Nitrate effects on N₂ fixation, growth and feed quality of lucerne and perennial lupin

Eilish BURROWS, Annamaria MILLS, Dan DASH, David JACK, Mitchell ANDREWS and Alistair D. BLACK*

Field Research Centre, Department of Agricultural Sciences, PO Box 85084, Lincoln University, Lincoln 7647, New Zealand
*Corresponding author: alistair.black@lincoln.ac.nz

Abstract
The effects of NO₃⁻ supply (0–500 kg N/ha) on total plant dry weight (DW), shoot N content and nutritional quality, and the proportion of plant N derived from the atmosphere (%Ndфа) were determined for lucerne and perennial lupin using ¹⁵NO₃⁻ under glasshouse conditions. Fodder beet was used as a non-legume reference plant. In both the initial and repeat experiments, total plant DW, shoot N% and shoot nutritional quality for lucerne and perennial lupin were unaffected by NO₃⁻ supply. Total plant DW increased 10-fold and shoot N% tripled for fodder beet with increased N supply. In the initial experiment, the %Ndфа for lucerne decreased from 89 to 37% with increasing N supply from 0 to 500 kg N/ha, while comparable values for perennial lupin were 96 to 64%. In the repeat experiment, %Ndфа decreased from 90 to 49% and 93 to 65% for lucerne and perennial lupin, respectively, with increasing NO₃⁻ supply from 0 to 500 kg N/ha. Both legumes showed an increased reliance on NO₃⁻ with increased soil NO₃⁻ level, but even at 500 kg N/ha (similar to amount of N in sheep urine patch), perennial lupin obtained much of its N from N₂ fixation.

Keywords: legume, plant N, nitrate assimilation, ¹⁵N isotope

Introduction
Most legumes (Fabaceae) can fix atmospheric nitrogen (N₂) via symbiotic bacteria (rhizobia) in root nodules and also utilise soil inorganic N (nitrate (NO₃⁻) and ammonium (NH₄⁺)) when available (Andrews et al., 2013). There are many reports for legumes having increased reliance on soil N in comparison with N₂ fixation as soil N levels increase, but the ability of legumes to use soil N is species dependent (Barron et al., 2011; Menge et al., 2015).

Lucerne (Medicago sativa) and perennial (‘Russell’) lupin (Lupinus polyphyllus) can fix substantial levels of N₂ under suitable conditions in high-country farming systems on the South Island of New Zealand (NZ) (Black et al., 2014; Berenji et al., 2018). Annual yield and nutritive values are greater for lucerne than perennial lupin under optimal conditions, but the latter can grow in acidic soils with high levels of aluminium, which lucerne cannot tolerate. In grazed crops of these legumes, substantial N will be returned to the soil as animal excreta that potentially (after transformation to NO₃⁻) could be leached from the soil into waterways (Andrews et al., 2007; Che et al., 2018). The ability of lucerne and perennial lupin to utilise soil N, and the effect of soil N on N₂ fixation could be important factors determining inputs and losses from the system. For example, if legumes can utilise substantial soil NO₃⁻ and, as a result, their N₂ fixation decreases, this would reduce N input into the system and should be a factor considered in nutrient budgeting models of the system. On-farm surveys of the proportion of plant N derived from the atmosphere (%Ndфа) for lucerne in Australia ranged from 17–90% with averages of 60–65% (Yang et al., 2011; Peoples et al., 2012). The sources of soil N taken up were not identified. The ability of perennial lupin to utilise NO₃⁻ and the impact of soil N/NO₃⁻ on its %Ndфа has not been tested.

Here, the effects of NO₃⁻ supply (0–500 kg N/ha) on total plant dry weight (DW), shoot N content and nutritional quality, and the proportion of plant N derived from the atmosphere (%Ndфа) were determined for lucerne and perennial lupin using ¹⁵NO₃⁻ under glasshouse conditions. Fodder beet (Beta vulgaris) was used as a non-legume reference plant. The objectives of the following study were to determine the ability of perennial lupin and lucerne to utilise soil NO₃⁻, and the impact of soil NO₃⁻ on growth, N₂ fixation and nutritional quality of two legume species.

Materials and Methods
Experimental procedures
Seed of lucerne cv. Force 4 and fodder beet cv. Brigadier was obtained from Seed Force Ltd, Christchurch, NZ. Perennial lupin seed was obtained from Gavin Loxton, Sawdon Station, Lake Tekapo, South Island, NZ.

