in dense canopy trials, indicating that selection for higher seed yields on single plants could efficiently be used to increase the seed yield per area. Using this information in breeding could increase the seed yield of future tetraploid cultivars. Based on experience and results gained during the present work, selection for higher seed yield of future tetraploid cultivars has been incorporated in Graminor's red clover breeding program. Seed yields of families after pair crosses are analyzed in the seed multiplication steps and progenies/families with low seed yield are discarded. Furthermore, we will establish spaced plant fields in parallel with the progeny testing in order to assess the seed yield of these families.

Lower seed yields of Norwegian tetraploid cultivars, grown in Norway, compared to Swedish cultivars, grown in Sweden, indicate that chromosome doubling should not be reintroduced into the Norwegian red clover breeding program. The genetic background of the Swedish red clover material was thought to be wider due to continuous introduction of neopolyploids by chromosome doubling of superior diploid materials. Results from our studies have revealed

that higher seed yield are obtained by cultivars that are developed by crossing of existing tetraploids rather than by making neopolyploids.

Future work will include cytological studies, which might reveal if any changes in frequency of irregularities took place from neopolyploids to advanced tetraploid breeding materials. In addition, studying the abundancy and species of bumblebees and their behavior could further explain the importance of pollinators for seed yield.

Furthermore, a natural continuation from our transcriptome analysis will be to identify Single Nucleotide Polymorphisms (SNPs) that are associated with seed yield. The markers will also be utilized to enrich existing genetic linkage maps of red clover in order to develop dense linkage maps for QTL mapping, and thus, marker assisted selection.

3. References

Aamlid, T. S. 2015. Dyrkningsveiledning. Frøavl av rødkløver.

Aamlid, T. S., A. Andersen, P. O. Lindemark, S. Jørgensen, L. O. Breivik, J. I. Øverland, A. M. Fremgård, Å Susort and A. A. Steensohn. 2009. Bekjemping av snutebiller i frøeng av rødkløver. Bioforsk FOKUS 4 (1).

Aamlid, T.S., A. Andersen, Å. Susort and A. A. Steensohn. 2010. Avpussing om forsommeren i økologisk frøeng av rødkløver. Bioforsk FOKUS 5 (1): 240-242.

Aase, A. L., F. Ødegaard, A. Mjelde, and Ø. Flagstad. 2011. *Bombus subterraneus* (L., 1758) (Hymenoptera, Apidae) rediscovered in Norway. Norwegian Journal of Entomology, 58: 15-19.

Adams, K. L., and J. F. Wendel. 2005. Polyploidy and genome evolution in plants. Curr Opin Plant Biol, 8: 135-141.

Annicchiarico, P., B. Barrett, E. C. Brummer, B. Julier, and A. H. Marshall. 2015. Achievements and challenges in improving temperate perennial forage legumes. Crit Rev Plant Sci, 34: 327-380.

Bakken, A. K., H. Bonesmo, A. S. Ekker, and A. Langerud. 2005. Fenologisk utvikling hos grovfôrvekstar vurdert etter ein numerisk skala. Grønn Kunnskap 9 (3): 80-90.

Boelt, B., B. Julier, Đ. Karagić and J. Hampton. 2015. Legume seed production meeting market requirements and economic impacts. Crit Rev Plant Sci, 34: 412-427.

Boller, B., F. X. Schubiger, and R. Kölliker. 2010. Red Clover In 'B. Boller et al. (eds.), Fodder Crops and Amenity Grasses, Handbook of Plant Breeding 5, DOI 10.1007/978-1-4419-0760-8_11, Springer Science + Business Media, pp 439-455.

Bragdø-Aas, M. 1970. Fertility of tetraploid red clover. Meld. Norg. LandbrHøgsk. 49(4): 1-34.

Büyükkartal, H. N. 2003. In vitro Pollen Germination and Pollen Tube Characteristics in Tetraploid Red Clover (*Trifolium pratense* L.). Turk J Bot, 27:57-61.

Büyükkartal, H. N. B. 2008. Causes of low seed set in the natural tetraploid *Trifolium pratense* L. Fabacae. Afr J Biotechnol. Vol 7 (9) pp. 1240-1249. doi: 10.5897/AJB07.771

Charlesworth, D. 2002. Self-incompatibility: how to stay incompatible. Curr Biol, 12: 424-426.

Cheung, A. Y., L. C. Boavida, M. Aggarwal, H. M. Wu, and J. A. Feijó. 2010. The pollen

tube journey in the pistil and imaging the in vivo process by two-photon microscopy. J Exp Bot, 61: 1907-1915.

Clifford, P. T. P. and I. J. Baird. 1993. Seed yield potential of white clover: characteristics, components and compromise. In Proceedings of the XVII International Grassland Congress, Palmerston North, New Zealand, 8-21 February. p. 1678-1679.

Darwin, C. R. 1888. The effects of cross and self fertilization in the vegatable kingdom.

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.

Dennis, B. A. 1975. Seed set and related problems in autotetraploid Trifolium pratense L. Doctoral thesis. University of Reading.

Dobin, A., C. A. Davis, F. Schlesinger, J. Drenkow, C, Zaleski, S. Jha, P. Batut, M. Chaisson and T. R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics, 29: 15-21.

Dresselhaus, T. And N. Franklin-Tong. 2013 Male–female crosstalk during pollen germination, tube growth and guidance, and double fertilization. Mol Plant, 6: 1018-1036.

Dunham, W. E. 1939. Insect pollination of red clover in western Ohio. Gleanings, 67: 486-488.

Ellerström, S., and J. Sjödin. 1966. Frequency and vitality of aneuploids in a population of tetraploid red clover. Hereditas, 55: 166-182.

Fairey, D. T., and J. G. Hampton. 1997. Temperate forage seeds: an introduction. Forage Seed Production, 1: 1-7.

Fergus, E. N., and E. A. Hollowel. 1960. Red clover. Adv Agron, 12:365-436.

Franklin-Tong, V. E. 2002. The difficult question of sex: the mating game. Curr Opin Plant Biol, 5: 14-18.

Gottschalk, W. 1978. Open problems in polyploidy research. Nucleus 21:99-112.

Herrmann, D., B. Boller, B. Studer, F. Widmer, and R. Kölliker. 2006. QTL analysis of seed yield components in red clover (*Trifolium pratense* L.). Theor Appl Genet, 112: 536-545.

Higashiyama, T., S. Yabe, N. Sasaki, Y. Nishimura, S. Y. Miyagishima, H. Kuroiwa, and

T. Kuroiwa. 2001. Pollen tube attraction by the synergid cell. Science, 293: 1480-1483.

Huss-Danell, K., E. Chaia, and G. Carlsson. 2007. N2 fixation and nitrogen allocation to above and below ground plant parts in red clover-grasslands. Plant Soil, 299: 215-226.

Isobe, S., I. Klimenko, S. Ivashuta, M. Gau and N. N. Kozlov. 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.

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 9:1-11.

Isobe, S. N., H. Hisano, S. Sato, H. Hirakawa, K. Okumura, K. Shirasawa, S. Sasamoto, A.Watanabe, T. Wada, Y. Kishida, H. Tsuruoka, T. Fujishiro, M. Yamada, M. Kohara and S. Tabata. 2012. Comparative genetic mapping and discovery of linkage disequilibrium across linkage groups in white clover (*Trifolium repens* L.). G3: Genes, Genomes, Genetics, 2: 607-617.

Isobe, S., R. Kölliker, B. Boller, and H. Riday. 2014. Red clover In: H. Cai et al. (eds), Genetics, genomics and breeding of forage crops. ISBN: 978-1-4822-0810-8. CRC Press, Taylor & Francis Group. p. 220-249.

Ištvánek, J., M. Jaroš, A. Křenek, and J. Řepková. 2014. Genome assembly and annotation for red clover (*Trifolium pratense*; Fabaceae). Am J Bot, 101: 327-337.

Julén, G. 1954. Aspects on the breeding of tetraploid red clover with special reference to the seed setting problem. European Grassland Conference, Paris, Organisation for European Economic Co-operation, pp. 69-72.

Julen, G. 1971. Breeding for increased yield of dry matter and of seed in tetraploid red clover. Proc Eucarpia Fodder Crop Sect Meet, 157-167.

Julén, G. 1975. The current situation in the tetraploid red clover. Ploidy in fodder plants. Eucarpia, Zürich, 79-89.

Julén, U. 1950. Fertility conditions of tetraploid red clover. Hereditas 36:151-160.

Kataoka, R., M. Hara, S. Kato, S. Isobe, S. Sato, S. Tabata and N. Ohmido. 2012. Integration of linkage and chromosome maps of red clover (*Trifolium pratense* L.). Cytogenet Genome Res, 137: 60-69.

Kjærgaard, T. 2003. A plant that changed the world: The rise and fall of clover 1000-2000. Landscape Res, 28(1): 41-49.

Kovi, M. R., M. Abdelhalim, A. Kunapareddy, Å. Ergon, A. M. Tronsmo, M. B. Brurberg, I. S. Hofgaard, T. Asp, and O. A. Rognli. 2016. Global transcriptome changes in perennial ryegrass during early infection by pink snow mould. Sci Rep, 6.

Kölliker, R., J. Enkerli and F. Widmer. 2006. Characterization of novel microsatellite loci for red clover (*Trifolium pratense* L.) from enriched genomic libraries. Mol Ecol Notes, 6: 50-53.

Lankinen, Å., and L. Öhlund. 2013. Forskning om pollens kvalitet som hjälpmedel för att förbättra avkastningen i rödklöverfröodlingen. LTJ- fakultetens faktablad. SLU, Alnarp.

Lawrence, M. J. 1996. Number of incopmatibility alleles in clover and other species. Heredity 76: 610-615.

Leduc, N., G. C. Douglas, M. Monnier and V. Connolly. 1990. Pollination in vitro: effects on the growth of pollen tubes, seed set and gametophytic self-incompatibility in *Trifolium pratense* L. and T. repens L. Theor Appl Genet, 80: 657-664.

Lee, M. R., A. L. Winters, N. D. Scollan, R. J. Dewhurst, M. K. Theodorou, and F. R. Minchin. 2004. Plant-mediated lipolysis and proteolysis in red clover with different polyphenol oxidase activities. J Sci Food Agr, 84: 1639-1645.

Lee, M. R., P. R. Evans, G. R. Nute, R. I. Richardson, and N. D. Scollan. 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 Sci, 81: 738-744.

Leffel, R. C. 1963. Pseudo-self-compatibility and segregation of gametophytic self-incompatibility alleles in red clover, *Trifolium pratense* L. Crop Sci, 3: 377-380.

Levan, A. 1945. The present state of plant breleding by induction of polyploidy. Sveriges Utsädesförenings Tidskrift, 55: 109-143.

Liu, B., and J. Wendel. 2002. Non- Mendelian phenomenon in allopolyploid genome evolution. Curr Genomics 3: 489-505.

Li and Hill 1989.

Li, Q. and P. M. J. Hill. 1989. A Study on Post-peak Flowering Shoot Manipulation for Seed Production of *Lotus corniculatus* L. Seed Production, 71.

Lindström, B. E. M.,B. E. Frankow-Lindberg, A. S. Dahlin, M. Wivstad and C. A. Watson,

2013. Micronutrient concentrations in relation to phenological development of red clover (*Trifolium pratense* L.), perennial ryegrass (*Lolium perenne* L.) and timothy (*Phleum pratense* L.). Grass Forage Sci, 69: 276-284.

Lundin, O., M. Rundlöf, H. G. Smith, and R. Bommarco. 2012. Towards integrated pest management in red clover seed production. J Econ Entomol, 1055: 1620-1628.

Lunnan, T. 1989. Effects of photoperiod, temperature and vernalization on flowering and growth in high-latitude populations of red clover. Nor J of Agri Sci, 3: 201-210.

Malengier, M., and J. Baert. 2007. Heriktability of seed yielding capacity in tetraploid red clover (*Trifolium pratense* L.). In: Breeding and seed production for conventional and organic agriculture. Proceedings of the XXVI meeting of the EUCARPIA Fodder Crops and Amenity Grasses Section, Perugia, Italy, 2-7 September 2006. p. 128-129.

Martin, J. N. 1913. The physiology of the pollen of *Trifolium pratense*. Bot Gaz, 112-126.

Meglic, V., and R. R. Smith. 1992. Self.-incompatibility and seed set in colhicine-, nitrous oxide-, and sexually derived tetraploid red clover. Crop Sci, 32: 1133-1137.

Merkenschlager, F. 1934b. Migration and distribution of red clover in Europe. Herbage Reviews, 2: 88-92.

Montardo, D. P., M. Dall'agnol, A. F. Crusius, and N. R. Paim. 2003. Path analysis for seed production in red clover (*Trifolium pratense* L.). In Portuguese, with English abstract). R Bras Zootec, 32: 1076-1082.

Naylor, R., H. Steinfield, W. Falcon, J. Galloway, V. Smith, E. Bradford, J. Alder, and H. Mooney. 2005. Losing the links between livestock and the land. Science, 210:1621-1622.

Oliva, R. N., J. J. Steiner and W. C. Young. 1994. Red clover seed production: II. Plant water status on yield and yield components. Crop Sci, 34: 184-192.

Osborn, T. C., J. C. Pires, J. A. Birchler, D. L. Auger, Z. J. Chen, H. S. Lee, L. Comai, A. Madlung, R. W. Doerge, V. Colot, and R. A. Martienssen. 2003. Understanding mechanisms of novel gene expression in polyploids. Trends Genet, 19: 141-147.

Pandey, K. K. 1955. Seed development in diploid, tetraploid and diploid-tetraploid crosses of *Trifolium pratense* L. Indian J Genet Plant Breed, 15: 25-35.

Pandey, K. K. 1956. Incompatibility in autotetraploid *Trifolium pratense*. Genetics, 41: 353.

Picard, J. and J. Berthaut. 1966. Fertility problems in autotetraploid red clover, and insight into the efficiency of selection at the diploid level. Proc. X Intern. Grassl Congr Helsinki, 664-666.

Plath, O. E. 1925. The role of bumblebees in the pollination of certain cultivated plants. The Ame Naturalist, 59: 441-451.

Povilaitis, B., and J. W. Boyes. 1956. A cytological study of autotetraploid red clover. Am J Bot, 43: 169-174.

Povilaitis, B., and J. W. Boyes. 1959. Embryo-sac production in relation to seed yields of diploid Dollard red clover. Can J Plant Sci, 39: 364-374.

Qin, Y., R. J. Wysocki, A. Somogyi, Y. Feinstein, J. Y. Franco, T. Tsukamoto, D. Dunatunga, C. Levy, S. Smith, R. Simpson, D. Gang, M. A. Johnson and R. Palanivelu. 2011. Sulfinylated azadecalins act as functional mimics of a pollen germination stimulant in Arabidopsis pistils. The Plant J, 68: 800-815.

Ramsey, J., and D. W. Schemske. 2002. Neopolyploidy in flowering plants. Annu Rev Ecol syst, 589-639.

Rao, S., and W. P. Stephen. 2009. Bumble Bee Pollinators in Red Clover Seed Production. Crop Sci, 49: 2207-2214.

Riday, H., and A. L. Krohn. 2010. Genetic map-based location of the red clover (*Trifolium pratense* L.) gametophytic self-incompatibility locus. Theor Appl Genet, 121: 761-767.

Ridout, M. S., X. M. Xu and K. R. Tobutt. 2005. Single-locus gametophytic incompatibility in autotetraploids. J Hered, 96: 430-433.

Sanderson, M. A., and W. F. Wedin. 1989. Phenological stage and herbage quality relationships in temperate grasses and legumes. Agron J, 81: 864-869.

Sato, S., S. Isobe, E. Asamizu, N. Ohmido, R. Kataoka, Y. Nakamura, T. Kaneko, N. Sakurai, K. Okumura, I. Klimenko, S. Sasamoto, T. Wada, A. Watanabe, M. Kothari, T. Fujishiro, S. Tabata. 2005. Comprehensive structural analysis of the genome of red clover (*Trifolium pratense* L.). DNA Res, 12: 301-364.

Semerikov, V., A. Belyaev, and M. Lascoux. 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.

Sjödin, J. and S. Ellerström. 1986. Autopolyploid forage crops. In: Svalöf 1886-1986. Research and Results in Plant Breeding. (Ed. G. Olsson) LTs förlag. Stockholm, Sweden, 102-113.

Sleper, D. A., and J. M. Poehlman. 2006. Breeding field crops (No. Ed. 5). Blackwell publishing.

Smith, R. R. 1994. Registration of 'Marathon' red clover. Crop Sci, 34:1125-1125.

Starling, T. M., C. P. Wilsie, and N. Gilbert.1950. Corolla tube length studies in red clover. Agron J, 42: 1-8.

Steiner, J. J., R. R. Smith, and S. C. Alderman. 1997. Red clover seed production: IV. Root rot resistance under forage and seed production systems. Crop Sci, 37:1278-1282.

Sullivan, M. L., and R. D. Hatfield. 2006. Polyphenol Oxidase and-Diphenols Inhibit Postharvest Proteolysis in Red Clover and Alfalfa. Crop Sci, 46: 662-670.

Takayama, S. and A. Isogai. 2005. Self-incompatibility in plants. Annu Rev Plant Biol, 56: 467-489.

Taylor, N. L. 2008. A century of red clover breeding developments in the United States. Crop Sci, 48:1-13.

Taylor, N. L., and K. H. Quesenberry 1996. Red Clover Science. Vol. 28. Current Plant Science and Biotechnology in Agriculture. Kluwer Academic Publishers, Dordrecht.

Taylor, N. L., and R. R. Smith. 1979. Red clover breeding and genetics. Adv in Agronomy, 31: 125-154.

Valle, O. 1961. Pollination problems in tetraploid red clover. In: S. Ellerström, editor, Precedings of the Symposium on fertility in tetraploid red clover, Eucarpia, Section "Fodder crops" group, Svalof, Sweden. 26-28 Sept. 1961, pp 28-33.

Van Minnebruggen, A., I. Roldan-Ruiz, J. Van Dingenen, E. Van Bockstaele, A. Rohde, and G. Cnops. 2013. Morphological and molecular characterization of branching in red clover (*Trifolium pratense*). In S. Barth and D. Milbourne (eds), Breeding strategies for sustainable forage and turf grass improvement. Springer Netherlands, 161-167.

Vasiljević, S., G. Š. Momirović, S. Katić, V, Mihailović, D. Lukić and T. Zivanović. 2000. Mutual relationships among green forage and seed yield components in genotypes of red clover (*Trifolium pratense* L.). Genetika, 32: 188-191.

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.

Vleugels, T., I. Roldán- Ruiz, and G. Cnops. 2015a. Influence of flower and flowering characteristics on seed yield in diploid and tetraploid red clover. Plant Breed, 134: 56-61.

Vleugels, T., I. Roldan-Ruiz, B. Ceuppens, G. Smagghe, and G. Cnops. 2015b. Are corolla tube dimensions the reason for low seed yield in tetraploid red clover? Proceedings of the 31st international EUCARPIA Symposium Section Fodder Crops and Amenity Grasses. Ghent, Belgium. In press. 6p.

Vleugels, T., B. Ceuppens, G. Cnops, P. Lootens, F. R. van Parijs, G. Smagghe and I. Roldán-Ruiz. 2016. Models with only two predictor variables can accurately predict seed yield in diploid and tetraploid red clover. Euphytica, 209: 507-523.

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

Wilsie, C. P., and N. W. Gilbert. 1940. Preliminary results on seed setting in red clover strains. J Am Soc Agron, 32: 231-4.

Wit, F. 1961. The cytological background to fertility disturbances in tetraploid clover. Proc Eucarpia Fodder Crop Sect, 13-19.

Yates, S. A., M. T. Swain, M. J. Hegarty, I. Chernukin, M. Lowe, G.G. Allison, T. Rutnik, M. T. Abberton, G. Jenkins and L. Skøt. 2014. De novo assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC genomics, 15:453.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J 2006. WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34:W293-297.

Young, N. D., J. Mudge, and T. N. Ellis. 2003. Legume genomes: more than peas in a pod. Curr Opin Plant Biol, 6: 199-204.

Zohary, M., and D. Heller. 1984. The genus Trifolium. Israel Academy of Sciences and Humanities.

Ødegaard, F., J. O. Gjershaug, S. Öberg and A. Mjelde. 2009. Status for humler (*Hymenoptera, Apidae*, Bombus spp.) i Norge. Fauna 62: 94-104.

Åman, P., and E. Nordkvist. 1983. Chemical composition and in-vitro degradability of major chemical constituents of red clover harvested at different stages of maturity. J Sci Food Agri, 34: 1185-1189

Published February 19, 2016

RESEARCH

Seed Yield of Norwegian and Swedish Tetraploid Red Clover (*Trifolium pratense* L.) Populations

Helga Amdahl, Trygve S. Aamlid, Åshild Ergon, Mallikarjuna R. Kovi, Petter Marum, Muath Alsheikh,* and Odd Arne Rognli

ABSTRACT

While tetraploid plants of red clover are taller, have thicker stems, and have broader leaves that altogether result in a higher forage yield compared to diploids, they generally have substantially lower seed yields than diploid plants. Tetraploid red clover can be induced chemically by colchicine or nitrous oxide (N_2O) and sexually by union of unreduced gametes. The average seed yield of tetraploid red clover in Norway is 60% of the diploid yield, while in Sweden it is 75%. One objective of this paper was to examine whether there is a difference in seed yield among chromosome doubled tetraploids and crossed tetraploids. A second objective was to investigate differences in seed yield and seed yield components in Norwegian and Swedish tetraploid populations. The third objective was to study which yield component most correlates with the seed yield per hectare. Seed production experiments were established at Landvik and Bjørke in Norway and Svalöv and Lännäs in Sweden. Populations made by crossings of tetraploids gave significantly greater yield ($p < 0.001$) compared to populations that were made by chromosome doubling. On average, Norwegian and Swedish varieties had equal yields in both experimental years. Norwegian and Swedish varieties differed mostly in earliness traits. Swedish populations began flowering on average 4 d earlier than Norwegian populations. Genotypic correlations showed that seed yield per flower head was the component with the highest correlation ($r = 0.956$ and $r = 0.977$) with yield per hectare in both experimental fields. Results from the second experimental year indicate a trend towards improved seed yield after several cycles of recurrent selection for higher seed yield per flower head.

H. Amdahl and M. Alsheikh, Graminor Breeding Ltd., Hommelstadvegen 60, NO-2322 Ridabu, Norway and Norwegian Univ. of Life Sciences, Dep. of Plant Sciences, P.O. Box 5003, N1432 Ås, Norway; T.S. Aamlid, NIBIO Landvik, NO-4886 Grimstad, Norway; Å. Ergon, M.R. Kovi, and O.A. Rognli, Norwegian Univ. of Life Sciences, Dep. of Plant Sciences, P.O. Box 5003, N1432 Ås, Norway; P. Marum, Graminor Breeding Ltd., Hommelstadvegen 60, NO-2322 Ridabu, Norway. Received 21 July 2015. Accepted 26 Oct. 2015. *Corresponding author (muath.alsheikh@graminor.no).

Abbreviations: Date5, beginning of flowering; P × L, Population × Location; HS, percent of hard seeds; Mat, percent of mature seed at harvest; NoMax, number of flower heads at maximum flowering; NS, percent of normal seedlings; SY, seed yield; SYFH, seed yield per flower head; TSW, thousand seed weight.

Red clover (*Trifolium pratense* L.) is a perennial forage legume grown mostly in mixtures with grasses such as timothy (*Phleum pratense* L.), meadow fescue (*Festuca pratensis* Huds.), tall fescue (*Festuca arundinacea* Schreb), and perennial ryegrass (*Lolium perenne* L.). The importance of red clover as a forage is due to its high protein content (Taylor and Quesenberry, 1996). Red clover increases the amount of crude protein and the amount of calcium in mixtures with grasses (Lunnan, 2000). There is a higher intake by farm animals of red clover compared to forage grasses, resulting in higher production of meat and milk (Randby, 1991; Peyraud et al., 2009). The ability of red clover to fix nitrogen makes it particularly attractive in organic forage production. Red clover does not tolerate grazing very well; hence, it is mostly used for cutting at two to four times during the harvest year (Boller et al., 2010). Red clover is a natural diploid, and both diploid and tetraploid varieties are grown. New tetraploid populations can be created in two ways: mitotically by doubling the chromosome number using colchicine, nitrous oxide, or mitosis inhibitors, and meiotically by crossing diploid plants that produces unreduced gametes (Meglic and Smith, 1992;

Published in Crop Sci. 56:603–612 (2016).
doi: 10.2135/cropsci2015.07.0441

Taylor and Quesenberry 1996). Tetraploid red clover is taller, has thicker stems and broader leaves, and produces greater forage yields compared to diploid red clover. In addition, tetraploids have larger flower heads and bigger seeds (Sjödin and Ellerström, 1986; Taylor and Quesenberry, 1996). Finally, tetraploids are more persistent than diploids due to better resistance to stem nematodes and crown rot (*Sclerotinia trifoliorum*) (Sjödin and Ellerström, 1986; Vestad, 1990; Taylor and Quesenberry, 1996; Vleugels et al., 2013). The first tetraploid varieties released in the Swedish market produced, on average, 10% more dry matter yield in the first harvest year, and up to 15% more dry matter yield in the second harvest year, compared to diploid varieties. The fresh weight of tetraploid single plants was, on average, 60% higher than that of diploid plants, whereas tetraploid families had 128% higher fresh weights than diploid families (Sjödin and Ellerström, 1986). Norwegian trials showed that tetraploid red clover is more winter hardy and that tetraploids have, on average, 20% higher dry matter yield in mountain areas compared to diploid varieties (Marum, 2006).

Approximately 100 Mg of red clover seed is sold annually in Norway, of which 5 to 10% are tetraploid varieties. In Sweden, approximately 480 Mg of diploid seeds and 120 Mg of tetraploid seeds are sold every year. In Norway, the average seed yield of tetraploid varieties was 164 kg ha^{-1} over the last 5 to10 years, as compared to 247 kg ha^{-1} of diploids (Felleskjøpet Agri and Strand Unikorn [Norwegian seed company], personal communication, 2015). In Sweden, the average seed yield of tetraploid varieties was 225 kg ha^{-1} for tetraploids and 300 kg ha^{-1} for diploids, and the seed yield of tetraploid red clover has decreased slightly during the last 5 to 10 years (I. Andersson, personal communication, 2015).

The main disadvantage of tetraploid red clover plants is the low seed yield compared to diploid clover (Sjödin and Ellerström, 1986; Vestad, 1990). The seed yield of tetraploids is approximately 40% lower than that of diploids under open pollination (Wexelsen and Vestad, 1954; Valle, 1961; Sjödin and Ellerström, 1986). Surprisingly however, in some years and at some places, tetraploid varieties have equal seed yield as diploid varieties (Sjödin and Ellerström, 1986; Marum, 2006; B. Andersson, personal communication, 2011). The possible explanations for the low seed set in the tetraploid red clover include poor pollination due to flower morphology (e.g., longer corolla tube), lower number of ovules per flower head or per plant, or pre- and post-fertilization developmental irregularities (Bragdø-Aas, 1970; Büyükkartal, 2003; Boelt et al., 2015).

In Norway, the practice of doubling the chromosome number of diploid red clover began in 1947. Since then, there have been six tetraploid varieties in commercial production. The last time chromosome doubling was performed in Norway was in the early 1990s. One of the last tetraploid varieties made by this method was 'Lavine' (Table 1). In contrast, the Swedish breeding program still includes chromosome doubling of the best performing diploid breeding lines or varieties (L. Öhlund, personal communication, 2015). Breeders started to work on improving the seed yield in Norwegian tetraploid red clover material in the early 1970s. From 1979 until 1987, two cycles of selection for seed weight per plant were done, in addition to several cycles of selection for forage yield and nematode resistance (Vestad, 1990). One result of this work was that breeders in 1995 and 1997 put together 170 single plant families to form the 'L-4374' breeding population. This population was then used to perform two cycles of selections for seed weight per flower head (Aamlid and Marum, unpublished data, 1998). The variety 'Lasang' was the result of this work while the 'LøRk0733' breeding population was the outcome after one cycle of natural selection for seed yield in 'Lasang'. The low seed yields of high-yielding and persistent varieties of tetraploid red clover is a major problem in Norwegian and Swedish forage production, and it has been questioned if this is due to ineffective breeding methods or to non-optimal localization of the tetraploid red clover seed production in the two countries. Thus, the objectives of this paper were to investigate (i) differences in seed yield between tetraploid varieties developed by chromosome doubling of diploids and varieties developed by crossing of tetraploids, (ii) the correlation between various seed yield components and seed yield per hectare of tetraploid varieties, and (iii) differences in seed yield between Swedish and Norwegian tetraploid varieties grown in various environments.

MATERIALS AND METHODS

Experimental Sites

Field trials with Swedish and Norwegian tetraploid red clover populations were established in 2011 (Exp. 1) and in 2012 (Exp. 2) and harvested in 2012 and 2013, respectively. Detailed descriptions of the varieties are presented in Table 1. Exp. 1 was conducted at four locations: Landvik (58.34° N, 08.52° E; 10 m.a.s.l.) and Bjørke (61.22° N, 20.42° E; 147 m.a.s.l.) in Norway, and Svalöf (55.56° N, 13.06° E; 70 m.a.s.l.) and Lännäs (63.09° N, 17.39° E; 26 m.a.s.l.) in Sweden. Exp. 2 was conducted only at Landvik and Bjørke. The soil type at all locations was silty loam (Orthic Humic Gleysols at Landvik [Hole and Solbakken, 1986] and Haplic Phaeozems at Bjørke [E. Solbakken, personal communication, 2015]; no available information for Svalöv and Lännäs).

All trials in both years were sown with 5 kg seed ha^{-1}. In the first year of study (Exp. 1), crops were sown at the four locations Landvik, Bjørke, Svalöv, and Lännäs on 22, 29, 5, and 13 June and harvested on 6 Sept., 1 Oct., 6 Sept., and 11 Oct., respectively. In the second year of study (Exp. 2), crops were sown at Landvik and Bjørke on 11 and 21 June and harvested on 13 and 5 Sept., respectively. On average for the last decade, red clover seed yields at the four experimental sites have

Table 1. Description of tetraploid red clover populations used in the seed production experiments.

Country of origin	Population	Year of approval	Population description	Population background	Exp.†	Breeding method‡
Norway	'Tripo'	1964	First Norwegian tretraploid (4×) variety. No longer in commercial production (Vestad, 1990).	No available information.	1, 2	D
	'Reipo'	2002	Low winter hardiness, high dry matter yield, middle late, good tolerance to diseases§	Developed after one and five selection cycles (SC) for seed yield and forage yield, respectively, on 3600 single tetraploid plants that belonged to nine unknown local populations (Vestad, 1990).	1, 2	C
	'Lavine'	2007	Winter hardy variety that never came to the market due to low seed yield§	Developed from doubling of 'Kongsvoll-2×' (a collection of several diploid populations from mountainous areas in Norway).	1, 2	D
	'L-4374'	Breeding population	Germplasm with improved seed yield resulting from selection for this trait (described earlier in the present study).	Developed after two SC for vegetative characters, four SC for high forage yield, one SC for nematode resistance and two SC for high seed yield on 3600 single tetraploid plants belonging to nine unknown populations (Vestad, 1990; Aamlid and Marum, unpublished data, 1998).	2	C
	'Lars'	2012	More winter hardy than 'Lasang', middle early, high dry matter yield and good tolerance to diseases. A promising variety in Norway§	A cross between 6 plants of the tetraploid variety 'Lone' and 6 plants of the tetraploid variety 'Betty'. All plants were selected for their competitive ability when grown with timothy.	1	C
	'Lasang'	2013	Winter hardy and persistent variety with medium earliness, high dry matter yield and good tolerance to diseases. Not in commercial production§	Developed from 'L-4374' after two SC for high number of seeds per flower head (Aamlid and Marum, unpublished data, 1998).	1, 2	C
	'LøRk0733'	Breeding population	No available information.	Developed after one natural SC on 'Lasang'.	2	C
Sweden	'Betty'	1992	Good winter hardiness and persistence with early flowering, has moderate dry matter yield, used in northern Sweden§	Developed from doubling the diploid variety 'Bjursele'. There was no particular selection for higher seed yield in either 'Bjursele' or 'Betty' (L. Öhlund, personal communication, 2015).	1, 2	D
	'SW Nancy'	2001	Middle late variety with good persistence, used in south Sweden (L. Öhlund, personal communication, 2015).	Developed from the crossings of 7 single plants of tetraploid 'Fanny'. Plants were selected for high seed yield combined with good vegetative characters. Variety 'Fanny' was made by crossing a certain number of plants selected for high nematode resistance and high seed yield (L. Öhlund, personal communication, 2015).	1, 2	C
	'SW Torun'	2002	Considerably late flowering and higher dry matter yield than 'Betty'. Used in north Sweden (L. Öhlund, personal communication, 2015).	Developed from doubling the diploid variety 'Jesper'. There was no particular selection for higher seed yield in 'Jesper' or 'SW Torun' (L. Öhlund, personal communication, 2015).	1, 2	D
	'Vicky'	2009	Better persistence than other varieties on the official variety list in Sweden (L. Öhlund, personal communication, 2015).	Developed from doubling diploid variety 'Eva'. 'Eva' was selected for improved persistency (L. Öhlund, personal communication, 2015).	1	C
	'SW RK1111'	Breeding population	No available information.	Developed from breeding population 'SW RK1051'. Plants from 'SW RK1051' were established as polycross. Seeds harvested from surviving plants after the second harvest year were put together to form 'SW RK1111' (L. Öhlund, personal communication, 2015).	1	C

† Experiment that included the population.

‡ Tetraploid populations developed by crossing existing tetraploid plants (C) or by chromosome doubling of diploid plants (D).

§ Population described by the Norwegian official program for variety testing for red clover: 'Reipo' and 'Betty' in 2002, 'Lavine' in 2007, 'Lars' in 2012, and 'Lasang' in 2013; Please contact the authors for additional source information.

Table 2. Traits of red clover seed production.

Type of trait	Trait	Description
Seed yield and seed yield components	Seed yield (SY)	All plots at each location were harvested at the same time with harvesters when the seeds were considered ripe. After harvest, the seeds were dehulled and cleaned roughly at each location before being sent to Landvik for final cleaning and laboratory analysis. Seed yields were corrected to 100% purity and 12% seed moisture content.
	Number of flower heads at maximum flowering (NoMax)	In the spring of the harvesting year, a 0.75 × 0.75 m subplot within each plot was designated. Flower heads were counted inside each subplot and numbers were converted to the number of flower heads per m^2.
	Seed yield per flower head (SYFH)	50 randomly chosen ripe flower heads were collected from each plot just before harvest. Seeds were threshed by hand, cleaned, weighed and the seed yield per flower head calculated.
	Thousand seed weight (TSW)	100 seeds were randomly selected from the pure seed yield of each of the four blocks per populations. The seeds were weighed and the thousand seed weight calculated.
Developmental earliness	Beginning of flowering (Date5)	Registered as the number of days after 31 May for the first five opened flower heads. The plots were observed every second to third day.
	Maturity (Mat)	Percent of black flower heads with the upper peduncle completely dry before harvest. Approximately 100 random flower heads in each plot were observed and classified as black/ripe, green/withered, or red/still flowering.
Seed quality	Normal seedlings (NS)	100 randomly chosen seeds from each variety were incubated on a germination table at 20°C. Normal seedlings were counted after 7 and 14 d.
	Hard seeds (HS)	Percent of hard seeds was calculated according to International Rules for Seed Testing (Norwegian Ministry of Agriculture and Food, 2015) after counting the normal seedlings.

been 350 kg ha^{-1} for diploid and 200 kg ha^{-1} for tetraploids at Landvik; 280 kg ha^{-1} for diploid and 150 kg ha^{-1} for tetraploid at Bjørke; and 320 kg ha^{-1} for both diploids and tetraploids at Svalöv. There were no historical data for seed production of red clover at Lännäs.

Experimental Protocol

The experiments were established according to a randomized complete block design with four blocks. The plot size was 1.5 × 6.6 m. The border plots of each block were seeded with the tetraploid variety 'Reipo' in Norway (Bjørke and Landvik) in both years, while the border plots of each block were seeded with the tetraploid variety 'Betty' at Lännäs and with the diploid red clover variety 'SW Ares' at Svalöv. Eight traits were quantified in the two experiemnts to assess the seed production of red clover poluations (Table 2).

Statistical Analysis

Analysis of variance was performed for each trial year using the PROC GLM in SAS version 9.4 (SAS Institute, 2014). Populations and locations were considered fixed effects, and block was a random effect nested under location. Differences between mean values were tested using the Ryan-Einot-Gabriel-Welsch multiple range test. Differences between populations resulting from doubling of diploids and populations resulting from crossing of tetraploids, and between Norwegian and Swedish populations were tested using orthogonal contrasts. For each trial year, correlation coefficients between all trait values averaged across blocks and environments were calculated using PROC CORR in SAS. The Population × Location (P × L) interaction for seed yield was analyzed using the Eberhart and Russell (1966) model of stability in the R package Plantbreeding (R-Forge, 2012), a software package in R for the analysis and visualization of data from plant breeding and genetics experiments.

RESULTS

Experiment 1

Main Effect of Location

The effect of location was highly significant ($P < 0.001$) for all traits. The highest seed yield (SY) was obtained at Landvik (459 kg ha^{-1}) and the lowest was at Svalöf (130 kg ha^{-1}). Seed yield per flower head (SYFH) was significantly higher ($P < 0.001$) at Svalöv (196 mg) compared to the other three locations (Landvik, 145 mg; Bjørke, 141 mg; Lännäs, 133 mg). The flowering period started earliest at Svalöv, the most southern location. There was a 14 d difference in the beginning of flowering (Date5) between the most southern (Svalöv) and the most northern location (Lännäs). The highest percentage of mature seed heads (Mat) at harvest (88%), was registered at Svalöv while this percentage was significantly lower ($P < 0.001$) at Landvik and Bjørke. Only seeds from Svalöv (92%) and Landvik (83%) met the international certification requirement, based on OECD and EEA regulations, of minimum 80% germination including a maximum of 20% hard seeds (Norwegian Ministry of Agriculture and Food, 2015).

Main Effect of Population

The effect of population was significant ($P = 0.020$) for normal seedlings (NS) and highly significant ($P < 0.001$) for the other traits, except for number of flower heads at maximum flowering (NoMax) and hard seeds (HS). The Norwegian varieties 'Lasang' and 'Lars' and the Swedish variety 'SW Nancy' gave the highest SY, which was significantly higher ($P < 0.001$) only from 'Lavine'. 'SW Nancy' and 'Lasang' showed the highest SYFH

Table 3. Mean values of seed production traits of each population of Norwegian and Swedish tetraploid red clover, mean values for populations by country of origin, mean values for popoluations developed by chromosome doubling of diploids vs. crossing of tetraploids, and the main effect of location on seed production traits (Exp. 1).

Population	Origin	Breeding method	SY†	NoMax	SYFH	TSW	Date5	Mat	NS	HS
			kg ha^{-1}	No. m^{-2}	mg		d	%		
Lasang	N‡	C§	305 a¶	540 a	174 a	3014 bc	41 bc	63 bc	54 ab	27 a
Lars	N	C	291 a	545 a	163 ab	2858 d	43 ab	68 abc	57 ab	26 a
Reipo	N	C	275 ab	521 a	152 ab	2967 cd	42 ab	62 bc	54 ab	29 a
Tripo	N	D	257 ab	487 a	147 b	2935 cd	44 a	56 c	52 b	27 a
Lavine	N	D	227 b	515 a	123 c	2891 cd	40 cd	72 ab	56 ab	29 a
SW Nancy	S	C	300 a	569 a	176 a	3024 bc	38 de	67 abc	53 ab	30 a
Vicky	S	D	281 ab	519 a	152 ab	3193 a	36 fg	68 abc	60 a	26 a
Betty	S	D	280 ab	533 a	151 ab	2916 cd	35 g	75 a	58 ab	28 a
SW RK1111	S	C	274 ab	525 a	161 ab	3106 ab	37 ef	63 abc	57 ab	24 a
SW Torun	S	D	251 ab	458 a	139 bc	2917 cd	40 cd	65 abc	55 ab	28 a
Norwegian	N		272 a	522 a	153 a	2962 a	41 a	64 a	55 a	27 a
Swedish	S		278 a	520 a	154 a	3013 a	37 b	69 b	56 a	28 a
Doubled		D	259 a	502 a	142 a	2971 a	39 a	67 a	56 a	28 a
Crossed		C	289 b	540 b	165 b	2994 a	40 b	65 a	55 a	27 a
Landvik			459 a	1113 a	145 b	2772 c	39 b	58 b	63 b	23 b
Bjørke			291 b	259 c	141 b	3113 a	40 b	52 b	33 d	55 a
Lännäs			215 c	481 b	133 b	2920 b	47 a	–	54 c	10 c
Svalöv			130 d	232 c	196 a	3123 a	33 c	88 a	72 a	23 b
F_{27} (P × L)			2.49**	1.29	1.37	0.86	2.12**	2.71**	1.37	0.99

** Significant at the 0.01 probability level.

† SY, seed yield; NoMax, number of flower heads at maximum flowering; SYFH, seed yield per flower head; TSW, thousand seed weight; Date5, beginning of flowering; Mat, percent of mature seeds at harvest; NS, percent of normal seeds; HS, percent of hard seeds.

‡ N, Norway; S, Sweden.

§ C, variety developed by crossing of existing tetraploid plants; D, variety developed by chromosome doubling of diploid plants.

¶ Within each column and by population, country of origin, breeding method, or location, means followed by the same letter are not significantly different at the 0.05 level.

significantly different ($P < 0.001$) from 'Tripo', 'Lavine' and 'SW Torun'. The population that started flowering earliest was 'Betty, it flowered 35 d after 31 May, while 'Tripo' flowered last, 9 d after 'Betty'. The Swedish populations started to flower 4 d earlier than Norwegian populations. SY, NoMax and SYFH were significantly higher ($P < 0.001$, $P = 0.020$ and $P < 0.001$, respectively) in populations developed by crossing of existing tetraploids than in populations developed by chromosome doubling of diploids (Table 3).

When trait values for each variety were averaged across locations, SY was highly positively correlated ($r = 0.956$, $P < 0.001$) with SYFH, and less strongly correlated, but still significant ($P = 0.023$) with NoMax ($r = 0.701$).

Population × Location Interaction

The P × L interaction was significant ($P < 0.01$) for Date5 and highly significant ($P < 0.001$) for SY and Mat, but not for the other traits (Table 3). The average SY of the varieties at the four locations is presented in Table 4, and the results of the analysis of variance (regression analysis) of the P × L interaction for seed yield in Tables 5 and 6. The general interpretation of the stability parameters estimated by the Eberhart and Russell (1966) regression method is that the most optimal variety has a high mean yield across locations, a regression coefficient (b_{ij}) of 1.0 indicating predictable performance across environments, and smallest possible deviations (sd_{ij}) from regression (stability). In the present analysis, the regression coefficient, b_{ij}, ranged from 0.65 for 'SW Torun' to 1.27 for 'SW Nancy' (Table 6); however, these coefficients were not significantly different as evident from the nonsignificant linear variance component of P × L (Table 5). The deviations from regression were significant ($P < 0.01$, Table 5) and the three varieties 'Reipo', 'Lavine' and 'SW Torun' had significant deviations from regression and were thus unstable (Table 5).

Table 4. Mean seed yield of tetraploid populations of red clover at four locations in Norway and Sweden and overall averages by population and location (Exp. 1).

	Norway		Sweden		
Population	Bjørke	Landvik	Lännäs	Svalöv	Average
	kg ha^{-1}				
Lasang	290	525	268	140	305
Lars	310	513	218	128	291
Reipo	333	438	183	155	275
Tripo	260	475	165	125	257
Lavine	195	380	228	105	227
SW Nancy	295	548	215	138	300
Vicky	275	465	263	128	281
Betty	320	463	210	128	280
SW RK1111	310	445	200	143	274
SW Torun	323	343	210	128	251
Average	291	459	215	130	274

Experiment 2

Main Effect of Location

The effect of location was significant ($P < 0.05$) for SYFH and Mat and highly significant ($P < 0.001$) for the remaining traits considered in this study (Table 7).

In contrast to 2012, in 2013 SY was higher at Bjørke than at Landvik. At Bjørke, flowering started 4 d earlier. The percentage of NS was higher at Landvik while the percentage of HS was higher at Bjørke.

Main Effect of Population

The effect of population was highly significant ($P < 0.001$) for SY, thousand seed weight (TSW), Date5, and for Mat, significant ($P < 0.01$) for SYFH and significant ($P < 0.05$) for percent HS level (Table 7).

'SW Nancy' had the highest SY in 2013, together with 'LøRk0733'. SYFH was highest in 'LøRk0733' and significantly higher ($P < 0.001$) than in 'SW Torun' and 'Lavine'. SYFH was the only significant ($P < 0.001$) trait positively correlated ($r = 0.977$) with the SY. TSW was significantly higher in 'Lasang' than in all other populations in the experiment. The population with the earliest Date5 was 'Betty', while 'Reipo' was 7 d later than 'Betty'. 'Betty' had the highest Mat at harvest, together with 'Lasang' and 'LøRk0733'. 'Tripo' had the lowest Mat (65%) but this was significantly lower ($P < 0.001$) only when compared to 'Betty' and 'Lasang'. 'Reipo' had the highest percentage of NS (58%) but was significantly different ($P < 0.05$) only from 'Betty', which had the lowest percentage of 46%. The percentage of HS was significantly higher in 'Betty' ($P < 0.01$) compared to 'Reipo', 'L-4374' and 'SW Torun'.

Table 5. Analysis of variance of average seed yield of tetraploid populations of red clover at four locations (Env) in Norway and Sweden (Exp. 1) to estimate phenotypic stability parameters according to the method used by Eberhart and Russell (1966).

Source	df	MS	*F*	*Pr > F*
Total	39	165.20		
Population	9	22.80	2.05	0.09
Location + (Population × Env)	30	207.90		
Location (linear)	1	5830.50		
Population × Location (linear)†	9	20.50	1.84	0.12
Pooled deviation‡	20	11.10	2.02	0.01**
Lasang	2	10.50	1.91	0.15
Lars	2	0.30	0.05	0.95
Reipo	2	17.20	3.13	0.05*
Tripo	2	9.20	1.67	0.19
Lavine	2	21.70	3.96	0.02*
SW Nancy	2	6.50	1.19	0.31
Vicky	2	11.00	2.00	0.14
Betty	2	3.50	0.64	0.53
SW RK1111	2	3.90	0.71	0.49
SW Torun	2	27.30	4.96	0.01**
Pooled error	120	5.50		

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† Tested against the pooled deviation mean square.

‡ Tested against the pooled error mean square.

Table 6. Stability parameters estimated for seed yield of tetraploid populations of red clover at four locations in Norway and Sweden.

Population	Origin†	Breeding method‡	b_{ij}§	sd_{ij}¶
Lasang	N	C	1.14	4.40
Lars	N	C	1.18	–5.84
Reipo	N	C	0.92	11.11*
Tripo	N	D	1.11	3.06
Lavine	N	D	0.77	15.63*
SW Nancy	S	C	1.27	0.43
Vicky	S	D	0.98	4.88
Betty	S	D	1.03	–2.57
SW RK1111	S	C	0.95	–2.19
SW Torun	S	D	0.65	21.15**

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† N, Norway; S, Sweden.

‡ C, variety developed by crossing of existing tetraploid plants; D, variety developed by chromosome doubling of diploid plants.

§ Regression coefficients resulting from the regression of seed yield of each variety on the environmental index.

¶ Deviation from regression on the environmental index.

Table 7. Mean values of seed production traits of each population of Norwegian and Swedish tetraploid red clover, mean values for populations by country of origin, mean values for populations developed by chromosome doubling of diploids vs. crossing of tetraploids, and the main effect of location on seed production traits (Exp. 2).

Population	Origin	Breeding Method	SY†	NoMax	SYFH	TSW	Date5	Mat	NS	HS
			kg ha^{-1}	No. m^{-2}	mg		d	%		
Tripo	N‡	D§	540 ef¶	539 a	170 ab	2924 bc	57 a	65 c	52 ab	36 ab
Reipo	N	C	566 def	617 a	178 ab	2855 c	58 a	73 bc	58 a	31 b
Lavine	N	D	424 g	574 a	140 c	2881 bc	53 b	77 abc	51 ab	36 ab
L-4374	N	C	616 bcd	617 a	182 ab	2846 c	54 b	74 bc	55 ab	31 b
Lasang	N	C	624 bc	579 a	185 ab	3033 a	53 b	81 ab	51 ab	37 ab
LøRk0733	N	C	665 ab	629 a	191 a	2955 b	57 a	76 abc	48 ab	38 ab
Betty	S	D	576 cde	538 a	170 ab	2845 c	48 c	86 a	46 b	44 a
SW Torun	S	D	516 f	562 a	169 b	2956 b	52 b	73 bc	54 ab	32 b
SW Nancy	S	C	691 a	574 a	177 ab	2948 b	54 b	70 bc	51 ab	38 ab
Norwegian	N		568 a	600 a	174 a	2913 a	38 a	75 a	53 a	35 a
Swedish	S		579 a	551 a	172 a	2921 a	35 b	76 a	51 a	37 a
Doubled		D	504 a	554 a	162 a	2904 a	36 a	75 a	51 a	35 a
Crossed		C	612 b	603 b	181 b	2912 b	38 b	75 a	52 a	35 a
Landvik			348 b	823 a	177 a	2881 a	39 a	77 a	62 a	19 b
Bjørke			794 a	339 b	170 a	2951 a	35 b	73 a	42 b	52 a
F_8 (P × L)			6.08***	1.25	1.31	6.84***	1.00	0.68	0.80	2.00

*** Significant at the 0.001 probability level.

† SY, seed yield; NoMax, number of flower heads at maximum flowering; SYFH, seed yield per flower head; TSW, thousand seed weight; Date5, beginning of flowering; Mat, percent of mature seeds at harvest; NS, percent of normal seeds; HS, percent of hard seeds.