The initial and repeat experiments were carried out in separate parts of the same glasshouse, and the setup, design and measurements taken were the same. Seed was sown from 13–15 March 2019 into 45 cm tall × 15 cm diameter tubes (‘pots’; 20 seeds/pot) with mesh at the base containing 5 l of N-free potting mix comprised of 80% composted bark and 20% pumice (1–4 mm) to which was added 1 g/l agricultural lime (primarily...
calcium carbonate), 0.3 g/l superphosphate (9% P, 11% S, 20% Ca; Ravensdown, NZ) and 0.3 g/l Osmocote (6 months, 0% N, 0% P, 37% K), 0.3 g/l Micromax trace elements and 1 g/l Hydraflo, all three obtained from Everris International, Geldermalsen, the Netherlands. The pH of the medium was 5.8. All lucerne and perennial lupin pots were watered by weight to field capacity every 3 days with a low NO$_3^-$ supply (0.5 mM KNO$_3$) until the first cut on 14 May. Commercial peat-based rhizobia inoculum for lucerne and perennial lupin (Nodulaid BASF, Canberra, Australia) was mixed with water into a slurry and applied at 5 ml/pot for the first three waterings. All legumes were nodulated at harvest. Fodder beet received 0.5 mM NO$_3^-$ until 4 April then 2 mM KNO$_3$ until 14 May, due to the plants showing N deficiency symptoms. Saucers were placed under each pot to collect leachate which was returned to the pots throughout the experiment. Plants were thinned out to 10 per pot for lucerne and five per pot for perennial lupin and fodder beet two weeks after sowing.

Lucerne and perennial lupin were cut to 4 cm in height using scissors on 14 May to simulate grazing. After this, the different rates of N were applied: 0, 25, 50, 100, 200 and 500 kg N/ha as K$^{15}$NO$_3$ labelled at 10 atom% in 100 ml of water for all treatments. The 500 kg N/ha was assumed to be similar to that within a sheep urine patch (Monaghan et al., 1989; Marsden et al., 2016). Thereafter, all pots were watered (tap water) by weight to field capacity every 3 days until harvest on 12 June and 20–21 June for the initial and repeat experiments, respectively. The temperature in the glasshouse ranged from 14 to 28°C during the experiments.

At harvest, plants from all pots were divided into shoot and root, dried at 60°C for 7 days then weighed. Shoot and root material was then ground, and total N content of 0.2 g samples of roots was determined using a CN elemental analyser (Elementar VarioMax CN Elemental Analyser, GmbH, Hanau, Germany). The ground shoot material was analysed for $^{15}$N/$^{14}$N with a Sercon (Crewe, UK) GSL (gas, liquid, solid) elemental analyser attached to a Sercon 20–22 isotope ratio mass spectrometer. The %Nd$^a$ was determined via the $^{15}$N isotope dilution method (Unkovitch et al., 2008) as:

$$\%\text{Nd}^a = \left(1 - \text{atom}\%^{15}\text{N} \times 1.00\right) \times 100$$

The nutritional quality of the lucerne and perennial lupin (dry matter digestibility (DMD), crude protein (CP) and metabolizable energy (ME)) was determined on ground shoot material using near infrared spectroscopy (NIRS; FOSS NIRSystems 5000, FOSS NIRSystems Inc., Laurel, MD, USA).

Experimental design and data analysis
The initial and repeat experiments were conducted as a fully randomised design with three replicate pots per N treatment for all three species. A two-way analysis of variance (ANOVA) was carried out on all data with plant species and N rate as fixed factors. All effects discussed had an F ratio with a probability P<0.01. Regression analysis was carried out on the data for fodder beet, and a quadratic model used with R$^2$ given. For legumes, exponential models were fitted to the %Nd$^a$ data.

Results
The main effects were the same for the initial and repeat experiments. Results for the initial experiment are presented in Figures 1 and 2. Total plant DW was greater and shoot to root DW ratio (S:R) was lower for perennial lupin than lucerne regardless of N supply (Figure 1a, b). For both legume species, neither total plant DW nor S:R were affected by N supply. For fodder beet, total plant DW increased four-fold and S:R three-fold with increased N supply from 0 to 500 kg N/ ha. Total plant DW was less for fodder beet than for perennial lupin or lucerne at 0–25 kg N/ha, but greater for fodder beet than the two legumes at 200–500 kg N/ha. S:R was greater for fodder beet than the two legumes, regardless of N supply.