‡ N, Norway; S, Sweden.

§ C, variety developed by crossing of existing tetraploid plants; D, variety developed by chromosome doubling of diploid plants.

¶ Within each column and by population, country of origin, breeding method, or location, means followed by the same letter are not significantly different at the 0.05 level.

Population × Location Interaction

The P × L interaction was highly significant ($P < 0.001$) for SY and TSW (Table 7). The ranking of populations as with regard to differences in SY was similar at both locations; therefore, the P × L interaction was not analyzed further.

DISCUSSION

Development Method: Doubling of Diploids vs. Crossing of Tetraploids

Tetraploid red clover populations made by crossing of tetraploids produced higher SY compared to those made by doubling of diploids. This finding is of interest for plant breeders who work with tetraploid red clover to improve SY. Low SY in newly produced tetraploids encouraged researchers to develop ways to enhance seed yield in future varieties. Picard and Berthaut (1966) and Julén (1954) suggested selection for higher SY at the diploid level as a means to improve SY in tetraploid red clover. Tetraploid populations in our study, denoted as doubled, originated from diploid varieties that were not deliberately selected for high SY. For that reason, the former suggestion could not be tested here. In their first seed yield trials with tetraploid red clover, Julén (1954), Pandey (1955), Eskilsson (1963), Sjödin (1981) and Vestad (1990) observed higher SY in tetraploids after several years of seed multiplication and selection for higher SY. The suggested explanation is that in raw tetraploids (first generations after doubling) multivalent formations occur in meiosis that could affect the fertility of tetraploids. In later generations, however, the frequency of multivalent formations is reduced in favor of bivalents, which affect seed formation in a positive way (Julén, 1954; Wexelsen and Vestad, 1954; Bragdø-Aas, 1970; Jauhar, 1970). Pandey (1955) studied colchicine derived tetraploids and detected collapsed endosperms in many ovules and degeneration of many ovules after reaching the cell formation stage. However, there was a large variation between different tetraploid plants, which suggests that selection for higher SY could increase the total SY per plant.

Seed Yield by Country of Origin

Our results show that, in Norway, there is a potential for higher SY in tetraploid red clover compared to achievements under practical seed production. Differences in SY between experimental fields and practical conditions were especially prominent at Norwegian locations where this difference varied between 43% and 289% in favor of the experimental fields (Felleskjøpet Agri and Strand Unikorn [Norwegian seed company], personal communication, 2015). SY is often higher in experimental trials compared to commercial seed production fields (Boelt and Gislum, 2011). Perhaps this could be due to the smaller size of experimental fields, which allows more optimal harvesting time and better pollinating conditions. In addition, harvesters designed especially for experimental fields could help reduce the loss of seeds. Among locations where our experiments were conducted, Svalöv should normally have the most favorable climatic conditions for seed production. The results, however, showed the lowest SY at Svalöv. A possible explanation could be that the border plots in the trial at Svalöv were seeded with the diploid red clover 'SW Ares', instead of a tetraploid variety. Diploid red clover has on average shorter corolla tube that might render them more attractive to pollinators than tetraploid red clover. A study by Vleugels et al. (2015) concluded that the length of corolla tube is most probably not one of the reasons for lower seed set in tetraploids. Presence of haploid pollen in a tetraploid field can, if involved in successful fertilization, lead to triploid embryo that often collapse after a short period (Julén, 1950). That Svalöv had the highest SYFH does not support this theory. Another reason for seed loss may be the intensive seed cleaning at Svalöv, as the amount of husks before the final cleaning of seeds was relatively low compared to the other three locations. Delayed harvest due to high and continuous precipitation could have negatively affected SY and NS at Lännäs. The percentage of HS is usually high when the harvesting conditions are unfavorable (Rolston, 1978). Results, however, show the lowest percentage of HS at Lännäs.

According to statistics from Norwegian and Swedish seed companies, tetraploid red clover has higher seed yields in Sweden (225 kg ha^{-1}, I. Andersson, personal communication, 2015) compared to Norway (164 kg ha^{-1}). Results from our trials did not indicate significant differences in SY between populations from the two countries. However, the Swedish populations started to flower earlier than Norwegians populations. Norwegian materials belong to the late type of red clover when following the Swedish classification into early, middle late and late types of plant development (Jetne, 1973). The only exception is 'Lavine' that develops relatively early and originates from a wild-type of diploid red clover. Wild red clovers in Norway and Sweden generally flower very early. Three of the Swedish populations ('SW Nancy', 'Vicky' and 'SW RK1111') belong to middle to late types of red clover, while 'Betty' is early due to its diploid ancestor 'Bjursele', which is a wild-type of red clover.

Population × Location Interaction

The P × L interaction for SY was highly significant in both experiments, indicating that different environments are suitable for seed production of different varieties. The variation in SY between populations was highest at Landvik in 2012, which indicates that the choice of variety is more important at locations with favorable climatic conditions. However, in 2013 the variation in SY between populations was highest at Bjørke, which this year had more optimal weather conditions than Landvik. It may have been the ice coverage in spring 2013 at Landvik that delayed not only the start of the growing period but also the start of flowering, and most probably reduced the seed yield compared to Bjørke.

Seed companies prefer to produce one or a few high yielding and stable varieties of crops, therefore the phenotypic stability of seed yield across the intended area of cultivation of the varieties is important. The regression analysis of P × L for seed yield showed that the highest yielding varieties, i.e., 'Lasang', 'SW Nancy' and 'Lars', behaved very similarly; they had above-average yield responses to the environment (although not significantly different from the others), and their deviation from regression is not significantly different from zero. These three varieties have all been bred by crossing existing tetraploids, and their good stability across the environments used in this study, in addition to their high yield capacity, lends further support to the general finding in this study that tetraploid varieties made in this way are most useful.

Selection for Higher Seed Yield

SYFH was the seed yield component with highest correlation with total seed yield. This information is important for the breeders working to improve SY in red clover. Previously, several studies found that the number of flower heads per plant (Taylor and Quesenberry, 1996; Montardo et al., 2003; Hermann et al., 2006; Malengier and Baert, 2007) and seed number per plant (Hermann et al., 2006) had the largest effect on SY per plant. However, these earlier studies were based on single plants, with the exception of Taylor and Quesenberry (1996). SYFH was only investigated by Hermann et al. (2006), who did not find a strong correlation between this component and SY per plant. Our results indicated significant correlation between SY and the number of flower heads per square meter; however, we found this component less suitable in practical breeding compared to SYFH. There is also a possible discrepancy in the registrations of the NoMax at different experimental sites. In our results, Landvik had

the highest NoMax, which is probably due to subjective assessment of the maximum flowering date and the fact that both the open and closed flower heads were counted at Landvik while only the open flowers were counted at the other three locations. We should have used Growing Degree Days to determine the date for counting the NoMax at all locations more objectively. Growing Degree Days would also be helpful in defining the optimal harvesting time.

Aamlid and Marum (unpublished data, 1998) performed two cycles of selection for higher SYFH on the 'L-4374' population (spaced plants). Results from Exp. 2 in the present study indicated a tendency to improved SY and SYFH from the selection within the 'L-4374' population to generate 'Lasang', and the selection within 'Lasang' to generate 'LøRk0733', but these differences were not significant (Table 7). The best performing Swedish variety, regarding SY, was 'SW Nancy' that has been made by crosses between tetraploids selected for high SY.

Future Perspectives

We believe that red clover researchers should continue to pursue tetraploid red clover varieties. Tetraploid varieties of red clover in this study produced acceptable SY in experimental trials but in a lesser degree in commercial seed production fields. More stable SY across locations and years is necessary to make tetraploid varieties more popular for seed companies. For Norwegian farmers, an increased profit from higher quality forage yields and a reduced need for nitrogen fertilization would exceed the additional costs of higher seed price of tetraploid red clover (Marum, 2006). Observed SY variation among tetraploid red clover germplasm supports the notion of improving SY through selection. Locations with optimal and stable climate, improved harvesting equipment, updated knowledge on harvesting conditions and cleaning the seed could contribute to increased and more stable SY.

To validate the results from this study we need additional SY trials with material with familiar background. We would also like to investigate how the genetic diversity of tetraploids is changing over time with respect to the chemical development of tetraploids vs. crossing of tetraploids. Plant breeders perform selection for higher seed yield in both single plant and field trials. Our next step will be to compare results from the present study with the results from single plant trials on the same tetraploid populations to identify the best way to select for higher SY.

CONCLUSIONS

Tetraploid populations, in this study made by crossings of existing tetraploids, gave higher SY compared to populations made by doubling the chromosome number. SYFH was the seed yield component with highest correlation with the total SY. This research has shown that there is a potential for higher SY of tetraploid red clover varieties in Norway and that Swedish populations were earlier than the Norwegian populations, but without significant differences in SY.

Acknowledgments

The authors would like to thank Graminor Breeding Ltd. and the Norwegian Research Council (the Industrial Doctorate Program) for funding to support this research. We are also grateful for the insightful discussion and input from Dr. Magne Gullord and for all the help from field technicians at Graminor during the study.

References

Boelt, B., and R. Gislum. 2011. Seed yield potential in Denmark. Paper presented at: Herbage seed production: Findings from research plots to commercial seed multiplication. NJF seminar 420, Ilmajoki, Finland. 28–29 June.

Boelt, B., B. Julier, Đ. Karagic, and J. Hampton. 2015. Legume seed production meeting market requirements and economic impacts. Crit. Rev. Plant Sci. 34(1–3):412–427. doi:10.1080/07352689.2014.898477

Boller, B., F.X. Schubiger, and R. Kölliker. 2010. Red Clover. In: B. Boller, editor, Handbook of Plant Breeding. Vol. 5. Springer, Dordrecht, Netherlands. p. 439–455.

Bragdø-Aas, M. 1970. Fertility of tetraploid red clover. Meld. Nor. Landbrukshøgsk. 49(4):1–34.

Büyükkartal, H.N. 2003. In Vitro Pollen Germination and Pollen Tube Characteristics in Tetraploid Red Clover (*Trifolium pratense* L.). Turk. J. Bot. 27(1):57–61.

Eberhart, S.T., and W.A. Russell. 1966. Stability parameters for comparing varieties. Crop Sci. 6(1):36–40. doi:10.2135/cropsci1966.0011183X000600010011x

Eskilsson, L. 1963. A method for estimating pollen quality in autopolyploid plants. Hereditas 49(1-2):185–188. doi:10.1111/j.1601-5223.1963.tb01874.x

Hermann, D., B. Boller, B. Studer, F. Widmer, and R. Kölliker. 2006. QTL analysis of seed yield components in red clover (*Trifolium pratense* L.). Theor. Appl. Genet. 112(3):536–545. doi:10.1007/s00122-005-0158-1

Hole, J. and E. Solbakken. 1986. Soil mapping at Landvik in Grimstad municipality. (In Norwegian). Norwegian Institute of Land Inventory (Jordregisterinstiuttet rapport nr 11/1986).

Jauhar, P. 1970. Chromosome behaviour and fertility of the raw and evolved synthetic tetraploids of pearl millet, Pennisetum typhoides stapf et hubb. Genetica (The Hague) 41(1):407–424.

Jetne, M. 1973. Grasboka. Landbruksforlaget, Oslo. Isbn: 8252900550. url: http://urn.nb.no/URN:NBN:no-nb_digibok_2007071801029 (verified 11 Dec. 2015).

Julén, U. 1950. Fertility conditions of tetraploid red clover. Hereditas 36(2):151–160. doi:10.1111/j.1601-5223.1950.tb03369.x

Julén, G. 1954. Aspects on the breeding of tetraploid red clover with special reference to the seed setting problem. European Grassland Conference, Paris, Organisation for European Economic Co-operation. p. 69–72.

Lunnan, T. 2000. Influence of clover mixtures in a meadow on forage quality parameters. (In Norwegian). Verknad av kløverinnblanding i eng på fôrkvalitetsparametrar. Husdyrforsøksmøtet 2000:349–352.

Malengier, M., and J. Baert. 2007. Heritability of seed yielding capacity in tetraploid red clover (*Trifolium pratense* L.). In:

Breeding and seed production for conventional and organic agriculture. Proceedings of the 26th meeting of the EUCARPIA fodder crops and amenity grasses section, Perugia, Italy, 2–7 Sept. 2006. p. 128–129.

Marum, P. 2006. Tetraploid red clover- advantages and disadvantages. In: A.Ø. Kristoffersen, editor, Plantemøtet Østlandet 2006. Bioforsk FOKUS 1:170–171.

Meglic, V., and R.R. Smith. 1992. Self-incompatibility and seed set in colchicine-, nitrous oxide-, and sexually derived tetraploid red clover. Crop Sci. 32(5):1133–1137. doi:10.2135/cropsci1992.0011183X003200050013x

Montardo, D.P., M. Dall'Agnol, A.F. Crusius, and N.R. Paim. 2003. Path analysis for seed production in red clover (Trifolium pratense L.). (In Portuguese, with English abstract). R. Bras. Zootec. 32(5):1076–1082. doi:10.1590/S1516-35982003000500007

Norwegian Ministry of Agriculture and Food. 2015. Regulations for seed certification (In Norwegian) Forskrift om såvare. https://lovdata.no/dokument/SF/forskrift/1999–09–13–1052, C Vedlegg om jordbruksfrø, Table 1B. (accessed 20 Oct. 2015).

Pandey, K.K. 1955. Seed development in diploid, tetraploid and diploid-tetraploid crosses of Trifolium pratense L. Indian J. Genet. Plant Breed. 15:25–35.

Peyraud, J.L., A. Le Gall, and A. Lüscher. 2009. Potential food production from forage legume-based-systems in Europe: An overview. Ir. J. Agric. Food Res. 48:115–135.

Picard, J., and J. Berthaut. 1966. Fertility problems in autotetraploid red clover, an insight into the efficiency of selection at the diploid level. Proceedings of the 10th International Grassland Congress, Helsinki, July 1966. p. 664–666.

R-Forge. 2012. Plantbreeding R software package. http://plantbreeding.r-forge.r-project.org (verified 3 Dec. 2015).

Randby, A.T. 1991. Grass meadow silage with or without red clover. (In Norwegian) Surfôr av graseng med eller uten rødkløver. Rep 6. Hellerud Forsoks-og eliteavlsgard, Skjetten.

Rolston, M.P. 1978. Water impermeable seed dormancy. Bot. Rev. 44(3):365–396. doi:10.1007/BF02957854

SAS Institute. 2014. Base SAS 9.4 Procedures Guide: Statistical Procedures. SAS Institute Inc., Cary, NC.

Sjödin, J. 1981. Selection for seed setting capacity in tetraploids of clover and grasses. Paper presented at: Eucarpia, Fodder Crop Section. Merelbeke, Belgium. 8-10 September. p.163–165.

Sjödin, J., and S. Ellerström. 1986. Autopolyploid forage crops. In: G. Olsson, editor, Svalöf 1886–1986: Research and results in plant breeding. Stockholm, Sweden. p. 102–113.

Taylor, N.L., and K.H. Quesenberry. 1996. Red Clover Science. Vol. 28. Current Plant Science and Biotechnology in Agriculture. Kluwer Academic Publishers, Dordrecht.

Valle, O. 1961. Pollination problems in tetraploid red clover. In: S. Ellerstöm, editor, Proceedings of the Symposium on fertility in tetraploid red clover, Eucarpia, Section "Foddercrops" group, Svalof, Sweden. 26-28 Sept. 1961. p 28–33.

Vestad, R. 1990. Red clover in Norwegian pastures. Past and future. In: O.A. Rognli, editor, Present status and future prospects of Norwegian plant breeding. Basic aspects of plant breeding. Norsk landbruksforskning, Suppl. no. 9. p. 165–172.

Vleugels, T., G. Cnops, and E. van Bockstaele. 2013. Screening for resistance to clover rot (Sclerotinia spp.) among a diverse collection of red clover populations (Trifolium pratense L.). Euphytica 194(3):371–382. doi:10.1007/s10681-013-0949-4

Vleugels, T., I. Roldán-Ruiz, and G. Cnops. 2015. Influence of flower and flowering characteristics on seed yield in diploid and tetraploid red clover. Plant Breed. 134(1):56–61. doi:10.1111/pbr.12224

Wexelsen, H., and R. Vestad. 1954. Observations on pollination and seed setting in diploid and tetraploid red clover. European Grassland Conference, Paris. p. 64–68.

Crop Science
SOCIETY OF AMERICA
5585 Guilford Road, Madison, WI 53711-5801
Tel. 608-273-8080 • headquarters@crops.org
www.crops.org

Crop Science Proofing Instructions

Following are the galley proofs for your *Crop Science* manuscript. Please review the galley proofs carefully; they will not be read by editorial staff. Review of the word-processing file at the query letter stage should not be considered a substitute for galley proof review.

Note that you are being asked to correct errors, not to revise the paper. Extensive alteration may require Editorial Board approval, possibly delaying publication.

Reviewing the proofs. Please follow these guidelines for reviewing the proofs:

- Mark your corrections directly on the galley proofs. Make sure that corrections are noticeable and easy to understand.
- Check all type on the galley proofs, including the footers that appear at the bottom of each page, the title and byline, the abbreviations list, and the author–paper documentation paragraph.
- If your galley proofs have tables, check the table data against that in your original tables.
- If your galley proofs have equations, check them against those in your original manuscript.
- If your galley proofs have figures, check to be sure that all type is legible, including any small-print text. If a figure requires alteration, *you must provide a PDF, .tiff, or .jpg file of the revised figure via email attachment*.

Returning the proofs. If the galley proofs require alterations, you may return the marked galleys using one of the following methods:

- Express mail;
- Fax. If you choose this option, you must also send an email message that includes a description of each alteration;
- Email attachment.

If the galley proofs do not require alterations, you may send an email message indicating that no changes are needed.

(more)

COLOR REPRINT COVERS AVAILABLE! For reprint covers, we are now offering high-quality color copies of the journal cover. Your covers will be taken from the issue in which your manuscript is published. Covers include manuscript title, byline, and page numbers.

A NOTE REGARDING PDF FILES: If you have a subscription to the online journal, you may download a PDF file of the article free of charge. The free file is intended for personal use only, and should not be distributed as a reprint or posted online. If you want to use a PDF file for reprints, please buy it here.

Thank you for your contribution to *Crop Science*.

Sincerely,

Elizabeth Gebhardt, Managing Editor
Crop Science
5585 Guilford Road
Madison, Wisconsin 53711 USA

608-268-4950 (phone), 608-273-2021 (fax), egebhardt@sciencesocieties.org

Revised 3/13

RESEARCH

Seed Yield Components in Single Plants of Diverse Scandinavian Tetraploid Red Clover Populations (*Trifolium pratense* L.)

Helga Amdahl, Trygve S. Aamlid, Petter Marum, Åshild Ergon, Muath Alsheikh,* and Odd Arne Rognli

ABSTRACT

Satisfactory seed yield of red clover (*Trifolium pratense* L.) cultivars is crucial for the availability of seeds on the market. Many breeders and researchers have used seed yield components to measure, compare, and explain differences in seed yield between diploid and tetraploid red clover cultivars and populations; however, the relative importance of each component varies between studies. In 2011 and 2012, single-plant trials with several tetraploid and one diploid red clover cultivar were established at the Norwegian plant breeding station in Bjørke. The goal was to study the impact of different seed-yield components on the seed yield of tetraploid plants. Seed weight per flower head was the seed-yield component that correlated best with the seed yield plant^{-1} (r = 0.91 and r = 0.68 in 2011 and 2012, respectively). Path coefficient analysis has also shown that the seed weight per flower head had the highest direct impact on seed yield plant^{-1} (direct path coefficients were 0.867 and 0.783 in 2011 and 2012, respectively). In comparison, the direct path coefficients for calculated number of flower heads, which was previously highlighted as the most important seed-yield component, were lower and more variable (0.739 and 0.392 in 2011 and 2012, respectively). Since previously seed yield per flower head was also identified as the most important seed-yield component in dense plant canopy, this component might have the potential to select for improved seed yield of new cultivars based on single plants. However, further studies are required to confirm this conclusion.

H. Amdahl, M. Alsheikh, and P. Marum, Graminor Breeding Ltd., Hommelstadvegen 60, NO-2322, Ridabu, Norway; T.S. Aamlid, Norwegian Inst. of Bioeconomy Research, Dep. of Urban Greening and Environmental Engineering, Landvik Research Centre, Reddalsveien 215, NO-4886 Grimstad, Norway; H. Amdahl, M. Alsheikh, Å. Ergon, and O.A. Rognli, Norwegian Univ. of Life Sciences, Dep. of Plant Sciences, P.O. Box 5003, N1432 Aas, Norway. Received 11 May 2016. Accepted 2 Aug. 2016. *Corresponding author (muath.alsheikh@graminor.no). Assigned to Associate Editor Heathcliffe Ridlay.

Abbreviations: CalFH, calculated number of flower heads; Corolla, length of the corolla tube; CountFH, counted number of flower heads plant^{-1}; FERT, Fertility; FFH, number of florets per flower head; FH197, number of flower heads by 19 July; SF, number of seeds floret^{-1}; SFH, seed number per flower head; SYFH, seed yield per flower head; SYP, seed yield plant^{-1}; TSW, thousand-seed weight.

Red clover (*Trifolium pratense* L.) is one of the most widely grown forage legumes in temperate areas, together with alfalfa (*Medicago sativa* L.) and white clover (*Trifolium repens* L.) (Boller et al., 2010). Red clover is cultivated in mixtures with grasses on approximately four million hectares worldwide (Isobe et al., 2014). Red clover in grass mixtures is important because its nitrogen-fixing ability, through symbiosis with *Rhizobium leguminosarum* biovar *trifolii*, reduces the need for nitrogen (N) input in fertilizers (Taylor and Quesenberry, 1996). Additionally, high content of crude protein and polyunsaturated fatty acids in red clover makes this species a popular component of legume-grass mixtures (Taylor and Quesenberry, 1996; Lee et al., 2009). Moreover, the high content of polyphenol oxidase (PPO) in red clover slows down the degradation of proteins in silage and thus reduces N losses from plant tissues (Lee et al., 2004; Sullivan and Hatfield, 2006).

Forage legume cultivars often produce low seed yields or seeds of low viability. Both diploid and tetraploid red clover cultivars

Published in Crop Sci. 56:1–10 (2016).
doi: 10.2135/cropsci2016.05.0321

are on the market, and satisfactory seed yields are crucial for seeds of new cultivars to be available for farmers (Sleper and Poehlman, 2006; Boelt et al., 2015). Tetraploid plants have larger leaves and flower heads, thicker stems, and are generally taller compared with diploid plants. Because of these properties, tetraploid cultivars normally produce higher forage yields than diploid cultivars. On the other hand, tetraploid red clover cultivars usually produce lower seed yields than diploid cultivars (Julén, 1975; Sjödin and Ellerström, 1986; Taylor and Quesenberry, 1996). In Norway, the seed yield of diploid red clover has also decreased from 1995 until 2005 (Aamlid, 2006; Havstad and Aamlid, 2015). The same tendency has also occurred in Denmark (B. Boelt, personal communication, 2016). Sufficient seed-yielding ability is considered as important as herbage-yield capacity in red clover breeding programs (Boelt et al., 2015). However, simultaneous breeding for forage and seed yield can represent a challenge, since these two traits are usually negatively correlated (Clifford and Baird, 1993; Steiner et al., 1997; Vasiljević et al., 2000; Herrmann et al., 2006; Sleper and Poehlman, 2006). Proposed reasons for lower seed set of tetraploid cultivars are: (i) low pollination efficiency due to a long corolla that which makes tetraploids less attractive for pollinators (Julén, 1954; Wexelsen and Vestad, 1954; Valle, 1961; Dennis, 1975); (ii) prefertilization problems, e.g., irregularities in the development of ovules and pollen grains due to the formation of aneuploid gametes (Julén, 1954; Povilaitis and Boyes, 1956; Bragdø-Aas, 1970); and (iii) postfertilization problems, e.g., irregularities during the development of seeds (Julén, 1954; Wexelsen and Vestad, 1954; Pandey, 1955; Bragdø-Aas, 1970; Büyükkartal, 2003; Boelt et al., 2015).

Many breeders and researchers have used seed-yield components to measure, compare, and explain differences in seed yield between diploid and tetraploid red clover populations and cultivars (Julén, 1954; Wexelsen and Vestad, 1954; Dennis, 1975; Vleugels et al., 2015b, 2016; Amdahl et al., 2016). Components that determine seed yield in red clover can be classified as primary components, i.e., the number of flower heads per unit area (sometimes recorded as number of flower heads plant^{-1} or shoot^{-1}), the number of florets per flower head, the number of seeds floret^{-1}, and weight seed^{-1} (often recorded as thousand-seed weight). In addition, there are secondary or composed components, i.e., seed number per flower head and seed yield per flower head. Seed yield is a result of different inter- and intra-plant processes that are taking place during seed crop development, in combination with biotic and abiotic environmental factors. These dynamic processes influence assimilate partitioning, and thus the contribution of each seed-yield component to the total seed yield (Fairey and Hampton, 1997). It is common to measure seed yield components on single plants since they are more practical and cheaper to handle compared with dense canopy. However, genotypes may differ in their response to competition in a dense plant canopy compared with single plants (Fairey and Hampton, 1997). In alfalfa, there was a poor correlation between seed yield plant^{-1} and seed yield in dense canopy (Annicchiarico, 2006). On the other hand, a relatively higher correlation was observed for seed yield between spaced plants and dense swards in white clover (Annicchiarico and Piano, 2000). There is limited information on the efficacy of increasing seed yield by selecting spaced plants using different seed-yield components, compared with selection in dense stands of red clover.

Clifford and Baird (1993) identified the number of inflorescences plant^{-1} as the most important seed-yield component in single plants of white clover. The same component was found to have the highest impact on seed yield in single plants of birdsfoot trefoil (*Lotus corniculatus* L.; Li and Hill, 1989). Montardo et al. (2003) performed a path coefficient analysis, which also showed that the number of flower heads plant^{-1} in diploid red clover had the greatest correlation with the seed yield through its direct effect. Herrmann et al. (2006) and Malengier and Baert (2007) found high correlations between the number of flower heads plant^{-1} and seed yield plant^{-1} in single plants of diploid and tetraploid red clover. Even for a dense canopy, Taylor and Quesenberry (1996) stated that it is the number of flower heads per unit area, rather than the seed number per flower head, that mostly influence the total seed yield. Conversely, Oliva et al. (1994) identified the seed number per flower head to have the highest direct effect on seed yield in a dense stand of diploid red clover. A recent study by Vleugels et al. (2016) highlighted that while both the number of mature flower heads plant^{-1} and the number of seeds per mature flower head are explanatory traits for seed yield plant^{-1}, the seed yield for diploids is most dependent on the number of mature flower heads, and for tetraploids, the number of seeds per mature flower head.

The main objective of the present paper was to study differences in seed-yield components in single plants of low and high seed-yielding tetraploid cultivars to identify the most efficient component to be used in selection for improved seed yield. To achieve this objective, in experiment one, we examined which components contributed the most to seed yield improvement of the Norwegian tetraploid material from the first cultivar 'Tripo' (approved in 1964), to one of the most recent cultivars 'Lasang' (approved 2013). In experiment two, our material was further expanded to include two tetraploid Norwegian breeding lines and three tetraploid Swedish cultivars. Since diploid red clover usually gives higher seed yields than the tetraploids, the diploid cultivar

'Lea,' which currently comprises about 70% of Norwegian seed production of red clover (Aamlid and Havstad, 2016), was also included as an industry standard in both experiments.

This is the first study of seed-yield components in single plants of Scandinavian tetraploid red clover cultivars. Path coefficient analysis (Li, 1975) had previously been used to understand the direct and indirect effects of various seed-yield components in diploid red clover (Oliva et al., 1994; Montardo et al., 2003), but not for tetraploid cultivars. Since we investigated the importance of different components for seed yield ha^{-1} of the same tetraploid cultivars and populations in a previous study (Amdahl et al., 2016), our objective was also to verify those findings on single plants.

MATERIALS AND METHODS

Experimental Site and Plant Material

Two single-plant field experiments were established in 2011 at the Graminor plant breeding station in Bjørke, Norway (61°22′ N, 20°42′ E). Cultivars used in experiment one were 'Tripo' (4×), 'Lasang' (4×), and 'Lea' (2×). The old cultivar 'Tripo' and the new cultivar 'Lasang' were selected to represent Norwegian tetraploids with low and high seed yields, respectively. The most common cultivar in Norway, 'Lea,' was selected as a diploid industry standard. Experiment two also included Norwegian breeding populations 'L-4374' (4×) and 'LøRk0733' (4×) and the Swedish cultivars 'Betty' (4×), 'SWTorun' (4×), and 'SWNancy' (4×), in addition to the cultivars used in experiment one. A description of all cultivars and their background is available in Amdahl et al. (2016).

Experimental Protocol

Single plants for both experiments were raised in 13-cm diameter pots in a greenhouse and transplanted into the field. For experiment one, 60 plants of each cultivar were seeded in August 2010, stored at −3°C during the winter, and transplanted into the field in May for seed production in 2011. For experiment two, 18 plants of each cultivar were seeded in May 2011 and transplanted into the field in September for seed production in 2012. In both experiments, a Partec CyFlow® Ploidy Analyser (Sysmex Europe GmbH, Norderstedt, Germany) was used to confirm the ploidy level of all plants before transplanting into the field.

In both experiments, the single plants of the two tetraploid cultivars were planted in a completely randomized design, while plants of the diploid cultivar 'Lea' were planted approximately 100 m from the nearest tetraploid plants to avoid formation of triploids. Despite the 100-m distance, the soil and surrounding vegetation were considered sufficiently uniform to allow a comparison of the tetraploids and the industrial diploid standard. The distance between individual plants in both experiments was 1.5 m in both directions. To keep the plants upright during flowering, and to facilitate the recording of seed-yield components, a 1 m high wire mesh cylinder, 60 cm in diameter, was placed around each single plant at the stem elongation. The following traits were observed and/or calculated (Table 1): number of early flower heads by 19 July (FH197), total number of flower heads (excluding buds) by early August (CountFH), calculated number of flower heads (CalFH), number of florets per flower head (FFH), length of the corolla tube (Corolla), seed number per flower head (SFH), seed yield per flower head (SYFH), thousand-seed weight (TSW), fertility (FERT), and seed yield $plant^{-1}$ (SYP). Measured traits were limited to nine randomly chosen plants per cultivar/population.

Statistical Analyses

Analyses of variance, including the tetraploids and the standard diploid cultivar were performed using PROC GLM in SAS version 9.4 (SAS Institute, 2014). Cultivars/populations were considered as fixed effects in the one-way ANOVA fixed effect model. Differences between mean values were tested using the Ryan–Einot–Gabriel–Welsch multiple range test. Pearson correlations between all the traits in tetraploid populations and for each experiment were analyzed using the PROC CORR procedure in SAS version 9.4. All correlations were based on mean values per single plant (e.g., SNF component mean of 81 data points: 9 plants $cultivar^{-1}$ and 9 flower heads $plant^{-1}$). A path coefficient analysis using the PROC REG procedure of SAS version 9.4 was performed to study the relative importance of the different seed-yield components. Direct path coefficients (Li, 1975) were estimated using multiple backward regression analysis, with SYP as the dependent variable, and components that showed significant correlations with SYP as regressors. The regression analysis was performed on normalized trait values (mean = 0 and SD = 1). Indirect effects via different paths were calculated as the product ($r\beta$) of the respective phenotypic correlation coefficients (r) and the associated direct path coefficients (β).

RESULTS

Experiment One

The two tetraploid cultivars differed significantly from each other ($P \leq 0.001$) regarding the start of flowering (Table 2). Single plants of 'Lasang' had, on average, 188 open flower heads by 19 July, while 'Tripo' had only 91. The diploid cultivar 'Lea' started to flower first and had 20% more FH197 compared with 'Lasang' and 60% more than 'Tripo.'

The CountFH was significantly higher by early August ($P \leq 0.01$) in 'Lasang' than in 'Tripo.' Half of all the CountFH in 'Lea' and 'Lasang' emerged after 27 July, and even later in 'Tripo' (Fig. 1). Despite the differences in CountFH, however, there was no significant difference ($P = 0.237$) between tetraploid cultivars in CalFH. No significant difference ($P = 0.243$) in FFH were detected, and 'Tripo' and 'Lasang' did not differ significantly in either Corolla or SFH. 'Lasang' produced 29% on average, and 'Tripo' 46%, less SFH compared with the diploid industrial standard, 'Lea.' This was also reflected in significantly higher FERT in 'Lasang' than in 'Tripo.' 'Tripo' and 'Lasang' had 50 and 28%lower FERT, respectively, compared with the industrial diploid standard. 'Lasang' had significantly higher ($P \leq 0.05$) SYFH (204 mg) than 'Tripo'.

Table 1. Description of traits measured and calculated in the two experiments with single plants of diploid and tetraploid red clover

Trait	Description	Experiment
Number of flower heads as of 19 July (FH197)	Newly opened flower heads with visible red florets were labeled and counted every third day from the start of flowering until 19 July to estimate earliness.	1, 2
Counted number of flower heads plant^{-1} (CountFH)	In 2011, the counting and labeling of newly opened flower heads every third day continued until 3 August on nine random plants per population.	1
Calculated number of flower heads (CalFH)	The number of flower heads on the nine chosen plants population^{-1} was calculated by dividing the seed yield plant^{-1} with the average seed yield per flower head.	1, 2
Number of florets per flower head (FFH)	Every third day during the flowering period, one flower head from each of the nine plants per population was cut and the number of florets was counted, regardless if they were open or not.	1, 2
Length of the corolla tube (Corolla)	After counting the number of florets per flower head, the corolla tube length of 25 randomly chosen florets was measured (Fig. 2).	1, 2
Seed number per flower head (SFH)	At maturity in September, one flower head from each labelling date was harvested separately from each of the nine plants population^{-1}. The seeds were threshed by hand, cleaned, and counted.	
Seed yield per flower head (SYFH)	Weight of the seeds that had been counted to determine SFH.	1, 2
Fertility (FERT)	Maximum fertility is assumed as two, since the maximum number of ovule per floret is two. The current fertility was then calculated by dividing SFH by FFH and expressed in percent.	1, 2
Seed yield plant^{-1} (SYP)	Each plant was cut on 1 October, dried, threshed, cleaned, and the seed yield determined.	1, 2

Table 2. Mean values and coefficients of variation (%, in parenthesis) of observed traits in two tetraploid cultivars, 'Tripo' and 'Lasang,' and one diploid cultivar, 'Lea,' in experiment one. Mean values are based on single plants.

Variety	Comparison of all cultivars									
	Traits									
	FH197†	CountFH	CalFH	FFH	Corolla	SFH	SYFH	TSW	FERT	SYP
					mm		mg	g	%	g
Tripo	91b‡ (29.6)	598b (27.0)	472a (26.9)	150a (8.6)	11.6a (7.3)	48b (16.6)	142b (23.2)	2.91b (13.4)	16c (17.7)	69b (37.6)
Lasang	188a (30.8)	741a (13.1)	536a (14.7)	138a (16.7)	11.2a (6.0)	63b (33.3)	204a (33.8)	3.25a (11.4)	23b (29.8)	107a (31.8)
Lea (2×)	236a (31.8)	822a (16.8)	597a (29.6)	140a (17.1)	10.2b (3.6)	89a (27)	179ab (28.5)	2.04c (10.8)	32a (22.3)	108a (37.0)
P-value	***	**	n.s.	n.s.	***	*	n.s.	***	***	*

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

† FH197, number of flower heads by 19 July; CountFH, counted number of flower heads; CalFH, calculated number of flower heads; FFH, number of florets per flower head; Corolla, length of corolla tube; SFH, number of seeds per flower head; SYFH, seed yield per flower head; TSW, thousand-seed weight; Fert, fertility; SYP, seed yield plant^{-1}.

‡ Within column, means followed by different letter are significantly different at $P \leq 0.05$; NS, not significant.

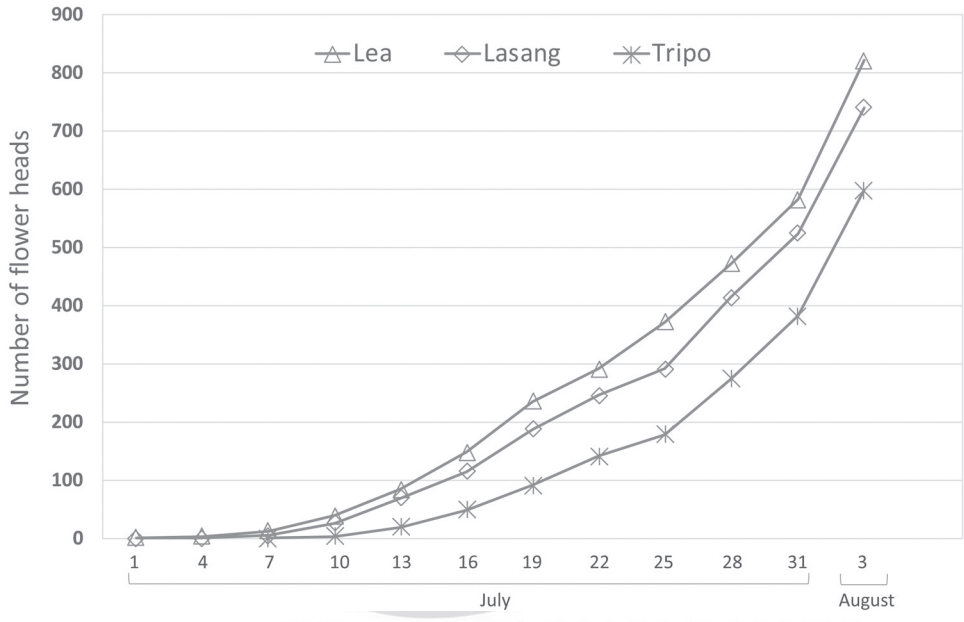


Fig. 1. Cumulative curves for number of flower heads during the flowering season of two tetraploid red clover cultivars, 'Lasang' and 'Tripo,' and one diploid cultivar, 'Lea.' The lines down to the x-axis represent the dates when 50% of flower heads were opened, while the lines to the y-axis represents 50% of the total number of flower heads registered during the flowering period.

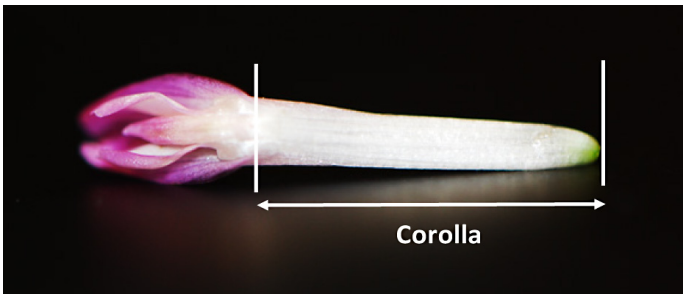


Fig. 2. A red clover floret with indication of length of the corolla tube (Corolla), as it was measured in the present study.

Correlation and Path Coefficient Analyses

Phenotypic correlation analysis on tetraploid plants showed that SYFH ($r = 0.91$, $P \leq 0.001$), SFH ($r = 0.85$, $P \leq 0.001$), and FERT ($r = 0.83$, $P \leq 0.001$) had the strongest correlation with SYP (Table 3). Counted number of flower heads plant^{-1} ($r = 0.64$, $P \leq 0.01$) and CalFH ($r = 0.49$, $P \leq 0.05$) were also correlated with the SYP. Among the measured traits that were significantly correlated with the SYP, path coefficient analysis found only SYFH ($P \leq 0.001$) and CalFH ($P \leq 0.001$) to have a direct effect on SYP. The direct effect of SYFH on the SYP was 0.867, while the direct effect of CalFH was 0.392 (Table 4). Together, these two traits explained 98% of the total phenotypic variation for SYP.

Experiment Two

The northern Swedish cultivar 'Betty' and the southern Swedish cultivar 'SW Nancy' had the highest number of heads in bloom by 19 July 2012 (Table 5). However, 'Betty' was only significantly earlier ($P \leq 0.01$) than 'Tripo,' 'L-4374,' and 'LøRk0733.' Tetraploids had, on average, 0.8 mm longer corollas, 39% fewer flower heads, 23% higher SYFH, and 22% lower SYP than the diploid reference, but the differences among tetraploid cultivars were not significant for any of these characteristics.

The population 'L-4374,' consisting of 170 single-plant families, was used by breeders to perform two cycles of selection for high SFH. This resulted in the cultivar 'Lasang.' In the present study, 'Lasang' had higher SFH, higher SYFH, higher FERT, and higher SYP compared with 'L-4374,' without significant differences. 'LøRk0733' is a result of one natural selection cycle of 'Lasang.' This population had slightly higher SFH, SYFH, FERT, and SYP compared with 'Lasang,' Even though these differences were not significant, SYFH improved 38% and SYP improved 25% from 'L-4374' to 'LøRk0733,' indicating an improvement in SYFH and SYP.

Table 4. Path coefficients showing direct and indirect effects of seed yield per flower head (SYFH) and calculated number of flower heads (CalFH) on seed yield plant^{-1} (SYP) in two tetraploid red clover cultivars, 'Tripo' and 'Lasang,' from experiment one.

Trait	Direct effect†	Indirect effect via SYFH	Indirect effect via CalFH	Total correlation with SYP
SYFH	0.867***	–	0.042	0.909
CalFH	0.392***	0.094	–	0.486

*** Significant at the 0.001 probability level.

† Residual effect = 0.019.

Correlation and Path Coefficient Analysis

When analyzing the tetraploid plants, SYFH, SFH, CalFH, and FERT were highly correlated ($P \leq 0.001$) with SYP (Table 6). Seed yield per flower head showed the strongest correlation ($r = 0.68$, $P \leq 0.001$), while SFH, CalFH, and FERT had correlations of 0.58 ($P \leq 0.001$), 0.57 ($P \leq 0.001$), and 0.53 ($P \leq 0.001$), respectively. The path coefficient analysis confirmed that SYFH was the most important seed-yield component, having a direct effect on SYP of 0.783 ($P \leq 0.001$). The other two traits with significant effects on SYP were CalFH (0.739, $P \leq 0.001$) and Corolla (0.095, $P \leq 0.01$) (Table 7). Together, these three components explained 95% of the phenotypic variation for SYP. The indirect effects of CalFH via SYFH and Corolla showed negative estimates, due to the negative correlations between CalFH and these two traits.

Table 3. Phenotypic correlation coefficients between mean values for each of the measured traits in two tetraploid red clover cultivars: 'Tripo' and 'Lasang.' Mean values are based on single plants in experiment one.

Trait	FH197†	CountFH	CalFH	FFH	Corolla	SFH	SYFH	TSW	FERT
CountFH	0.288 n.s.								
CalFH	0.001 n.s	0.612**							
FFH	−0.327 n.s	−0.213 n.s	−0.554*						
Corolla	−0.391 n.s	−0.212 n.s	−0.201 n.s	0.167 n.s					
SFH	0.304 n.s	0.348 n.s	0.048 n.s	0.238 n.s	0.056 n.s				
SYFH	0.360 n.s	0.448 n.s	0.110 n.s	0.177 n.s	0.060 n.s	0.952***			
TSW	0.373 n.s	0.499*	0.279 n.s	−0.174 n.s	−0.203 n.s	0.081 n.s	0.370 n.s		
FERT	0.444 n.s	0.408 n.s	0.212 n.s	−0.176 n.s	0.038 n.s	0.903***	0.869***	0.103 n.s	
SYP	0.271 n.s	0.640**	0.487*	−0.028 n.s	0.024 n.s	0.845***	0.910***	0.411 n.s	0.827***

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

† FH197, number of flower heads by 19 July; CountFH, counted number of flower heads; CalFH, calculated number of flower heads; FFH, number of florets per flower head; Corolla, length of corolla tube; SFH, number of seeds per flower head; SYFH, seed yield per flower head; TSW, thousand-seed weight; Fert, fertility; n.s., not significant; SYP, seed yield plant^{-1}.

Table 5. Mean values and coefficients of variation (%, in parenthesis) for measured and calculated traits in seven tetraploid cultivars/populations and one diploid red clover cultivar in experiment two. Mean values are based on single plants from experiment two.

	Comparison of all populations								
	Trait								
Population	FH197†	CalFH	FFH	Corolla	SFH	SYFH	TSW	FERT	SYP
				mm		mg	g	%	g
Betty	137a‡ (46.7)	420b (28.8)	153a (10.5)	9.6ab (3.1)	62a (25.8)	187a (28.3)	3.02a (33.8)	20ab (22.3)	78a (37.7)
SW Nancy	127ab (45.6)	427b (31.8)	162a (14.0)	9.4ab (5.7)	58a (29.5)	191a (26.7)	3.30a (9.4)	18ab (28.9)	80a (39.9)
SW Torun	77ab (45.4)	442b (26.7)	154a (17.5)	10.0a (6.1)	59a (18.3)	196a (17.9)	3.33a (9.4)	19ab (21.5)	87a (34.7)
Tripo	64b (40.2)	450b (26.0)	169a (17.9)	9.7ab (6.5)	66a (40.9)	216a (43.7)	3.24a (9.6)	19ab (36.5)	90a (35.0)
L-4374	66b (40.1)	465b (27.3)	150a (11.5)	10.0a (8.2)	51a (34.4)	172a (34.3)	3.40a (7.7)	17b (24.9)	80a (36.0)
Lasang	83ab (49.0)	426b (20.4)	163a (11.5)	9.9a (7.3)	70a (18.4)	227a (24.3)	3.26a (9.5)	22ab (18.8)	97a (33.9)
LøRk0733	71b (76)	423b (37.6)	158a (12.4)	10.1a (8.6)	72a (28.6)	239a (32.3)	3.33a (13.2)	23ab (26.6)	100a (47.1)
Lea (2×)	104ab (41.2)	713a (29.9)	156a(16.8)	9.0b (3.7)	77a (30.7)	166a (33.9)	2.15b (9.1)	24a (19.6)	113a (30.5)
P-value	**	***	n.s.	**	n.s.	n.s.	***	*	n.s.

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

† FH197, number of flower heads by 19 July; CalFH, calculated number of flower heads; FFH, number of florets per flower head; Corolla, length of corolla tube; SFH, number of seeds per flower head; SYFH, seed yield per flower head; TSW, thousand-seed weight; Fert, fertility; SYP, seed yield plant^{-1}.

‡ Within column, means followed by the different letter are significantly different at $P \leq 0.05$; n.s., not significant.

Table 6. Phenotypic correlation coefficients between mean values for each of the measured trait in two tetraploid red clover cultivars: 'Tripo' and 'Lasang.' Mean values are based on single plants from experiment two.

Trait	FH197†	CalFH	FFH	Corolla	SFH	SYFH	TSW	FERT
CalFH	0.043 n.s							
FFH	–0.078 n.s	–0.091 n.s						
Corolla	–0.296*	–0.355**	0.214 n.s					
SFH	0.030 n.s	–0.246 n.s	0.427***	0.209 n.s				
SYFH	–0.075 n.s	–0.174 n.s	0.446***	0.217 n.s	0.946***			
TSW	–0.010 n.s	–0.134 n.s	–0.133 n.s	–0.105 n.s	–0.184 n.s	–0.169NS		
FERT	0.015 n.s	–0.197 n.s	–0.015 n.s	0.147 n.s	0.887***	0.815***	–0.156 n.s	
SYP	–0.025 n.s	0.569***	0.269*	0.003 n.s	0.580***	0.676***	–0.236NS	0.529***

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

† FH197, number of flower heads by 19 July; CountFH, counted number of flower heads; CalFH, calculated number of flower heads; FFH, number of florets per flower head; Corolla, length of corolla tube; SFH, number of seeds per flower head; SYFH, seed yield per flower head; TSW, thousand-seed weight; Fert, fertility; n.s., not significant; SYP, seed yield plant^{-1}.

DISCUSSION

Comparison of Seed-Yield Components in Several Tetraploid Red Clover Cultivars

In experiment one, FH197, CountFH, SYFH, TSW, and FERT were all significantly higher in 'Lasang' compared with 'Tripo.' In experiment two, SFH, SYFH and FERT showed a tendency of being higher in cultivars that gave the highest SYP; however, these differences were not significant. One explanation for this discrepancy could be the weather, as the first part of growing season was warmer and had more rainfall in 2011 than in 2012 (the mean temperature and rainfall for May plus June were 12.2°C and 159 mm in 2011 versus 11.1°C and 93 mm in 2012). This was also reflected by more than twice as many flower heads per plant at the first count on 19 July in 2011 than in 2012. As evidenced in our dense plant canopy trials (Amdahl et al., 2016), differences in seed yield between tetraploid cultivars are likely to be more prominent under optimal weather conditions.

Table 7. Path coefficients showing direct and indirect effects of calculated number of flower heads (CalFH), seed yield per flower head (SYFH), and length of the corolla tube (Corolla) on the seed yield plant^{-1} (SYP) in seven tetraploid red clover populations from experiment two.

		Indirect effect via			
Trait	Direct effect†	CalFH	SYFH	Corolla	Total correlation between trait and SYP
CalFH	0.739***	–	–0.136	–0.033	0.570
SYFH	0.783***	–0.128	–	0.020	0.675
Corolla	0.095**	–0.262	0.169	–	0.002

** Significant at the 0.01 probability level; *** significant at the 0.001 probability level.

† Residual effect = 0.050.

'Lasang' started to flower earlier than 'Tripo' in experiment one. Steiner et al. (1997) stated that the most important factor for high seed yield in western Oregon, where the harvest took place at the beginning of September, is the number of flower heads at the end of July. Our results showed that, at the end of July, only half of the total number of flower heads were open. Flower heads that begin to flower during August have less of a chance to

mature before harvesting, which takes place in September in Norway. Thus, earlier flowering would be an expected positive trait for seed yield. However, as long as flowering started in July and the differences were not bigger than presented in Tables 2 and 5, this trait did not have a significant effect on the seed yield. This finding corresponds well with the finding by Dennis (1975). Therefore, we consider traits other than start of flowering to have a stronger impact on seed yield.