Shoot %N was greater for lucerne than perennial lupin at all N levels, but for both species was unaffected by N supply (Figure 2a). Shoot %N for fodder beet increased four-fold with higher N supply from 0 to 500 kg N/ha. Values were lower for fodder beet than perennial lupin or lucerne at 0–100 kg N/ha, but greater for fodder beet than perennial lupin and similar for fodder beet and lucerne at 500 kg N/ha. In the initial experiment, the %Nd$^a$ for lucerne decreased from 89 to 37% with increasing N supply from 0 to 500 kg N/ha: comparable values for perennial lupin were 96 to 64% (Figure 2b). In the repeat experiment, %Nd$^a$ decreased from 90 to 49% for lucerne and 93 to 65% for perennial lupin with increasing NO$_3^-$ supply from 0 to 500 kg N/ha.

Nutritional quality of the legumes was unaffected by N supply. Across both experiments, values were 71–73% DMD, 29–30% CP and 11–12 MJ/kg DM ME for lucerne and 76–78% DMD, 23–25% CP and 11 MJ/kg DM ME for perennial lupin. This represented high nutritional quality for both species (Machado et al., 2007; Ryan-Salter 2019) and indicated that increased reliance on NO$_3^-$ with decreased N$_2$ fixation did not affect the nutritional quality of the shoots.

Discussion
The results indicated that both lucerne and perennial lupin have increased reliance on NO$_3^-$ nutrition with increased soil NO$_3^-$ level. Increased uptake of soil NO$_3^-$
was matched by a similar size decrease in N\textsubscript{2} fixation, such that total plant N and DW changed little with N supply. In contrast, fodder beet, the control plant, showed substantial increases in total plant DW, S:R and tissue N content with increased supply, as would be expected for most non-legume species (Andrews et al., 2013). At comparable NO\textsubscript{3} supply, the %N\text{dfa} was greater for perennial lupin than lucerne and, even at 500 kg N/ha (similar to a sheep urine patch), perennial lupin derived the major proportion of its N from N\textsubscript{2} fixation. This indicated that, under grazing in the field, perennial lupin will still maintain high levels of N\textsubscript{2} fixation, but this needs to be further tested on mature plants under field conditions.

Ryan-Salter (2019) carried out a \textsuperscript{15}NO\textsubscript{3} experiment similar to that described here, and reported that %N\text{dfa} was 38% for perennial lupin and 26% for lucerne at 600 kg NO\textsubscript{3}\textsuperscript{-}/N/ha. These results were lower than predicted from the values obtained in the current study, but can be explained, at least in part, by the use of a non-legume reference plant in the current experiment, but not in the Ryan-Salter (2019) study.

On a per pot basis (≡ area), dry matter growth was greater for lucerne than perennial lupin as there was twice as many plants per pot (10) with lucerne. This was more indicative of the field situation in high-country farming systems on the South Island of NZ, where dry matter yields/area/annum are likely to be greater for lucerne than perennial lupin.

At pot level, the amount of N\textsubscript{2} fixed was greater with lucerne than perennial lupin at 0–100 kg N/ha, but similar for the two species at 200 and 500 kg N/ha, due to a greater decrease in N\textsubscript{2} fixation associated with a higher increase in NO\textsubscript{3}\textsuperscript{-} assimilation for lucerne. Scaling this to field level, the range of values for %N\text{dfa} for lucerne in pots, 90–37%, was in the range quoted for lucerne on farms in Australia (Yang et al., 2011; Peoples et al., 2012). Yang et al. (2011) estimated annual average N\textsubscript{2} fixation of 322 kg N/ha/annum with uptake of 181 kg N/ha/annum from the soil. The sources of soil N were not identified, but nitrate reductase (a substrate (NO\textsubscript{3}\textsuperscript{-}) induced enzyme) was substantially greater in lucerne shoots than in associated species, which indicated that it assimilated substantial NO\textsubscript{3}\textsuperscript{-} in the shoot.
Overall, the current results obtained, and in the literature, indicated that lucerne has a high capacity to utilise soil NO$_3^-$, This uptake and its assimilation resulted in decreased N$_2$ fixation, and these responses were of a magnitude that should be considered when modelling N inputs and losses from lucerne systems. The results indicated that perennial lupin has a lower ability to utilise soil NO$_3^-$ and the effect of soil NO$_3^-$ on its N$_2$ fixation is less. Further studies are required to confirm this in the field.

**Conclusions**

Both lucerne and perennial lupin showed more reliance on soil NO$_3^-$ with increased soil NO$_3^-$ level. Higher uptake of soil NO$_3^-$ was matched by a similarly sized decrease in N$_2$ fixation, such that total plant N and DW changed little with N supply. At comparable NO$_3^-$ supply levels, perennial lupin utilised less soil NO$_3^-$ than lucerne. Even at 500 kg N/ha (about equal to N in sheep urine patches) perennial lupin utilised the major proportion of its N from N$_2$ fixation. Nutritional quality of the legumes was unaffected by N supply.