The most important seed yield component in tetraploids

Low seed yield in tetraploid red clover has been a challenge since the first tetraploid cultivars were developed. For that reason, identifying the most efficient seed-yield component(s) to use in selection for improved seed yield is necessary. Seed yield per flower head (SYFH) was the seed-yield component with the strongest correlation and with the strongest direct effect on SYP in the present tetraploid material. Seed yield per flower head is a secondary seed yield component consisting of FFH, number of seeds per floret (FERT), and TSW. Our results did not indicate that FFH is a limiting factor for seed yield in tetraploids, since there was no difference between different tetraploid cultivars for this trait. On the other hand, FERT was significantly lower in the low seed-yielding 'Tripo' compared with the higher seed-yielding 'Lasang' in the first year (experiment one). In the study by Oliva et al. (1994), SF (number of seeds floret^{-1}) was identified as the most important component directly influencing SYP. Keeping in mind that red clover has a maximum of two ovules per floret^{-1}, it is obvious that the seed-yield potential is not fully utilized. Reasons for this could be insufficient pollination, deficient fertilization, and/or incomplete seed development. Different utilization degree of the seed-yield potential seems to distinguish between high and low seed-yielding tetraploid red clover populations/cultivars.

In addition to SYFH, CalFH was also an important seed-yield component in the present study. In most of the earlier studies on the effect of different seed-yield components, number of flower heads plant^{-1} was pointed out as a component most strongly correlated with SYP (Dennis, 1975; Herrmann et al., 2006; Malengier and Baert, 2007; Vleugels et al., 2015b). Taylor and Quesenberry (1996) had the same conclusion in dense plant canopy. In these studies, however, the number of flower heads was not counted but calculated by dividing the SYP with the SYFH. Our results showed that the calculation of the number of flower heads overestimated the actual number of flower heads plant^{-1} in experiment one; this may also have been the case in experiment two. This could mean that not all of the flower heads counted by 3 August 2011 produced mature seeds. Figure 1 shows that about 50% of flower heads opened after 3 August 2011. Possible explanation for overestimation of SYFH followed by underestimation of CalFH could have been that we unintentionally harvested the biggest flower heads. However, flower heads for determination of SYFH were harvested based on the date on the label, and not on the size of the flower head. Flower heads were harvested during the whole flowering period. At harvesting, flower heads were wilted, making it difficult to differentiate between big and small flower heads. Thus, we can exclude nonrandom selection of flower heads in favor of larger flower heads as having affected the results in this study. Contrary to the CountFH, CalFH could not significantly differentiate between the different tetraploid populations in experiment one. Additionally, the number of flower heads in our trials was up to 1200, which is quite a high number to be efficiently used in selection work. Therefore, we do not consider the number of flower heads plant^{-1} as the most optimum indictor for improved seed yield.

Even though the highest seed yield among tetraploids was occasionally obtained by plants that had the longest corolla tube (a corolla tube mean of nine plants belonging to 'LøRk0733' that gave, on average, 100 g seed plant^{-1}, Table 5), Corolla was not correlated with SYP in either of the experiments. However, this component had a slight direct effect on the seed yield, as revealed by path coefficient analysis. One possible reason for this could be that, during backward selection of uncorrelated seed components, the correlation coefficient of the remaining components in the model was slightly changing. We believe that this happened with the corolla tube that gained slight effect on the SYP.

Furthermore, seed number plant^{-1} was also identified as a component that is highly correlated with SYP (Herrmann et al., 2006; Malengier and Baert, 2007). However, both seed number plant^{-1} and head number plant^{-1} were not directly determined but calculated, based on other measured traits. Our experiments indicate that counting and calculating the number of seeds plant^{-1} and flower heads plant^{-1} to quantify, and thus optimize, yield is not a practical and economically viable solution for breeders. Time- and cost-efficient alternatives must be considered, which we discuss in our final conclusion.

Clifford and Scott (1988) observed a general relationship between the number of flower heads and the total seed yield. They proposed that each plant was not able to support more than 70 fertilized ovules inflorescence^{-1}, which corresponds to 70 seeds. In the present study, the highest SFH among tetraploids was recorded in 'Lasang' and 'LøRk0733.' Both cultivars were initially selected for their improved seed yield. The average SFH in 'Lasang' and 'LøRk0733' was 70 and 72, respectively, which is in agreement with the proposal of Clifford and Scott (1988). Seventy seeds per flower head is approximately 23% of the seed-yield potential (if we consider 150 FFH on average with two ovules per floret), and this corresponds well with the observed FERT in our study.

Tetraploids versus Diploid Industry Standard

Overall, the diploid industry standard cultivar, 'Lea,' flowered earlier than both tetraploid cultivars (Fig. 2). However, earlier flowering did not result in higher SYP. This result is in agreement with the results of Dennis (1975), where flowering date did not have any direct impact on seed yield. Number of flower heads plant^{-1} (CountFH and CalFH) and FERT were higher in the diploid 'Lea,' but this did not result in higher SYP compared with the tetraploid cultivars. Tetraploid plants produced higher SYFH compared with the diploid control; however, this did not result in higher SYP. Lower FERT but higher SYFH in tetraploids are presumably a result of high seed weight of tetraploid seeds (Anderson, 1971; Taylor and Quesenberry, 1996). Length of the corolla in diploids and tetraploids has been measured in numerous studies with contradictory conclusions, stressing the effect of weather conditions on the development of flowers (Starling et al., 1950; Bingefors and Eskilsson, 1962; Vleugels et al., 2015a, 2015b). In both of our trials, the length of corolla tube was longer in tetraploids, yet the total seed yield did not differ significantly between tetraploid cultivars and 'Lea.' Only one diploid cultivar is certainly insufficient to perform comparison with tetraploids, and we are therefore unable to draw a final conclusion based on our measurement of corolla tube. However, Vleugels et al. (2016) compared 15 diploid and 15 tetraploid cultivars, and concluded that Corolla is not likely the reason for lower seed yields in tetraploid cultivars.

Potential Seed-Yield Improvement

To extend on the experiments performed in Norway for improving seed yield in either diploid or tetraploid red clover, Aamlid and Marum (unpublished data, 1998) performed two cycles of selection for higher seed number per flower head on an 'L-4374' population. This population consisted of 170 single-plant families that were previously selected, among other traits, for higher seed yield (two cycles of selection). This work resulted in cultivar 'Lasang,' and after one cycle of natural selection in 'Lasang,' the 'LøRk0733' population was developed. Even though, the differences between these three mentioned populations were not significant in experiment two, there was a tendency toward improved SFH, SYFH, FERT, and SYP from the starting 'L-4374' population to the development of 'Lasang.' Though SFH was not significantly different between various tetraploid cultivars in the present experiments, it might be valuable to continue selection for higher SFH to improve SYP. Seed number per ripe flower head was also shown by Vleugels et al. (2016) to be the most important seed-yield component in tetraploid red clover.

Despite the fact that the 'LøRk0733' population does not have a background in 'Tripo,' we found it interesting to compare these two populations, since they represent the first ('Tripo') and, at the time of the trial, the most recent ('LøRk0733') material in the Norwegian tetraploid red clover breeding program. Results from experiment two did not reveal significant differences in any studied trait. Here, however, there was also a tendency toward improved SFH, SYFH, FERT, and SYP in 'LøRk0733,' indicating improvement in seed yield in the advanced Norwegian tetraploid material. This was also confirmed in practical seed production.

Single-Plant versus Dense Plant Canopy

Seed yield per flower head was also identified as the component most strongly correlated (r = 0.96 and 0.98) with seed yield per area in two multisite dense-stand seed production experiments with twelve tetraploid cultivars and breeding populations (Amdahl et al., 2016). Similarly, in the current study, SYFH was also identified to be the most strongly correlated (r = 0.91 and 0.68 in experiments one and two, respectively) with the SYP. Since the same component was identified as the most important for SYP, initial conclusions can be drawn here that SYFH is the component of choice to select for higher seed yield in single-plant to improve seed yield in dense canopy. It is important to note that the number of flower heads plant^{-1} has a larger effect in single-plant trials compared with the number of flower heads per area in dense canopy trials. It is known that single plants are usually larger in size when standing alone and have more flower heads, as compared with being in dense canopy, thus reflected in the final seed yield. Therefore, further studies are needed to confirm that the SYFH in a single plant can be used to improve seed yield of potential cultivars in dense canopy.

CONCLUSIONS AND FUTURE PERSPECTIVES

Among examined seed yield components, FH197, CountFH, SYFH, TSW, and FERT could significantly differentiate between low and high seed-yielding populations, though only in the experiment one. We consider SYFH to be more suitable to use in selection for higher seed yield compared with number of flower heads plant^{-1}. Counting of flower heads per plant, as done in experiment one, is not practical from an economic point of view. It is also difficult to choose the right time, as flowering occurs over a period of approximately 2 mo. Determining the SYFH by collecting 50 flower heads from a plant or population of interest is less time consuming compared with counting or threshing and cleaning several single plants to determine the number of flower heads plant^{-1}. Therefore, we suggest SYFH to be more suitable in selection for higher seed yields. We also believe that selecting for SYP would not be the best estimate for the performance of the same plant in a dense canopy. Seed yield per flower head was also

identified in our previous study to be the most strongly correlated seed-yield component in in dense canopy (Amdahl et al., 2016). However, more research is required to validate the use of the SYFH component in single-plant trials to improve the seed yield in dense canopy. In addition, we do not consider Corolla to be a possible reason for lower seed yields in tetraploid red clover, and this is in agreement the finding by Vleugels et al. (2016).

To further understand seed yield in tetraploid red clover, two tetraploid red clover cultivars ('Tripo' and 'Lasang') were chosen for further genomic analyses. By reciprocal hand-pollination of these two contrasting plants (regarding most of the examined seed-yield components), a mapping population was developed. This population was then planted in a single-plant trial to study their seed-yield components. DNA and RNA from the parent plants and their progeny (the mapping population) were extracted and analyzed to identify potential regions in chromosomes that are most linked to seed yield in tetraploid red clover.

Acknowledgments

The authors would like to thank the Norwegian Research Council (Industrial PhD program) and Graminor Breeding Ltd. for funding this work. We are also grateful for all of the help from the field technicians at Graminor and NIBIO during the course of this study.

References

Aamlid, T.S. 2006. Tiårsjubileum for Norsk frøavlsnytt: Hvor står vi og hvor går vi? (In Norwegian.) Norsk frøavlsnytt 11:4–5.

Aamlid, T.S., and L. Havstad. 2016. Oversikt over norsk frøavl og frøavlsforskning 2014–2015. (In Norwegian.). Jord og Plantekultur 2016 1: 170-174.

Amdahl, H., T.S. Aamlid, Å. Ergon, R.K. Mallikarjuna, P. Marum, M. Alsheikh and O.A. Rognli. 2016. Seed yield of Norwegian and Swedish tetraploid red clover (*Trifolium pratense* L.) populations. Crop Science 56:603–612. doi:10.2135/cropsci2015.07.0441

Anderson, L.B. 1971. A study of some seedling characters and the effects of competition on seedlings in diploid and tetraploid red clover (*Trifolium pratense* L.). N. Z. J. Agric. Res. 14:563–571. doi:10.1080/00288233.1971.10421650

Annicchiarico, P. 2006. Prediction of indirect selection for seed and forage yield of lucerne based on evaluation under spaced planting. Plant Breed. 125:641–643. doi:10.1111/j.1439-0523.2006.01295.x

Annicchiarico, P., and E. Piano. 2000. Response of white clover genotypes to evaluation environments of dense and spaced planting, and implications for selection. Euphytica 111:111–120. doi:10.1023/A:1003898913594

Bingefors, S., and L. Eskilsson. 1962. Pollination problems in tetraploid red clover in central Sweden. Pflanzenz 48:205–214.

Boelt, B., B. Julier, Đ. Karagić, and J. Hampton. 2015. Legume seed production meeting market requirements and economic impacts. Crit. Rev. Plant Sci. 34:412–427. doi:10.1080/07352689.2014.898477

Boller, B., F.X. Schubiger, and R. Kölliker. 2010. Red clover. In: B. Boller, editor, Handbook of plant breeding. Vol. 5. Springer, Dordrecht, the Netherlands. p. 439–455.

Bragdø-Aas, M. 1970. Fertility of tetraploid red clover. Meld. Nor. Landbrukshøgsk. 49:1–34.

Büyükkartal, H.N. 2003. In vitro pollen germination and pollen tube characteristics in tetraploid red clover (*Trifolium pratense* L.). Turk. J. Bot. 27:57–61.

Clifford, P.T.P., and I.J. Baird. 1993. Seed yield potential of white clover: Characteristics, components and compromise. In: Proceedings of the 17th International Grassland Congress, Palmerston, New Zealand, Hamilton, New Zealand, Lincoln, New Zealand, Rockhampton, Australia. 8–21 Aug. New Zealand Grassland Association, Dunedin, New Zealand. p. 1678–1679.

Clifford, P. T. P., and D. Scott. 1988. Inflorescence, bumble bee, and climate interactions in seed crops of a tetraploid red clover (*Trifolium pratense* L.). J. Appl. Seed Prod. 7:38–45.

Dennis, B.A. 1975. Seed set and related problems in autotetraploid *Trifolium pratense* L. Ph.D. diss., University of Reading, Reading, UK.

Fairey, D.T., and J.G. Hampton. 1997. Temperate forage seeds: An introduction. Forage Seed Prod. 1:1–7.

Havstad, L.T., and T.S. Aamlid. 2015. Oversikt over norsk frøavl og frøavlsforsking 2013–2014. (In Norwegian.) Bioforsk Fokus 10:176–180.

Herrmann, D., B. Boller, B. Studer, F. Widmer, and R. Kölliker. 2006. QTL analysis of seed yield components in red clover (*Trifolium pratense* L.). Theor. Appl. Genet. 112:536–545. doi:10.1007/s00122-005-0158-1

Isobe, S., R. Kölliker, B. Boller, and H. Riday. 2014. Red clover In: H. Cai, T. Yamada, and C. Kole, editors, Genetics, genomics and breeding of forage crops. CRC Press, Boca Raton, FL. p. 220–249.

Julén, G. 1954. Aspects on the breeding of tetraploid red clover with special reference to the seed setting problem. In: European Productivity Agency, European Grassland Conference, Paris. 21–24 June. European Productivity Agency, Organisation for European Economic Co-operation, Paris. p. 69–72.

Julén, G. 1975. The current situation in the tetraploid clover. In: B. Nüesch, editor, Ploidy in fodder plants. Eucarpia, Zürich. p. 79–89.

Lee, M.R., P.R. Evans, G.R. Nute, R.I. Richardson, and N.D. Scollan. 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 Sci. 81:738–744. doi:10.1016/j.meatsci.2008.11.016

Lee, M.R., A.L. Winters, N.D. Scollan, R.J. Dewhurst, M.K. Theodorou, and F.R. Minchin. 2004. Plant-mediated lipolysis and proteolysis in red clover with different polyphenol oxidase activities. J. Sci. Food Agric. 84:1639–1645. doi:10.1002/jsfa.1854

Li, C.C. 1975. Path analysis: A primer. Boxwood Press, Pacific Grove, CA.

Li, Q. and P. M. J. Hill. 1989. A study on post-peak flowering shoot manipulation for seed production of *Lotus corniculatus* L. Seed Prod. p. 71–75.

Malengier, M., and J. Baert. 2007. Heritability of seed yielding capacity in tetraploid red clover (*Trifolium pratense* L.). In: F. Veronesi and D. Rosellini, editors, Breeding and seed production for conventional and organic agriculture. Proceedings of the 26th meeting of the EUCARPIA fodder crops and amenity grasses section, Perugia, Italy, 2–7 September 2006. p. 128–129.

Montardo, D.P., M. Dall'agnol, A.F. Crusius, and N.R. Paim. 2003. Path analysis for seed production in red clover (*Trifolium pratense* L.). (In Portuguese, with English abstract.) R. Bras. Zootec. 32:1076–1082. doi:10.1590/S1516–35982003000500007

Oliva, R.N., J.J. Steiner, and W.C. Young. 1994. Red clover seed production: II. Plant water status on yield and yield components. Crop Sci. 34:184–192. doi:10.2135/cropsci1994.0011183X003400010033x

Pandey, K.K. 1955. Seed development in diploid, tetraploid and diploid-tetraploid crosses of Trifolium pratense L. Indian J. Genet. Plant Breed. 15:25–35.

Povilaitis, B., and J.W. Boyes. 1956. A cytological study of autotetraploid red clover. Am. J. Bot. 43:169–174. doi:10.2307/2438673

SAS Institute. 2014. Base SAS 9.4 procedures guide: Statistical procedures. SAS Institute Inc., Cary, NC.

Sjödin, J., and S. Ellerström. 1986. Autopolyploid forage crops. In: G. Olsson, editor, Svalöf 1886–1986: Research and results in plant breeding. Svalöv, Stockholm, Sweden. p. 102–113.

Sleper, D.A., and J.M. Poehlman. 2006. Breeding field crops. 5th ed. Wiley–Blackwell, Hoboken, NJ.

Starling, T.M., C.P. Wilsie, and N.W. Gilbert. 1950. Corolla tube length studies in red clover. Agron. J. 42:1–8. doi:10.2134/agronj1950.00021962004200010001x

Steiner, J.J., R.R. Smith, and S.C. Aldermann. 1997. Red clover seed production: IV. Root rot resistance under forage and seed production systems. Crop Sci. 37:1278–1282. doi:10.2135/cropsci1997.0011183X003700040042x

Sullivan, M.L., and R.D. Hatfield. 2006. Polyphenol oxidase and *o*-diphenols inhibit postharvest proteolysis in red clover and alfalfa. Crop Sci. 46:662–670. doi:10.2135/cropsci2005.06-0132

Taylor, N.L., and K.H. Quesenberry. 1996. Red clover science: Current plant science and biotechnology in agriculture. Vol. 28. Kluwer Academic Publishers, Dordrecht, the Netherlands. doi:10.1007/978-94-015-8692-4

Valle, O. 1961. Pollination problems in tetraploid red clover. In: S. Ellerström, editor, Proceedings of the Symposium on fertility in tetraploid red clover, Eucarpia, Section "Foddercrops" group, Svalof, Sweden. 26–28 Sept. p. 28–33.

Vasiljević, S., G.Š. Momirović, S. Katić, V. Mihailović, D. Lukić, and T. Zivanović. 2000. Mutual relationships among green forage and seed yield components in genotypes of red clover (*Trifolium pratense* L.). Genetika 32:188–191.

Vleugels, T., B. Ceuppens, G. Cnops, P. Lootens, F. R. van Parijs, G. Smagghe, and I. Roldán-Ruiz. 2016. Models with only two predictor variables can accurately predict seed yield in diploid and tetraploid red clover. Euphytica 209: 507–523.

Vleugels, T., I. Roldan-Ruiz, B. Ceuppens, G. Smagghe, and G. Cnops. 2015a. Are corolla tube dimensions the reason for low seed yield in tetraploid red clover? In: Breeding in a world of scarcity. Proceedings of the 31st International EUCARPIA Symposium Section Fodder Crops and Amenity Grasses, Ghent, Belgium. 13–17 Sept. (in press).

Vleugels, T., I. Roldán-Ruiz, and G. Cnops. 2015b. Influence of flower and flowering characteristics on seed yield in diploid and tetraploid red clover. Plant Breed. 134:56–61. doi:10.1111/pbr.12224

Wexelsen, H., and R. Vestad. 1954. Observations on pollination and seed setting in diploid and tetraploid red clover. In: European Productivity Agency, European Grassland Conference, Paris. 21–24 June. European Productivity Agency, Organisation for European Economic Co-operation, Paris. p. 64–68.

***De novo* and reference transcriptome assembly of transcripts expressed during flower development provide insight into the seed setting ability of tetraploid red clover (*Trifolium pratense* L.)**

Mallikarjuna Rao Kovi[1], Helga Amdahl[1,2], Muath Alsheikh[1,2,§] and Odd Arne Rognli[1]

[1]*Department of Plant Sciences, Norwegian University of Life Sciences, NO-1432 Ås, Norway*

[2]*Graminor Breeding AS, Hommelstadvegen 60, NO-2322, Ridabu, Norway*

[§]Corresponding author: muath.alsheikh@graminor.no

Abstract

Red clover (*Trifolium pratense* L.) is one of the most important legume forage species in temperate livestock agriculture. Tetraploid red clover cultivars are generally producing less seed than diploid cultivars. Improving the seed setting potential of tetraploid cultivars is necessary in order to utilize the high forage quality and environmentally sustainable nitrogen fixation of red clover. To date, genomic resources for investigating the seed setting ability of red clover are scarce. In the current study, our aim was to identify candidate genes involved in seed setting in order to develop molecular tools to be used in breeding for improved seed setting. Two genotypes, one from cv.'Tripo' with weak seed setting and one from cv.'Lasang' with strong seed setting, were selected based on data from field experiments for transcriptome analysis of developing flower buds. *De novo* and reference based analyses of transcriptome assemblies were conducted to study the global transcriptome changes from early to late developmental stages of flower development of the two contrasting red clover genotypes. Transcript profiles, gene ontology enrichment and KEGG pathway analysis indicate that genes related to flower development, pollen pistil interactions, photosynthesis and embryo development are differentially expressed between the 'Tripo' and 'Lasang' genotypes. A significant number of genes related to pollination was overrepresented in 'Lasang', which might be a reason for its good seed setting ability. The candidate genes detected in this study might be used to develop molecular tools for breeding tetraploid red clover varieties with improved seed yield potentials.

Key words. *Trifolium pratense* L., seed yield, RNA sequencing, Transcriptomics, Differentially expressed genes

Introduction

Red clover (*Trifolium pratense* L.) is a perennial forage legume species. It is an outcrossing species with a gametophytic self-incompatibility (GSI) system cultivated mostly in the temperate regions. Normally red clover is diploid (2n=2X=14), however, artificially induced tetraploid varieties (2n=4X=28) are also in commercial use. Tetraploid plants were first developed in 1939 by treating germinating seeds, young seedlings or apical meristem of diploids with the chemical colchicine (Sjödin and Ellerström, 1986; Taylor and Quesenberry, 1960; Boller et al., 2010). In addition to colchicine treatment, new tetraploid plants can also arise by treating the plants with nitrous oxide (N_2O) and by gametic non-reduction (Meglic and Smith, 1992; Taylor and Quesenberry, 1996; Boller et al., 2010). However, red clover breeders develop new tetraploid varieties mainly by crossing plants from two or more tetraploid varieties or breeding lines.

Advantages of cultivating tetraploid compared to diploid red clover is its higher forage yield, better persistency and tolerance to some diseases like *Sclerotinia trifoliorum* Eriks. (Sjödin and Ellerström, 1986; Vestad, 1990; Taylor and Quesenberry, 1996; Vleugels et al., 2013). Lower seed yield of tetraploid varieties is the major disadvantage compared to diploid cultivars (Wexelsen and Vestad, 1954; Valle, 1961; Sjödin and Ellerström, 1986).

Seed yield of red clover, especially tetraploids, has not improved in Scandinavia for a long time. The reasons for this are probably complex. A main reason is that several studies indicate that forage and seed yield are negatively correlated which makes seed yield improvement difficult (Clifford and Baird, 1993; Steiner et al., 1997; Vasiljević et al., 2000; Herrmann et al., 2006; Sleper and Pehlman, 2006). Red clover is primarily grown for forage and forage yield is the main breeding goal; however, seed yield is crucial for the commercial value of new varieties (Ravagnani et al., 2012; Annicchiarico et al., 2015). The outcrossing nature and strong self-incompatibility system of red clover prevent the development of inbred lines and hybrids, thus only a proportion of potential heterosis for seed yield can be captured in the usual synthetic varieties (Sleper and Poehlaman, 2006; Annicchiarico et al., 2015).

Genomic resources related to seed yield are scarce in red clover compared to other model legume plants. Currently, four genetic linkage maps for identification of markers linked to important traits have been developed in red clover (Isobe et al., 2003; Sato et al., 2005; Herrmann et al., 2006; Isobe et al., 2009). Several QTL studies of seed yield and seed yield components have been

conducted in species like white clover (*Trifolium repens* L.), soybean (*Glycine max* L.) and perennial ryegrass (*Lolium perenne* L.) (Barrett et al., 2005; Mansur et al., 1996; Cogan et al., 2005). However, so far only one QTL study of seed yield has been performed in red clover, and this study identified 38 QTL (Herrmann et al., 2006).

Rapid advancements in next generation sequencing (NGS) technology allow characterization and quantification of RNA through cDNA sequencing at massive scale (Shendure and Ji, 2008). A draft assembly of the red clover genome based on 16 different genotypes of red clover was recently published (Ištvánek et al., 2014). Furthermore, Yates et al. (2014) performed *de novo* transcriptome studies in red clover and provided insights into the drought response. De Vega et al. (2015) have newly assembled a red clover genome to the chromosome level, estimating its size to be ~309 Mb. They annotated 40,868 genes and identified clusters involved in forage quality and livestock nutrition.

With the availability of new genomic resources in red clover and the advancements in RNA-seq technologies, we performed both *de novo* and reference (red clover genome) based transcriptome analysis of the global transcriptome response during flower and seed development in two red clover genotypes with contrasting seed setting ability. The aim of this study was to identify molecular responses and to elucidate genes determining seed setting ability in red clover.

Materials and methods

Plant material

In 2011, nine single plants of each of the low seed yielding variety 'Tripo', and the high seed yielding variety 'Lasang' were scored for the following seed yield components: number of flower heads per plant, number of florets per flower head, number of seed per flower head, fertility, seed weight per flower head, and length of the corolla tube (Amdahl et al; 2016b). The two lowest ranking plants of 'Tripo' and the two highest ranking plants of 'Lasang' for the majority of registered seed yield components, were selected for further analysis.

RNA sampling

A total number of 12 flower buds, one bud from each of three flower development periods (early – 12th of July, middle – 21st of July and late – 27th of July) from each of the four selected

plants, was picked, flash frozen in liquid N_2 and stored at -80°C until RNA extraction. The frozen flower bud samples were crushed with a pestle and mortar. Using Sigma Spectrum™ Plant Total RNA Kit (Sigma Life Science), total RNA was extracted from the 12 flower buds. On-Column DNAse I Digestion Kit (Sigma Life Science) was used to remove DNA contamination. The quality and concentration of RNA was measured using NANODROP (Nanodrop Technologies, Wilmington, DE, USA) and BIOANALYZER (Agilent Technologies, Palo Alto, CA, USA).

RNA-seq library preparation and Illumina sequencing

Twelve flower bud RNA samples with RIN (RNA Integrity Number) values above seven were used to construct separate cDNA libraries with fragment lengths of 200 bp (±25 bp). Single-end sequencing was performed at the Norwegian Sequencing Centre (NSC), University of Oslo using the Illumina sequencing platform (HISEQ 2000) generating single-end reads with a length of 50 bp. The FastQC program (http://www.bioinformatics.babraham. ac.uk/projects/fastqc/) was used to analyse the quality of the raw sequencing reads.

***De novo* transcriptome analysis**

The *de novo* assembly was performed in a similar manner as described by Kovi et al. (2016). Briefly, adapter sequences and low quality reads were removed using the sickle program (https://github.com/najoshi/sickle/blob/master/README.md). The clean reads derived from the four individual genotypes named Tripo42 and Tripo55, Lasang77 and Lasang108, were used to construct separate *de novo* assemblies for each genotype using the Trinity assembler (release 2013-02-25) (Grabherr et al., 2011). The *de novo* assembled transcriptome was then used as a reference to map the individual reads using the Bowtie program (Langmead et al., 2009). Transcript abundance was measured for each genotype and time point combination as the expected number of fragments per kilobase of transcript sequence per million mapped reads (FPKM) (Trapnell et al., 2010) using RSEM version 1.1.11 (Li and Dewey, 2011).

Identification of differentially expressed genes (DEGs), annotation and gene ontology (GO) analysis

The edgeR program (Robinson et al., 2010) was employed to identify DEGs and a false discovery rate (FDR) of 0.05 was further used to determine the significant DEGs. Transcripts showing

differential expression at any flower development time-point were clustered using a K-means clustering algorithm. The annotation of the DEGs were performed using the Blast2GO program (Conesa and Gotz, 2008). Initially BLASTx was performed with an E-value threshold of 10e-06, followed by annotation with a cut-off value of 55 and GO weight Hsp-hit value of 20. The GO enrichment analysis was performed with a p-value of 0.01. The GO classification of DEGs in the two genotypes were generated using the WEGO program (Ye et al., 2006). KEGG Pathway analysis was performed with the Blast2GO program (Conesa and Gotz, 2008).

Validation of *de novo* assembly by CEGMA

The CEGMA software (version 2.4) (Parra et al., 2007) was used to evaluate the quality of the four transcriptome assembly datasets. Several genome and transcriptome assembly studies have used CEGMA for evaluating the quality of assemblies (Kovi et al., 2016). CEGMA detects the presence of 248 extremely conserved core eukaryotic genes (CEGs) and their coverage in transcriptome assemblies for evaluation of the completeness of the assembly.

Red clover reference based transcriptome analysis, detecting DEGs and functional annotation

Using a reference-based approach, we mapped all the clean reads from two genotypes ('Tripo' and 'Lasang') and three time point (early, middle and late flower development) combination to the red clover reference genome (De Vega et al., 2015) using STAR, an ultrafast universal RNA-seq. aligner program (Dobin et al., 2012). The Cufflinks program (Trapnell et al., 2010) was used to assemble the transcriptomes and to estimate the transcript abundance, followed by the cuffmerge and cuffdiff programs, which is included with the Cufflinks package. The Cuffmerge program merge the transcriptome assemblies from the three flower development time- points of each genotype for performing differential expression analysis. The cuffdiff program compare the expression levels of genes and transcripts between the three time-points for each genotype, and detect genes that are up- or down-regulated between the time-points. The merged GTF files obtained from the cuffmerge program was used in the TransDecoder program (Haas et al., 2013) to find the coding regions within transcripts. The longest homology coding sequences obtained from TransDecoder were blasted against the Viridiplantae database extracted from NCBI to find the gene names for the coding sequences. Further annotation was performed using the SWISS-

PROT database and GFF3 (generic feature format) annotation file describing genomic features, was generated using in-house developed python scripts.

Comparison of significant DEGs to seed yield related QTL

To compare the DEGs with the QTL for seed yield and seed yield traits described by Hermann et al. (2006), we identified flanking SSR markers associated with the QTL and downloaded the marker sequences from the NCBI database. The chromosome locations of markers and DEGs were identified using the BLAST program with the marker sequences and DEG sequences as the query and the red clover genome sequence (De Vega et al., 2015) as the subject. A genetic map was created based on the physical location of the DEGs in the red clover genome. Briefly, all the physical location (bp) of DEGs were converted to centimorgan (cM) by an average of 450 kb/cM in red clover and spanned 440 cM across seven linkage groups (LGs), approximately similar to 444 cM of Hermann et al. (2006). The genetic map was created by the MapDraw software (Liu and Meng, 2003). The detected SSR markers of QTL for seed yield trait (Hermann et al., 2006) are highlighted on the DEGs genetic map.

Results

***De novo* assembly**

The low seed yielding 'Tripo' and the high seed yielding 'Lasang' genotypes (Fig. 1) were sequenced and characterized by the *de novo* transcriptome assembly (Table 1). A total number of 218 million reads of 50bp were generated for the four genotypes (Tripo42, Tripo55, Lasang77 and Lasang108). 112 million reads were from the 'Tripo' genotypes and 106 million reads from the 'Lasang' genotypes. Individual transcriptome assemblies were generated for each genotype. The number of contigs observed in Lasang108 was 80,328 (N50 of 930 bp), in Lasang77 83,489 (N50 of 982 bp), in Tripo42 84,545 (N50 of 1016 bp), and in Tripo55 84,442 (N50 of 982 bp). The longest contig sizes were 7469, 7295, 7447 and 7339 bp for Lasang108, Lasang77, Tripo42 and Tripo55, respectively. CEGMA analysis determined the complete CEGs (Core Eukaryotic Genes) in Lasang108, Lasang77, Tripo42 and Tripo55 transcriptome assemblies to be 89.11, 92.34, 92.34 and 92.34 %, respectively, while the percentage of partially complete CEGs ranged from 97.18 to 97.98 (Table 2). The average number of orthologues per CEG in the four assemblies ranged from

3.18 to 3.30, while the percentage of CEGs that had more than one orthologue ranged from 89.59 to 95.20 (Table 2).

DEGs identified by *de novo* and reference based methods

Clean reads from each sample were mapped onto their respective genotype specific *de novo* assemblies and to the reference genome (red clover genome sequence) to estimate the expression levels of transcripts at different flower development time-points, early (EF), middle (MF) and late (LF) flower development. The DEGs identified in a series of pairwise comparisons between the three flower development time-points EF-LF, EF-MF and LF-MF were 15,000, 7,204 and 7,903, respectively, in Tripo42; 18,105, 6,050 and 10,100, respectively, in Tripo55; 12,040, 8,426 and 2,304, respectively, in Lasang77; and 10,986, 7,492 and 2,430, respectively, in Lasang108 with a false discovery rate (FDR) < 0.05 (Fig. 2B). In the reference-based analysis, 875 and 932 DEGs were observed between EF-LF samples; 279 and 586 between EF-MF samples and 331 and 93 between MF-LF samples in the 'Tripo' and 'Lasang' genotypes, respectively, including up- and down-regulated transcripts (Fig. 2A)

To determine the sample relations, differential expression data from the edgeR program was used to generate heat maps (Fig. 3). EF and MF grouped together in the low seed yielding 'Tripo' genotypes, while MF and LF grouped together in the high seed yielding 'Lasang' genotypes, indicating that unique genes expressed during late flower development (LF) in 'Tripo' and early flower development (EF) in 'Lasang' were playing major roles in their flowering and seed setting abilities.

Blast, annotation and GO of differentially expressed genes

BLASTx was performed for all the DEGs against the Viridiplantae database derived from NCBI. Approximately 80% of the DEGs had blast hits and 60% were annotated using the Blast2GO program (Conesa and Gotz, 2008). The top blast hit species were *Trifolium subterraneum*, followed by *Medicago truncatula*, both species closely related to red clover. Gene ontology (GO) classification of DEGs of 'Tripo' and 'Lasang' were represented as three main GO categories, i.e. cellular component, molecular function and biological process in a histogram (Fig. 4) using the WEGO (Web Gene Ontology Annotation Plot) graphical tool (Ye et al., 2006). GO comparisons between 'Tripo' and 'Lasang' showed some differences regarding the cellular component and molecular function categories, while relatively small differences were observed for the biological

process category. DEGs involved in membrane-enclosed lumen and translation regulator were present only in 'Tripo', while DEGs involved in structural molecule were present only in 'Lasang'.

GO enrichment analysis by Fischer's exact test was conducted to determine over- or under represented GO terms using Blast2GO and the REVIGO tool for reducing and visualising gene ontologies (Supek et al., 2011). Six GO terms were enriched when comparing the 'Tripo' and 'Lasang'. Out of these, four GO terms, i.e. plasma membrane, pollination, transport and Golgi apparatus were overrepresented in the high seed yielding 'Lasang'. Transcripts assigned to DNA metabolic processes and nucleic acid binding were overrepresented in the low seed yielding 'Tripo' (Fig. 5, Supplementary figure 1).

Several genes, putatively involved in flower and seed development were detected in these studies, e.g. walls are thin related protein (WAT1), tubby-like F-box protein, gibberellin (GA) 2–beta-dioxygenase, putative aquaporin NIP4-1, zinc finger protein 4, which all were significantly upregulated from the EF to the MF stage, and significantly downregulated from MF to LF (Table 3, Fig. 2). Ethylene-responsive transcription factor (ERF106), probable inorganic phosphate transporter 1-4 (OsPht1;4) were significantly downregulated from the EF to the MF stage, while they were upregulated from MF to LF (Table 3, Fig. 2). Furthermore, the Kyoto encyclopedia of genes and genomes (KEGG) database detected different pathways between Tripo and Lasang at EF-MF and MF-LF stages. In total 1196 DEGs were involved in 87 pathways (Supplementary table 1). Pathways with highest representation among the genes were involved in starch and sucrose metabolism (4.84 %, 58 genes), pentose and glucoronate interconversions (2.84 %, 34 genes), phenylpropanoid biosynthesis (2.75 %, 33 genes), purine metabolism (2.34 %, 28 genes) and thiamine metabolism (2 %, 24 genes).

DEGs compared to the seed yield QTL

The DEGs identified in this study were compared to seed yield related QTL to see if any of our genes are co-located with the seed yield QTL described by Hermann et al. (2006). Out of 15 SSR markers flanking the seed yield QTL, six SSR markers are located in corresponding regions as six DEGs detected in this study across 4 four linkage groups (Fig. 6). The six DEGs are myb-related protein MYBAS2, 4-coumarate--CoA ligase-like 2, protein cornichon homolog 3, ethylene-responsive transcription factor ERF113, protein DETOXIFICATION 45, and UDP-glucuronate 4-epimerase 4.

Discussion

Comparative analysis between *de novo* and reference based transcriptome assays

In order to understand the molecular responses related to seed yield, we performed RNA sequencing of 'Tripo' and 'Lasang' genotypes with contrasting seed yield ability, and performed *de novo* and red clover reference based mapping. When a reference genome is available, reference-based approaches have been considered more effective than *de novo* assembly (Martin and Wang, 2011) but very few studies have compared the two strategies (Kovi et al., 2016; Ward et al., 2012). In the present transcriptome analysis of red clover, we compared both strategies. The CEGMA analysis showed that the *de novo* assemblies were very complete in terms of gene content since they captured high percentages of ultra-conserved CEGs in all assemblies of the 'Tripo' and 'Lasang' genotypes. *De novo* and reference based (red clover reference genome) gene expression data indicated that genes expressed during the early flower development stage (EF) in 'Lasang' and during the late flower development stage (LF) in 'Tripo' might play key roles in their differential seed setting abilities. In both *de novo* and the reference based mapping, the pattern of the differentially expressed transcripts was similar. A larger number of differentially expressed transcripts was observed in early vs late flower development stage than in middle vs late and middle vs early stage (Fig. 2 and Fig. 3). This might be due to the presence of several differentially expressed transcripts expressed at all three stages. Additionally, there was a larger number of differentially expressed transcripts at the early vs middle flower development stage in 'Lasang' than in 'Tripo', while there was a more differentially expressed transcripts at the early vs late stage in 'Tripo' compared to 'Lasang' (Fig. 2). Further, we found the proportion of differentially expressed transcripts to be very high in *de novo* compared to the reference based mapping, similar to the findings of Kovi et al. (2016). The trinity *de novo* assembler yield more transcripts due to the lack of strand-specific information. However, most of the differentially expressed genes identified in both these methods were related to the *Medicago truncatula* (Supplementary figure 2), which is the most closely related species to red clover. Furthermore, both methods identified many similar candidate genes putatively involved in flower and seed development (Table 3), thus demonstrating the potential of the *de novo* method of capturing genes even in the absence of a reference genome.

Potential candidate genes involved in flower and seed development

Several genes putatively involved in flower development were detected in these studies (Table 3). WAT1 related protein is a cell wall protein mainly responsible for transmembrane transporter activity (http://www.uniprot.org/uniprot/Q94AP3). Ranocha et al. (2010) reported that stem apices in the mutant *wat1* produced significantly lower seed yields in *Arabidopsis thaliana* compared to wild type stem apices. It might be that the downregulated expression of this gene in 'Tripo' flower buds in the early and middle flower development periods negatively affected its seed setting ability and thus seed yield.

Tubby-like proteins are involved in abscisic acid (ABA) signaling pathways and plays a key role in seed germination and early seed growth (Bao et al., 2014). In a recent study, Verma et al. (2016) identified a tubby-like F-box protein as a potential candidate gene for the seed weight QTL *qSW* in chickpea (*Cicer arietinum* L.). Gibberellin 2–beta-dioxygenase was highly expressed in EF and MF. According to Xue et al. (2012), genes that encodes gibberellin 2-beta-dioxygenase 1 were highly expressed in rice embryo.

NIP4-1 belongs to the aquaporin gene family, which are small integral membrane proteins that facilitate water and solute movement across different tissues throughout development and growth (Khan et al., 2015). Regulation of water and nutrient state is very relevant for pollen development, pollen tube growth and germination (Firon et al., 2012). Recently Di Giorgio et al. (2016) showed that NIP4-1 and NIP4-2 are required for pollen development and pollination in *Arabidopsis thaliana*. Furthermore, single *nip4;1* mutant plants showed a significantly higher frequency of abnormal, stunted siliques and fewer seeds when compared with the wild type (Di Giorgio et al., 2016). This indicate that NIP4-1 plays a prominent role in determining seed yield. In our studies, significant upregulation of this gene in 'Lasang' during the EF and MF stages might play a key role in determining the better seed yielding capacity of this cultivar.

Zinc finger proteins (ZFP) play an important role in various biological functions, such as plant growth and development (flower, shoot, seed, pistil and leaf) (Luo et al., 1999; Kubo et al., 2000). Recently it was found that ZFP3, ZFP4 and the related ZFP subfamily of zinc finger factors regulate light and ABA responses during germination and early seedling development (Joseph et al., 2014). Higher expression of ZFP4 during the EF and MF stages indicate that it might be important for seed setting in our tetraploid red clover genotypes.

The gene ERF106 belongs to the APETALA2 (AP2) gene family, which controls seed weight (Ohto et al., 2009), was overexpressed during the EF and MF stages in 'Lasang' flower buds. APETALA2 influences the development of embryo, endosperm and seed coat (Ohto et al., 2009). According to Xue et al. (2012), genes involved in ethylene mediated signaling were highly expressed in rice developing seeds.

The rice gene OsPht1,4 belongs to a group of genes that regulate phosphorus homeostasis in plant cells (Ying et al., 2015). Jia et al. (2011) reported that suppression of OsPT4 in rice resulted in lower P content in unfilled rice grains, which again resulted in lower seed yields. This gene was overexpressed during the MF and LF stages in 'Lasang' flower buds indicating its positive effect on seed yield. This gene is also involved in the embryo development in rice (Zhang et al., 2015).

GO differences in 'Tripo' and 'Lasang'

Gene ontology (GO) comparisons between the two genotypes showed differences regarding the cellular component and molecular function categories (Fig. 4). Six GO terms were enriched between these two genotypes. Out of these, four GO terms, i.e. plasma membrane, pollination, transport and Golgi apparatus were overrepresented in the high seed yielding 'Lasang'. Transcripts assigned to DNA metabolic processes and nucleic acid binding were overrepresented in the low seed yielding 'Tripo' (Fig. 5; supplementary figure 1). Genes such as pollen-specific leucine-rich repeat extensin-like protein 1-like (PEX1), pollen profiling variant 1, phd finger protein male sterility 1–like (MS1-PHD), and polypyrimidine tract-binding protein were observed in the pollination GO term. PEX1 reported to be involved in reproduction with in the pollen tube wall during its rapid growth (Rubinstein et al., 1995). Another gene, MS1-PHD encodes a PHD-type transcription factor and regulates pollen and tapetum development and pollen wall biosynthesis (Yang et al., 2007). In the GO term plasma membrane, genes like aberrant pollen transmission (APT1), flotillin-like protein, sodium transporter hkt1-like were observed. Xu and Dooner (2006) showed that the APT1 protein is involved in membrane trafficking and is required for the high secretory demands of tip growth in pollen tubes. Most of the overrepresented genes are linked to pollen development, which is crucial for fertility and seed setting, thus likely involved in determining the higher seed yield capacity of 'Lasang' compared with 'Tripo'.

Validation of the DEGs by comparing to previous red clover seed yield QTL

Comparative mapping studies are powerful tools to validate the detected DEGs by comparing the sequences of the genes to the sequences of markers located inside or flanking QTL (Man et al., 2016). The genetic map of DEGs was created based on their physical locations (bp) in the red clover genome. However, there is no constant ratio to convert between bp and cM, as some regions of the genome with frequent recombinations, have fewer bp per cM and the regions with rare recombinations (e.g. centromeres), have more bp per cM. The best approach might be picking the most detailed genetic map (in our case Hermann et al. 2006) and fetching the sequences for each SSR markers in the map, and BLAST these markers against the genome (red clover genome). From the BLAST analysis, we were able to translate each of these cM distances into a bp distance between the points of alignment from the markers to the chromosome, thus calculated as 450 kb/cM. A similar approach was carried out in *Arabidopsis* by estimating genetic distance as 250 kb/cM (Lukowitz et al., 2000). In this study, we detected six DEGs that mapped to the seed yield QTL regions identified by Hermann et al. (2006) on four linkage groups LG1, LG2, LG3 and LG6 respectively. (Fig. 6). Among them, MYB transcription factors plays a key role in plant development, pollen development (Phan et al., 2011), pollen tube differentiation (Leydon et al., 2014), floral initiation and seed development (Woodger et al., 2003). The gene 'protein cornichon homolog' belongs to a conserved protein family found in eukaryotes demonstrated to participate in the selection of integral membrane proteins as cargo for their correct targeting (Rosas-Santiago et al., 2015). Further Man et al. (2016), detected protein cornichon homolog as a potential gene encoding for the yield related QTL in cotton. 4-coumarate--CoA ligase-like 2, belongs to a group of essential enzymes involved in the phenylpropanoid-derived compound (PDC) pathway, which generates various secondary compounds like lignin, anthocyanins, and isoflavonoids. Doughty et al. (2014) suggests that flavonoids may play a fundamental role in regulating communication between the seed coat and the endosperm also.

Conclusions

In this study, transcriptome analysis was conducted for 'Tripo' with inferior seed setting ability and two from cv. 'Lasang' with improved seed setting ability, and several DEGs were identified. Many genes related to pollination, flower and seed development were upregulated during the early to middle (EF-MF) flower development stage in the 'Lasang' and downregulated during

the middle to late (MF-LF) flower development stage in the 'Tripo', indicating their major role in determining seed setting and potential seed yield. GO enrichment analysis further confirmed that plasma membrane, pollination, transport and Golgi apparatus related genes are overrepresented in the 'Lasang'. Further, comparative mapping, co-located, six seed yield related QTL to the six DEGs on the same linkage groups, thus validating the detected DEGs in this study. Putative candidate genes detected in this study might provide a basis for future research in understanding the seed yield biology in red clover.

Accession codes

The raw Illumina sequencing data generated in this study were deposited in the EMBL-EBI ArrayExpress Archive, under accession number E-MTAB-5117.