**ACKNOWLEDGEMENTS**

Funding for Alistair Black and Annamaria Mills was provided by Beef + Lamb New Zealand, Ministry of Business, Innovation and Employment, Seed Force New Zealand and PGG Wrightson Seeds under the ‘Hill Country Futures’ research programme (BLNZT1701).

**REFERENCES**


Developing genomic selection for dry matter yield in white clover

Grace O. EHOCE1,2, Sai Krishna AROJU1, Anna LARKING1, Ruy JAUREGUI1, Greig COUSINS2, Jessica O’CONNOR1, M. Z. Zulfi JAHUFER1, and Andrew G. GRIFFITHS1,*

1AgResearch, Grasslands Research Centre, Private Bag 11008, Palmerston North, New Zealand
2PGG Wrightson Seeds Ltd, Grasslands Research Centre, Palmerston North, New Zealand
*Corresponding author: andrew.griffiths@agresearch.co.nz

Abstract
Genomic selection (GS) integrates DNA markers and trait data to develop a model that enables prediction of performance (genomic-estimated breeding values; GEBVs). GS can improve the effectiveness of breeding programmes, especially for complex traits, such as dry matter yield (DMY). DMY data were generated from a training population of 200 white clover half-sibling (HS) families assessed in multi-location field trials over two years. This generated a GS prediction model after integrating genotyping-by-sequencing marker data from parents of HS families with HS DMY data. Two selection strategies were compared: a conventional method where individuals were chosen randomly from the phenotypically highest ranked HS families (HSp); or where GEBVs were used to select the best individuals within the highest ranked HS families (ApWFGS). Mean predicted DMY GEBVs of the selected plants, as well as the predicted response to selection, were compared with those of the base population. This study showed that, compared with conventional selection (HSp), incorporating genomic selection using ApWFGS was predicted to double the increase in DMY and responses to selection relative to the base population. Synthetic populations based on these selections have been generated and will be tested in field trials to empirically evaluate selection for a complex trait.

Keywords: Trifolium, GBLUP, genotyping, prediction, modelling

Introduction
White clover (Trifolium repens L.) is an integral component of temperate pastoral agriculture, where it provides a low-cost, high quality feed source throughout the year (Caradus et al., 1997; Jahuf er et al., 2002). The advantages of white clover are not limited to providing a rich source of protein for livestock, but improve soil fertility through nitrogen fixation (Woodfield and Caradus 1996). Effective inoculation of white clover by the symbiotic soil bacterium Rhizobium leguminosarum var. trifolii, underpins nitrogen fixation. This source of nitrogen from the atmosphere becomes available to the companion sward, thereby reducing reliance on synthetic fertiliser (Gibson and Cope 1985). The growing emphasis on environmental sustainability and the anticipated shift to low input agriculture make white clover an attractive forage crop due to the environmental and economic benefits it provides.

Conventional white clover breeding is complicated as many important traits, including dry matter yield (DMY), are under quantitative genetic control and are highly influenced by the environment. This results in low heritability estimates, which make trait improvement challenging. Not only is DMY traditionally difficult to phenotype, it is also usually assessed at late growth stages, thereby increasing the length of the breeding cycle, with consequent negative impacts on the rate of genetic gain and cultivar development. The annual rate of genetic gain in white clover and other forage, such as perennial ryegrass, has been estimated to be less than 1% per year (Hayes et al., 2013; Hoyos-Villegas et al., 2019). Breeding strategies, such as half-sibling (HS) family selection, can access 25% of the total additive variation among families but leaves the remaining 75% additive variation within-family unexploited.

Application of a breeding method that can access this 75% within-family variation may increase the magnitude of genetic gain and reduce the time frame for releasing new cultivars (Vogel and Pedersen 1993).

Genomic selection (GS) is a recent technology that is routinely applied to animals and crops and can improve the efficiency and effectiveness of breeding programmes, especially for complex traits such as DMY. GS is based upon a statistical model that is trained by integrating trait and genome-wide DNA marker data acquired from a training population of plants or families sampled from a breeding programme. The model can subsequently be used to predict trait phenotype/breeding values (Genomic-Estimated Breeding Values; GEBVs) for non-phenotyped individuals in the breeding programme, based solely on DNA marker information. This underpins breeding efficiencies including: 1) Selecting best individuals early for traits that are usually measured over multiple years, which reduces selection cycle time and increasing genetic gain (AG) per unit time; 2) Enhanced ability to screen large numbers of selection candidates, increasing selection intensity; 3) Enabling access to within-family additive variation in genetically structured populations.