Author contribution statement

M.A. and O.A.R. designed the study with inputs form H.A and M.R.K. H.A performed phenotype experiments and collected plant material. M.R.K. was responsible for RNA sequencing, bioinformatics and expression analysis. M.R.K. and H.A. drafted the manuscript with inputs from M.A. and O.A.R. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by the Norwegian Research Council (NRC) grant No. 209702 (Industrial PhD program) and by Graminor Breeding AS. We are also grateful to Elena Gusakova for the help with RNA extraction. We sincerely acknowledge the efforts of Torben Asp from Arhus University and Tina Graceline Kirubakaran from CIGENE, NMBU for providing valuable suggestions of bioinformatics analysis.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Amdahl H, Aamlid TS, Marum P, Ergon Å, Alsheikh M, Rognli OA (2016) Seed yield components in single plants of diverse Scandinavian tetraploid populations (*Trifolium pratense* L.). Crop Science, Accepted, in press (First Look: doi: 10.2135/cropsci2016.05.0321)

Annicchiarico P, Barrett B, Brummer EC, Julier B, Marshall AH (2015) Achievements and Challenges in Improving Temperate Perennial Forage Legumes. Crit Rev Plant Sci 34:327-380

Bao Y, Song WM, Jin YL, Jiang CM, Yang Y, Li B, Huang WJ, Liu H, Zhang HX (2014) Characterization of Arabidopsis Tubby-like proteins and redundant function of AtTLP3 and AtTLP9 in plant response to ABA and osmotic stress. Plant Mol Biol 86:471-483

Barrett BA, Baird IJ, Woodfield DR (2005) A QTL Analysis of White Clover Seed Production. Crop Sci 45:1844-1850

Boller B, Schubiger FX, Kölliker R (2010). Red Clover In 'B. Boller et al. (eds.), Fodder Crops and Amenity Grasses, Handbook of Plant Breeding 5, DOI 10.1007/978-1-4419-0760 8_11, Springer Science+Business Media, pp 439-455

Clifford PTP, Baird IJ (1993) Seed yield potential of white clover: characteristics, components and compromise. In Proceedings of the XVII International Grassland Congress, Palmerston North, New Zealand, 8-21 February. p. 1678-1679

Cogan NOI, Smith KF, Yamada T, Francki MG, Vecchies AC, Jones ES, Spangenberg GC, Forster JW (2005) QTL analysis and comparative genomics of herbage quality traits in perennial ryegrass (Lolium perenne L.). Theor Appl Genet 110:364-380

Conesa A, Gotz S (2008) Blast2GO: A comprehensive suite for functional analysis in plant genomics. Int J Plant Genomics 2008:619832

De Vega JJ, Ayling S, Hegarty M, Kudrna D, Goicoechea JL, Ergon Å, Rognli OA, Jones C, Swain M, Geurts R, Lang C, Mayer KFX, Rössner S, Yates S, Webb KJ, Donnison IS, Oldroyd GED, Wing RA, Caccamo M, Powell W, Abberton MT, Skøt L (2015) Red clover (Trifolium pratense L.) draft genome provides a platform for trait improvement. Sci Rep 5:17394

Di Giorgio JA, Bienert GP, Ayub ND, Yaneff A, Barberini ML, Mecchia MA, Amodeo G, Soto GC, Muschietti JP (2016) Pollen-Specific Aquaporins NIP4;1 and NIP4;2 Are Required for Pollen Development and Pollination in Arabidopsis thaliana. Plant Cell 28:1053-1077

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2012) STAR: ultrafast universal RNA-seq aligner. Bioinformatics

Doughty J, Aljabri M, Scott RJ (2014) Flavonoids and the regulation of seed size in Arabidopsis. Biochem Soc Trans 42:364-369

Firon N, Nepi M, Pacini E (2012) Water status and associated processes mark critical stages in pollen development and functioning. Ann Bot 109:1201-1214

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29:644-652

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, Couger MB, Eccles D, Li B, Lieber M, MacManes MD, Ott M, Orvis J, Pochet N, Strozzi F, Weeks N, Westerman R, William T, Dewey CN, Henschel R, LeDuc RD, Friedman N, Regev A (2013) De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity. Nat protoc 8:10.1038/nprot.2013.1084

Herrmann D, Boller B, Studer B, Widmer F, Kölliker R (2006) QTL analysis of seed yield components in red clover (Trifolium pratense L.). Theor Appl Genet 112:536-545

Isobe S, Klimenko I, Ivashuta S, Gau M, Kozlov NN (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

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 9:1-11

Istvanek J, Jaros M, Krenek A, Repkova J (2014) Genome assembly and annotation for red clover (Trifolium pratense; Fabaceae). Am J Bot 101:327-337

Jia H, Ren H, Gu M, Zhao J, Sun S, Zhang X, Chen J, Wu P, Xu G (2011) The phosphate transporter gene OsPht1;8 is involved in phosphate homeostasis in rice. Plant Physiol 156:1164-1175

Joseph MP, Papdi C, Kozma-Bognár L, Nagy I, López-Carbonell M, Rigó G, Koncz C, Szabados L (2014) The Arabidopsis ZINC FINGER PROTEIN3 Interferes with Abscisic Acid and Light Signaling in Seed Germination and Plant Development. Plant Physiol 165:1203-1220

Khan K, Agarwal P, Shanware A, Sane VA (2015) Heterologous Expression of Two Jatropha Aquaporins Imparts Drought and Salt Tolerance and Improves Seed Viability in Transgenic Arabidopsis thaliana. PLoS ONE 10:e0128866

Kovi MR, Abdelhalim M, Kunapareddy A, Ergon Å, Tronsmo AM, Brurberg MB, Hofgaard IS, Asp T, Rognli OA (2016) Global transcriptome changes in perennial ryegrass during early infection by pink snow mould. Sci Rep 6:28702

Kubo K-i, Kanno Y, Nishino T, Takatsuji H (2000) Zinc-Finger Genes That Specifically Express in Pistil Secretory Tissues of Petunia. Plant Cell Physiol 41:377-382

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol10:1-10

Leydon AR, Beale KM, Woroniecka K, Castner E, Chen J, Horgan C, Palanivelu R, Johnson MA (2013) Three MYB transcription factors control pollen tube differentiation required for sperm release. Curr Biol 23:1209-1214

Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323

Liu RH, Meng JL (2003) MapDraw: a microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data. Hereditas 25:317-321

Lu G, Coneva V, Casaretto JA, Ying S, Mahmood K, Liu F, Nambara E, Bi YM, Rothstein SJ (2015) OsPIN5b modulates rice (Oryza sativa) plant architecture and yield by changing auxin homeostasis, transport and distribution. Plant J 83:913-925

Lukowitz W, Gillmor CS, Scheible W-R (2000) Positional Cloning in Arabidopsis. Why It Feels Good to Have a Genome Initiative Working for You. Plant Physiol 123:795-806

Luo M, Bilodeau P, Koltunow A, Dennis ES, Peacock WJ, Chaudhury AM (1999) Genes controlling fertilization-independent seed development in Arabidopsis thaliana. Proc Natl Acad Sci USA 96:296-301

Man W, Zhang L, Li X, Xie X, Pei W, Yu J, Yu S, Zhang J (2016) A comparative transcriptome analysis of two sets of backcross inbred lines differing in lint-yield derived from a Gossypium hirsutum x Gossypium barbadense population. Mol Genet Genomics 291:1749-1767

Mansur LM, Orf JH, Chase K, Jarvik T, Cregan PB, Lark KG (1996) Genetic Mapping of Agronomic Traits Using Recombinant Inbred Lines of Soybean. Crop Sci 36:1327-1336

Martin JA, Wang Z (2011) Next-generation transcriptome assembly. Nat Rev Genet 12:671-682

Meglic V, Smith RR (1992) Self-Incompatibility and Seed Set in Colchicine-, Nitrous Oxide-, and Sexually Derived Tetraploid Red Clover. Crop Sci 32:1133-1137

Ohto M-a, Floyd SK, Fischer RL, Goldberg RB, Harada JJ (2009) Effects of APETALA2 on embryo, endosperm, and seed coat development determine seed size in Arabidopsis. Sex Plant Reprod 22:277-289

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061-1067

Phan HA, Iacuone S, Li SF, Parish RW (2011) The MYB80 Transcription Factor Is Required for Pollen Development and the Regulation of Tapetal Programmed Cell Death in Arabidopsis thaliana. Plant Cell 23:2209-2224

Ranocha P, Dima O, Nagy R, Felten J, Corratgé-Faillie C, Novák O, Morreel K, Lacombe B, Martinez Y, Pfrunder S, Jin X, Renou J-P, Thibaud J-B, Ljung K, Fischer U, Martinoia E, Boerjan W, Goffner D (2013) Arabidopsis WAT1 is a vacuolar auxin transport facilitator required for auxin homoeostasis. Nat Commun 4:2625

Ravagnani A, Abberton MT, Skøt L (2012) Development of Genomic Resources in the Species of Trifolium L. and Its Application in Forage Legume Breeding. Agronomy 2:116

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140

Rosas-Santiago P, Lagunas-Gomez D, Barkla BJ, Vera-Estrella R, Lalonde S, Jones A, Frommer WB, Zimmermannova O, Sychrova H, Pantoja O (2015) Identification of rice cornichon as a possible cargo receptor for the Golgi-localized sodium transporter OsHKT1;3. J Exp Bot 66:2733-2748

Rubinstein AL, Broadwater AH, Lowrey KB, Bedinger PA (1995) Pex1, a pollen-specific gene with an extensin-like domain. Proc Natl Acad Sci USA 92:3086-3090

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 Res 12:301-364

Shendure J, Ji H (2008) Next-generation DNA sequencing. Nat Biotechnol 26:1135-1145

Sjödin J, Ellerström, S (1986) Autopolyploid forage crops. In: Svalöf 1886-1986. Research and Results in Plant Breeding. (Ed. G. Olsson) LTs förlag. Stockholm, Sweden. 102-113.

Sleper DA, Poehlman JM (2006) Breeding field crops (No. Ed. 5). Blackwell publishing

Steiner JJ, Smith RR, Alderman SC (1997) Red Clover Seed Production: IV. Root Rot Resistance under Forage and Seed Production Systems. Crop Sci 37:1278-1282

Supek F, Bošnjak M, Škunca N, Šmuc T (2011) REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. PLoS ONE 6:e21800

Taylor NL, Quesenberry KH (1996) Red Clover Science. Vol. 28. Current Plant Science and Biotechnology in Agriculture. Kluwer Academic Publishers, Dordrecht

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotech 28:511-515

Valle O (1961) Pollination problems in tetraploid red clover. In: S. Ellerström, editor, Precedings of the Symposium on fertility in tetraploid red clover, Eucarpia, Section “Fodder crops” group, Svalof, Sweden. 26-28 Sept. 1961, pp 28-33

Vasiljević S, Momirović GŠ, Katić S, Mihailović V, Lukić D, Zivanović T (2000) Mutual relationships among green forage and seed yield components in genotypes of red clover (Trifolium pratense L.). Genetika 32: 188-191

Verma S, Gupta S, Bandhiwal N, Kumar T, Bharadwaj C, Bhatia S (2015) High-density linkage map construction and mapping of seed trait QTL in chickpea (Cicer arietinum L.) using Genotyping-by-Sequencing (GBS). Sci Rep 5:17512

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

Vleugels T, Cnops G, van Bockstaele E (2013) Screening for resistance to clover rot (Sclerotinia spp.) among a diverse collection of red clover populations (Trifolium pratense L.). Euphytica 194:371-382

Ward JA, Ponnala L, Weber CA (2012) Strategies for transcriptome analysis in nonmodel plants. Am J Bot 99:267-276

Wexelsen H, Vestad R (1954) Observations on pollination and seed setting in diploid and tetraploid red clover. European Grassland Conference, Paris. p. 64-68

Woodger FJ, Gubler F, Pogson BJ, Jacobsen JV (2003) A Mak-like kinase is a repressor of GAMYB in barley aleurone. Plant J 33:707-717

Xu Z, Dooner HK (2006) The Maize aberrant pollen transmission 1 Gene Is a SABRE/KIP Homolog Required for Pollen Tube Growth. Genetics 172:1251-1261

Xue L-J, Zhang J-J, Xue H-W (2012) Genome-Wide Analysis of the Complex Transcriptional Networks of Rice Developing Seeds. PLoS ONE 7:e31081

Yang C, Vizcay-Barrena G, Conner K, Wilson ZA (2007) MALE STERILITY1 is required for tapetal development and pollen wall biosynthesis. Plant Cell 19:3530-3548

Yates SA, Swain MT, Hegarty MJ, Chernukin I, Lowe M, Allison GG, Ruttink T, Abberton MT, Jenkins G, Skøt L (2014) De novo assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC Genomics 15:1-15

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J (2006) WEGO: a web tool for plotting GO annotations. Nucleic Acids Res 34:W293-297

Zhang F, Sun Y, Pei W, Jain A, Sun R, Cao Y, Wu X, Jiang T, Zhang L, Fan X, Chen A, Shen Q, Xu G, Sun S (2015) Involvement of OsPht1;4 in phosphate acquisition and mobilization facilitates embryo development in rice. Plant J 82:556-569

Figure Legends

Figure 1. A) Cumulative curve for number of flower heads for four tetraploid red clover genotypes: two low seed yielding 'Tripo' genotypes and two high seed yielding 'Lasang' genotypes. The x-axis indicates the date when counting of flower heads was performed. B) Seed yield per plant in low seed yielding tetraploid red clover variety 'Tripo' and high seed yielding tetraploid red clover variety 'Lasang'.

Figure 2. The number of differentially expressed transcripts identified using reference based assembly (A1) and *de novo* based assembly (A2). Venn diagrams represents the number of up- and down-regulated transcripts that were common and specific for the pairwise comparisons using the reference based assembly in pairwise comparisons between early vs. late flowering, early vs. mid flowering, and mid vs. late flowering in genotypes 'Tripo' and 'Lasang' using a false discovery rate (FDR) of <0.05.

Figure 3. Heat maps of differentially expressed genes detected using *de novo* assemblies for each genotype and grouped according to their expression patterns. Y-axis represents the experimental conditions. A) Tripo42 B) Tripo55 C) Lasang77 D) Lasang108. EF; early flowering, MF; middle flowering. LF; late flowering.

Figure 4. Gene ontology classifications of differentially expressed genes observed during pairwise comparisons of 'Tripo' and 'Lasang' genotypes generated by the WEGO tool (http://wego.genomics.org.cn/cgi-bin/wego/index.pl) using the newest GO archive provided. The results are distributed in three main GO categories: cellular component, molecular function and biological process. The right y-axis indicates the number of genes in a GO category for each genotype. The left y-axis indicates the percentage of a specific category of genes in the respective main categories for each genotype.

Figure 5. Gene ontology (GO) enrichment analysis by Fischer's exact test. The scatterplot of GO terms which are associated with differentially expressed genes in 'Tripo' (A) and in 'Lasang' (B) shows the cluster representatives (i.e. terms remaining after the redundancy reduction) in a two dimensional space derived by applying multidimensional scaling to a matrix of the GO terms' semantic similarities. Bubble color indicates p-value (-log10 p-value); size indicates the frequency of the GO term in the underlying GOA database (bubbles of more general terms are larger). In 'Tripo' (low seed yield) (A) has higher representation of GO of flower and shoot system development. In 'Lasang' (high seed yield) (B) has higher representation of GO for pollination or pollen-pistel interactions (multi-organism process).

Figure 6. Comparative mapping of significant differentially expressed genes (DEGs) detected in this study to the red clover seed yield related QTL (Hermann et al., 2006). The distribution of DEGs on the seven linkage groups of red clover. The SSR markers (denoted in brackets) (Hermann et al., 2006), linked to QTL (denoted in bars) which are co-located to the DEGs are highlighted. In LG1, MYBA2 mapped to the QTL for thousand seed weight (TSW). In LG2, 4CLL2 and CNIH3 mapped to QTL for time of flowering (TOF) and TSW. In LG3, EF113 mapped to QTL for seed yield per plant (SYP). In LG6, GAE4, mapped to the QTL for seed yield per head (SYH).

Tables

Table 1. Characteristics of the *de novo* transcriptome assemblies

Genotypes	Total number of contigs	N50 (bp)	Maximum contig length (bp)	Total number of reads
Lasang108	80,328	930	7469	51,000,000
Lasang77	83,489	982	7295	55,000,000
Tripo42	84,545	1016	7447	57,000,000
Tripo55	84,442	982	7339	55,000,000

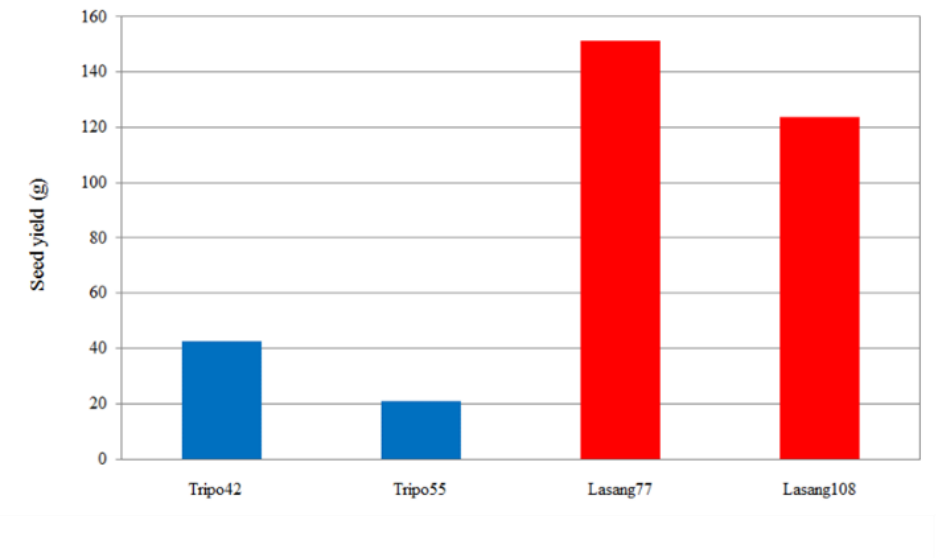
Table 2. Results of CEGMA analysis for *de novo* assembly validation.

Out of 248 CEGs[1]	**Lasang108**	**Lasang77**	**Tripo42**	**Tripo55**
% of fully represented	89.11	92.34	92.34	92.34
% of at least partially represented	97.98	97.58	97.18	97.98
Average number of orthologues per CEG	3.19	3.28	3.18	3.30
% of detected CEGs with more than 1 orthologue	89.59	95.20	90.39	89.96

[1]CEGs: Core Eukaryotic Genes

Table 3. List of differentially expressed genes that can be considered as potential candidate genes involved in seed setting in two red clover genotypes, 'Tripo' and 'Lasang'.

Sequence ID	Description	Chromosome position	LogFC EF/MF	LogFC MF/LF
XLOC_000673	Probable inorganic phosphate transporter 1-4	LG1:16251287-16253813	-100.00	100.00
XLOC_001313	Polyadenylate-binding protein 8	LG1:2381924-2385791	-4.37	-2.69
XLOC_001371	Polyadenylate-binding protein 2	LG1:3705560-3712487	4.44	-1.06
XLOC_002319	Peroxidase 15	LG1:27442419-27444838	-4.42	-4.35
XLOC_002704	Ethylene-responsive transcription factor ERF106	LG1:5667372-5667944	-100.00	100.00
XLOC_005163	Uncharacterized protein	LG2:31978172-31979622	-4.12	-0.79
XLOC_008531	Long chain acyl-CoA synthetase 3	LG3:4387754-4394714	4.78	-0.45
XLOC_009167	Type III polyketide synthase A	LG3:18921285-18925387	-5.42	-3.29
XLOC_009463	Gibberellic acid methyltransferase 2	LG3:26837375-26842634	-4.62	-2.65
XLOC_009550	Protein DETOXIFICATION 45, chloroplastic	LG3:29591163-29602019	5.69	-2.03
XLOC_013980	Tubby-like F-box protein 13	LG4:5131256-5132175	100.00	-2.95
XLOC_015206	Tubby-like F-box protein 11	LG4:5130358-5131150	100.00	-2.86
XLOC_016540	Non-specific lipid-transfer protein C6	LG5:486967-488591	-6.78	-5.40
XLOC_016934	Potassium transporter 19	LG5:12411777-12412424	100.00	-2.14
XLOC_017036	Inorganic phosphate transporter 1-1	LG5:725199-729206	4.13	-0.70
XLOC_019592	Cucumisin	LG6:5874355-5878558	6.48	-2.67
XLOC_020680	Alpha-L-fucosidase 3	LG6:8011079-8011470	100.00	-0.23
XLOC_022367	Elongation factor TuA, chloroplastic	LG7:23203090-23207956	5.30	-2.26
XLOC_022717	Uncharacterized protein	LG7:1020381-1022267	100.00	-1.77
XLOC_024498	Zinc finger protein 4	LG7:11459754-11460460	100.00	-0.20
XLOC_025967	Putative aquaporin NIP4-1	scaf_10543:0-1249	100.00	-2.40
XLOC_028325	Probable E3 ubiquitin-protein ligase ARI10	scaf_1364:21411-23094	100.00	-5.16
XLOC_028939	Glycerophosphodiester phosphodiesterase GDPDL4	scaf_14591:517-949	100.00	-1.10
XLOC_029741	Cystinosin homolog	scaf_15837:737-851	100.00	-2.98
XLOC_029976	Nudix hydrolase 1	scaf_1621:8252-8547	100.00	-1.89
XLOC_034559	Probable magnesium transporter NIPA2	scaf_25880:5-754	100.00	-1.62
XLOC_037847	1-aminocyclopropane-1-carboxylate oxidase homolog 1	scaf_34610:27-594	100.00	-0.94
XLOC_039083	Probable 2-oxoglutarate/Fe(II)-dependent dioxygenase	scaf_3804:1914-2262	100.00	-3.64
XLOC_039720	17.2 kDa class II heat shock protein	scaf_4026:0-243	100.00	-5.90
XLOC_040357	B3 domain-containing protein	scaf_440:86915-89742	100.00	-3.55
XLOC_041752	Cytochrome P450 94B1	scaf_521:81827-84098	5.26	-1.65
XLOC_043901	L-type lectin-domain containing receptor kinase VIII.2	scaf_6561:0-1615	-100.00	100.00
XLOC_044539	Gibberellin 2-beta-dioxygenase	scaf_702:17734-18233	100.00	-0.65
XLOC_045146	WAT1-related protein	scaf_743:105103-107134	100.00	-2.56



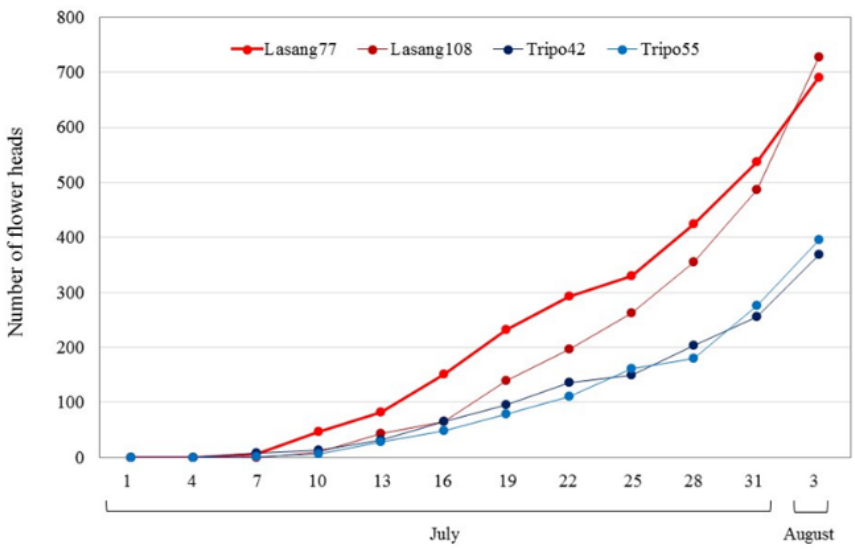


Figure 1. A) Cumulative curve for number of flower heads for four tetraploid red clover genotypes: two low seed yielding 'Tripo' genotypes and two high seed yielding 'Lasang' genotypes. The x-axis indicates the date when counting of flower heads was performed. B) Seed yield per plant in low seed yielding tetraploid red clover variety 'Tripo' and high seed yielding tetraploid red clover variety 'Lasang'.

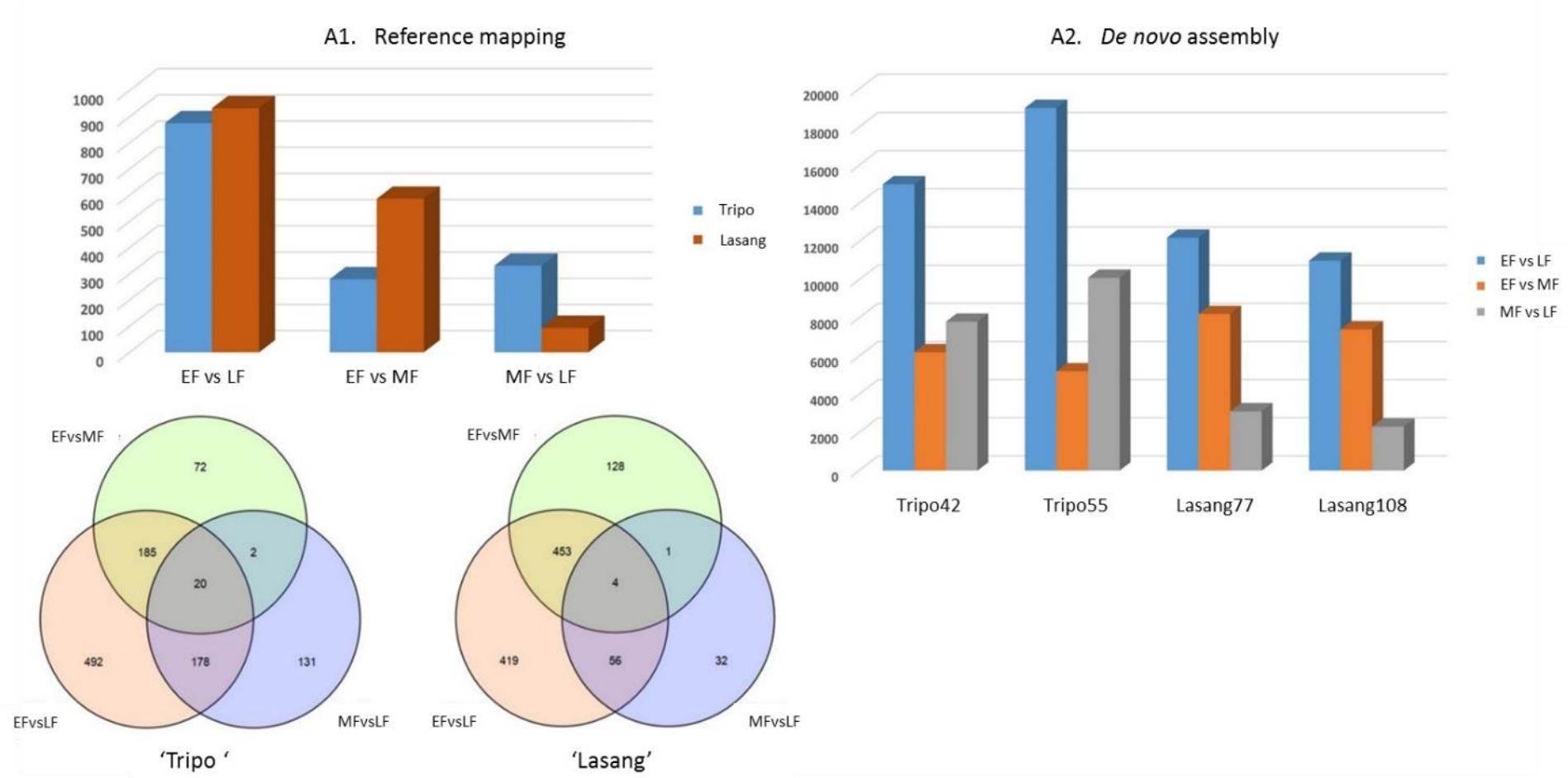


Figure 2. The number of differentially expressed transcripts identified using reference based assembly (A1) and *de novo* based assembly (A2). Venn diagrams represents the number of up- and down-regulated transcripts that were common and specific for the pairwise comparisons using the reference based assembly in pairwise comparisons between early vs. late flowering, early vs. mid flowering, and mid vs. late flowering in genotypes 'Tripo' and 'Lasang' using a false discovery rate (FDR) of <0.05.

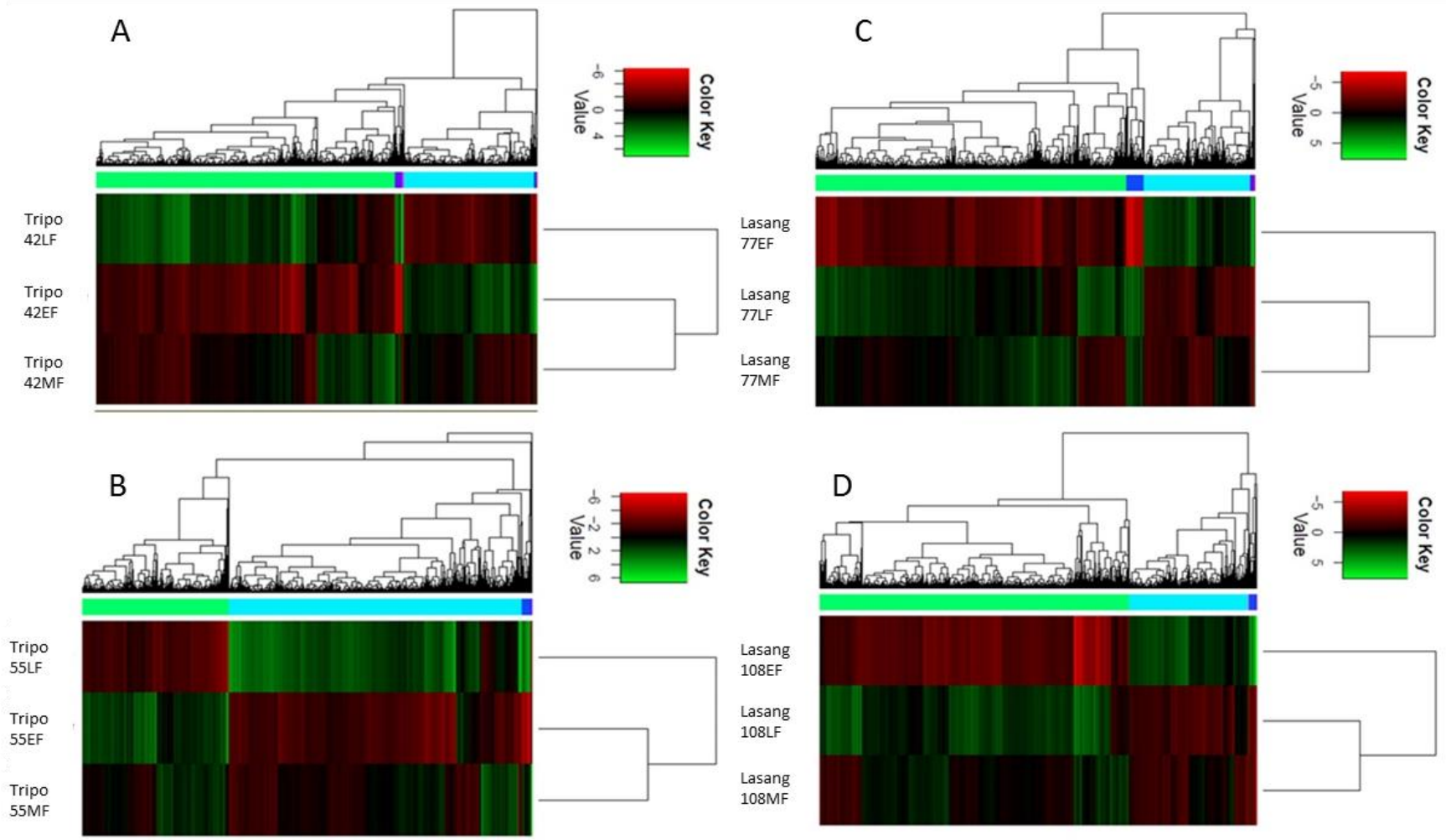


Figure 3. Heat maps of differentially expressed genes detected using *de novo* assemblies for each genotype and grouped according to their expression patterns. Y-axis represents the experimental conditions. A) Tripo42, B) Tripo55, C) Lasang77, D) Lasang108. EF; early flowering, MF; middle flowering, LF; late flowering.

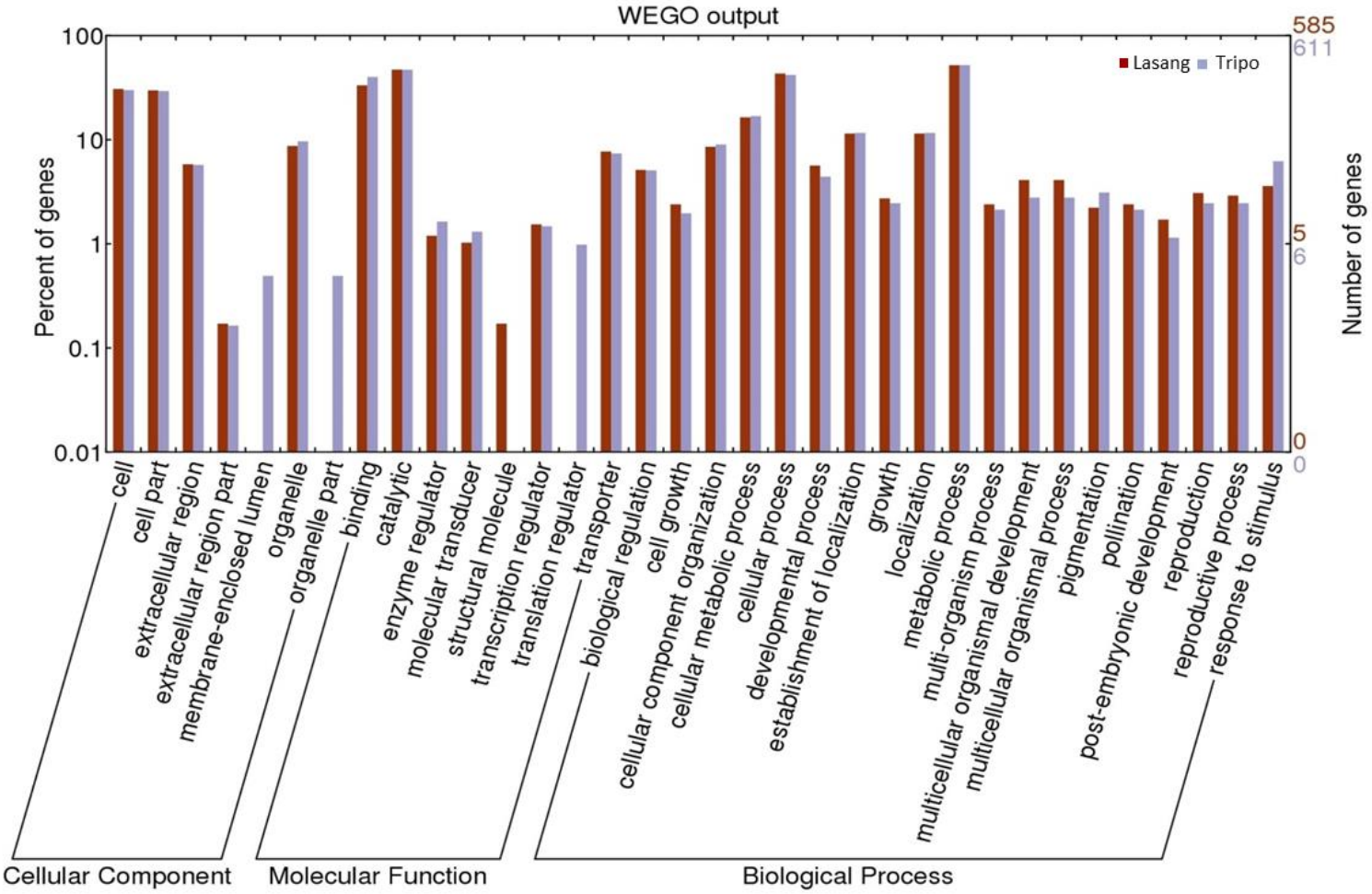


Figure 4. Gene ontology classifications of differentially expressed genes observed during pairwise comparisons of 'Tripo' and 'Lasang' genotypes generated by the WEGO tool (http://wego.genomics.org.cn/cgi-bin/wego/index.pl) using the newest GO archive provided. The results are distributed in three main GO categories: cellular component, molecular function and biological process. The right y-axis indicates the number of genes in a GO category for each genotype. The left y-axis indicates the percentage of a specific category of genes in the respective main categories for each genotype.

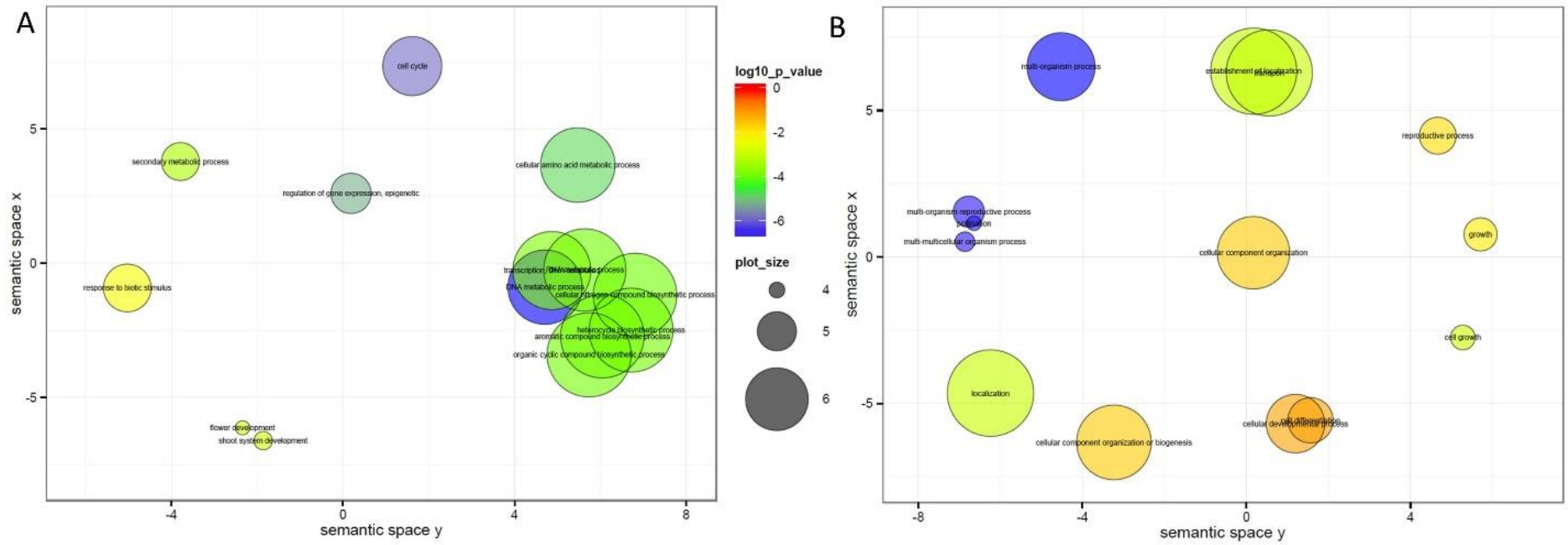


Fig 5. Gene ontology (GO) enrichment analysis by Fischer's exact test.The scatterplot of GO terms which are associated with differentially expressed genes in 'Tripo' (A) and in 'Lasang' (B).shows the cluster representatives (i.e. terms remaining after the redundancy reduction) in a two dimensional space derived by applying multidimensional scaling to a matrix of the GO terms' semantic similarities. Bubble color indicates p-value (-log10 p-value); size indicates the frequency of the GO term in the underlying GOA database (bubbles of more general terms are larger). In 'Tripo' (low seed yield) (A) has higher representation of GO of flower and shoot system development. In 'Lasang' (high seed yield) (B) has higher representation of GO for pollination or pollen-pistel interactions (multi-organism process).

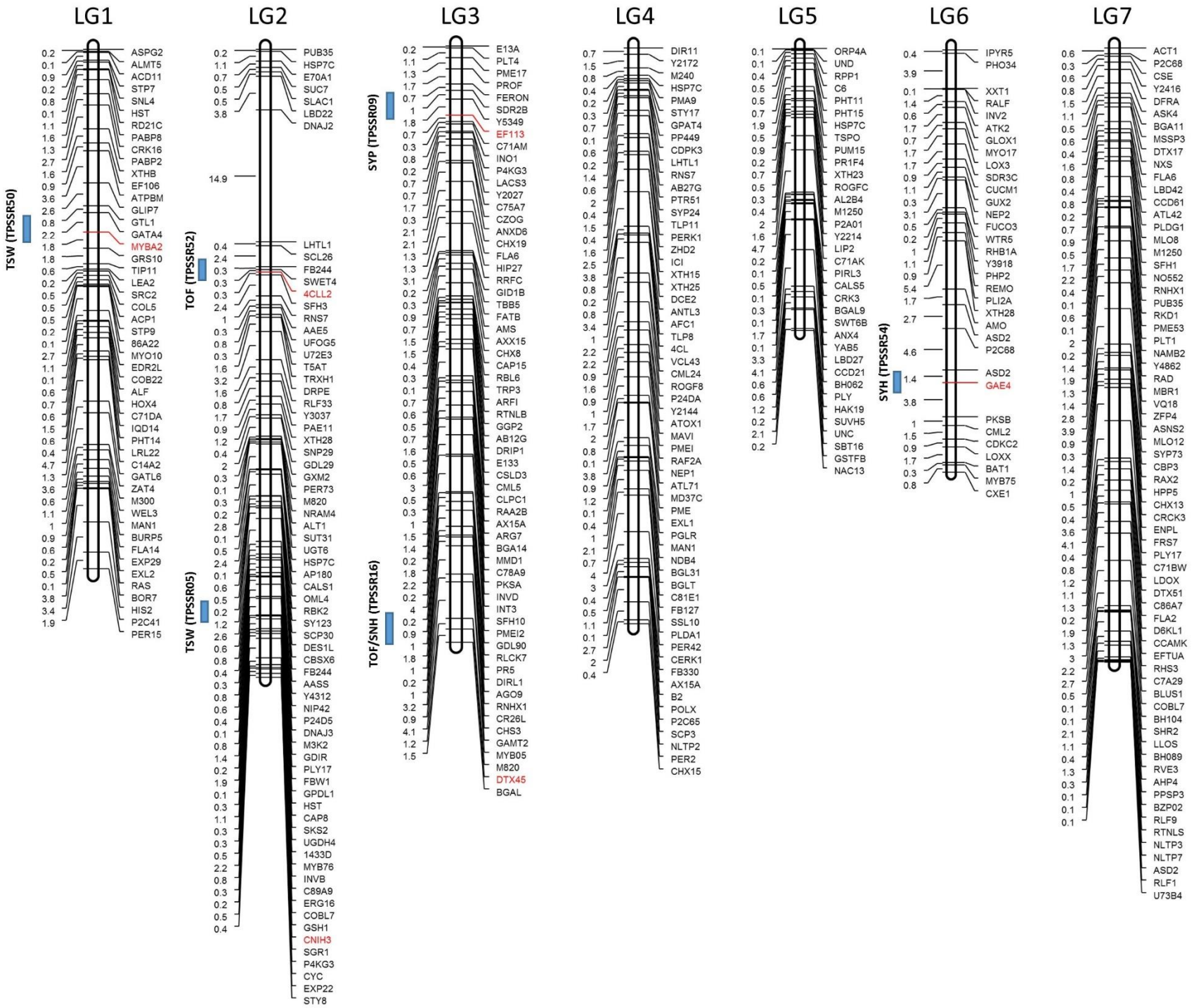
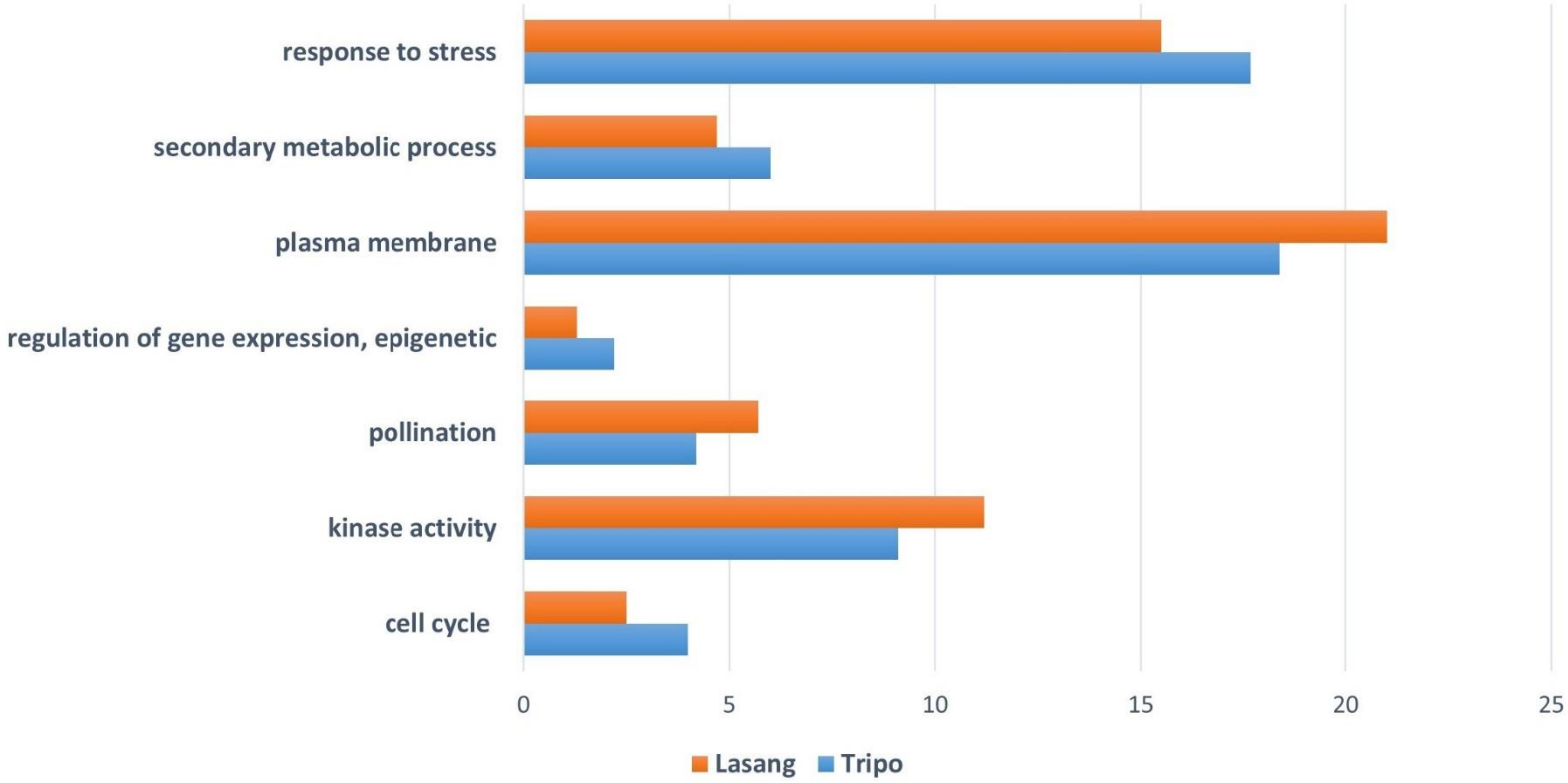

LG1
LG2
LG3
LG4
LG5
LG6
LG7
TSW (TPSSR50)
TOF (TPSSR52)
TSW (TPSSR05)
SYP (TPSSR09)
TOF/SNH (TPSSR16)
SYH (TPSSR54)

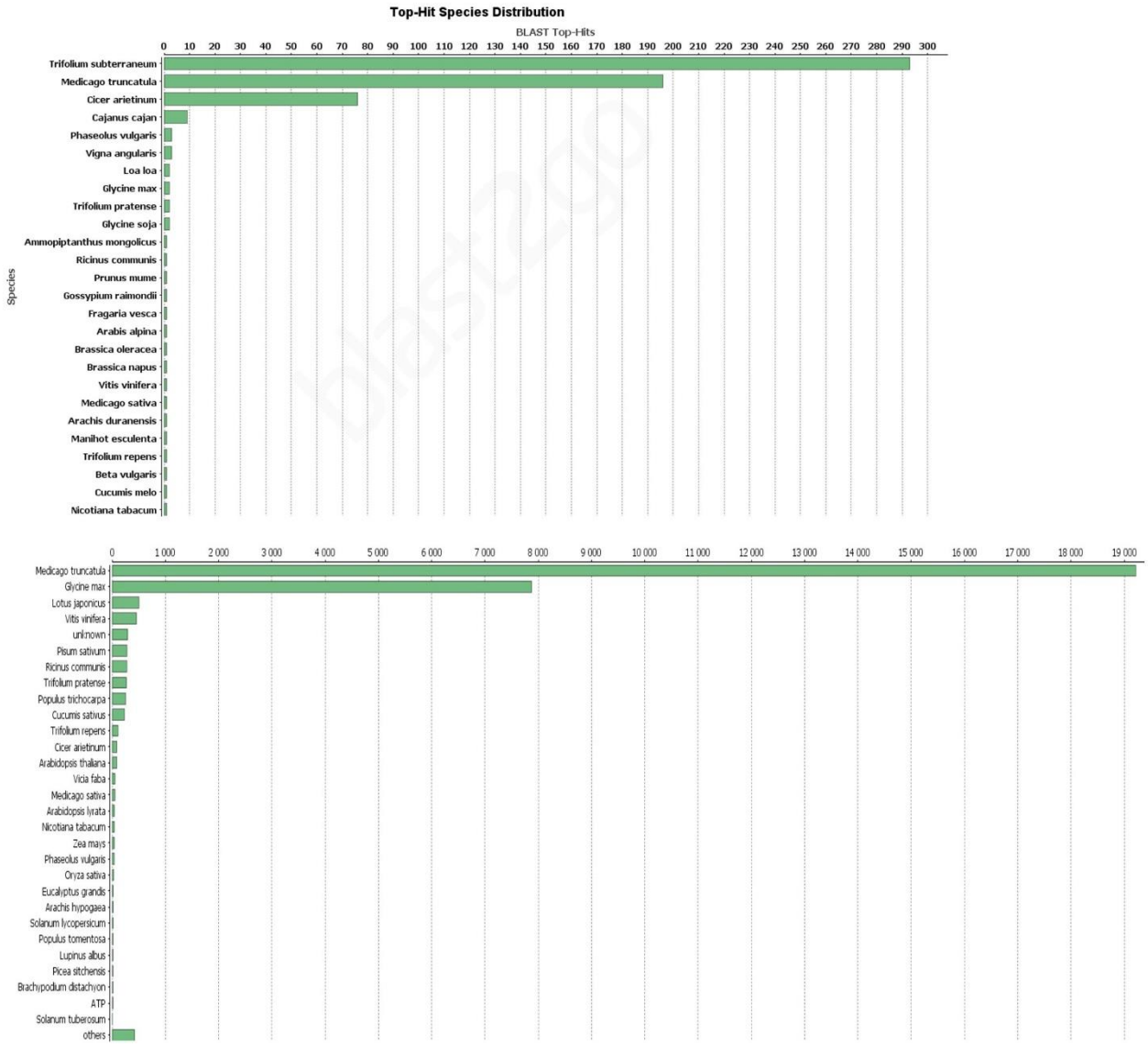
Figure 6. Comparative mapping of significant differentially expressed genes (DEGs) detected in this study to the red clover seed yield related QTLs (Hermann et al., 2006). The distribution of DEGs on the seven linkage groups of red clover. The SSR markers (denoted in brackets) (Hermann et al., 2006), linked to QTLs (denoted in bars) which are co-located to the DEGs are highlighted. In LG1, MYBA2 mapped to the QTL for thousand seed weight (TSW). In LG2, 4CLL2 and CNIH3 mapped to QTLs for time of flowering (TOF) and TSW. In LG3, EF113 mapped to QTL for seed yield per plant (SYP). In LG6, GAE4, mapped to the QTL for seed yield per head (SYH).

Supplementary Table S1. List of pathway-enriched differentially expressed genes in top ten pathways between ‘Tripo’ and ‘Lasang’ at EF-MF and MF-LF stages.

Pathway	Differentially expressed genes (No./%)	Pathway ID
Starch and sucrose metabolism	58 (4.84%)	map00500
Pentose and glucuronate interconversions	34(2.84%)	map00040
Phenylpropanoid biosynthesis	33 (2.75%)	map00940
Purine metabolism	28 (2.34%)	map00230
Thiamine metabolsim	24 (2.00%)	map00730
Amino sugar and nucleotide sugar metabolism	23 (1.92%)	map00520
Drug metabolism - other enzymes	22 (1.83%)	map00983
Aminobenzoate degradation	18 (1.50%)	map00627
Biosynthesis of antibiotics	18 (1.50%)	map01130
Flavonoid biosynthesis	18 (1.50%)	map00941



Suppl. Fig 1. Annotation differences between ‘Lasang’ and ‘Tripo’ genotypes detected by Fischer’s exact test.



Suppl. Fig. 2: The top blast hit-species distribution in reference based and *de novo* based assembly