Several simulated and empirical studies have shown that GS can outperform phenotypic selection,
resulting in more genetic gain per breeding cycle or unit of time (Massman et al., 2013; Faville et al., 2018; Annicchiarico et al., 2019; Esfandyari et al., 2020). Currently, no studies evaluating or validating the use of GS for white clover DMY have been published. White clover cultivars are often developed using among HS family phenotypic selection (HS\(_p\)) breeding methodology.

The objective of the following study was to enhance this strategy by integrating GS in a two-step process. First, best HS families were identified based on trial phenotypic data, then GS was applied to select the best individuals within each family based on their DNA profile (\(A_pW_{GS\text{HS}}\)). This paper reports the development of a GS model for DMY in white clover, its application in selection and the predicted response to selection.

**Materials and Methods**

**Plant material and field trial**

A white clover (Trifolium repens) training population, comprising half-sibling (HS) families derived from a polycross of 274 \(F_1\) maternal parents, was generated in the summer of 2015/2016 in a bee-proof cage using bumble bees (Bombus spp.) which had been pre-washed to remove wild pollen. The highest 200 seed-yielding \(F_1\) HS families were established in row-column, replicated (three replicates), multi-year field trials. Two locations were used for the trial: AgResearch Grasslands Research Centre in Palmerston North, Manawatu (Aorangi) (40.38˚S, 175.61˚E); and the AgResearch Ruakura Research Farm in Hamilton, Waikato (37.77˚S, 175.31˚E). The soil types at the Palmerston North and Ruakura sites were Kairanga fine sandy loam and peaty silt loam soil, respectively. Plant material was prepared by germinating a random sample of seed from each of the 200 HS families and grown in standard glasshouse conditions. Fifteen plants from each HS family were then transplanted into 0.5 m by 0.75 m plots in a sward of perennial ryegrass (Lolium perenne) cv ‘Ceres One50’ with AR37 endophyte in August/September 2016. No irrigation was applied post-trial establishment. The trial plots were grazed by cattle when herbage mass was between 2500–2800 kg/ha DM to residuals of 1100 – 1200 kg/ha DM measured by a plate meter (Jenquip, Feilding, New Zealand) after each grazing. To assess DMY, two mechanical harvests in each plot were performed, one each in November 2017 and 2018 at herbage mass accumulation of 2500–2800 kg/ha DM. A 0.2 m\(^2\) quadrant was randomly placed in each plot and the above-ground biomass removed. Harvested samples were then separated into white clover and ryegrass components, oven-dried and weighed. Due to the size of the trial (672 plots including repeated checks) at each site, giving a total of 1,344 plots, full DM harvests were made once a year, with focus on the spring growth phase. These annual harvests were interspersed with seasonal calibration cuts, which were not included in this set of genomic predictions.

**Statistical analysis**

Residual Maximum Likelihood (REML) (Patterson and Thompson 1971; Harville 1977) was conducted using DeltaGen software (Jahufer and Luo 2018) and enabled estimation of variance components for genetic and nongenetic effects and Best Linear Unbiased Predictors (BLUPs) for each HS family. In a mixed linear model, years, sites and repeated checks were considered fixed effects, while the HS families, replicates, rows and columns of the experimental design were considered random effects. The statistical significance of the variance components was estimated using deviance of log-likelihood, as suggested by Galwey (2006) and significance was indicated at \(P<0.05\).

**Among and within family selection**

Among and within family selection pressure was 10% and 5% respectively. The highest ranked 10% (\(n = 20\)) HS families based on the DMY BLUPs across years and locations were selected. A random sample of 20 seeds was sampled from remnant seed of each of the selected 20 HS families, resulting in 400 selection candidates that were scarified, germinated, and grown under standard glasshouse conditions until three trifoliate leaves were present. The aim was to generate synthetic populations based on polycrosses of 20 plants, comprising either one plant selected randomly or based on genomic-estimated breeding values (GBEVs; described below) from each of the 20 highest ranked HS. This represented 5% within HS family selection pressure. In addition, 20 individuals were grown from each of the lowest 10% (\(n = 20\)) HS families to provide material (\(n = 400\)) that, when combined with the highest ranked 20 families, represented a base population against which derived GEBVs could be compared.

**Genotyping-by-sequencing, genomic relationship matrix and genomic selection**

Genotype data was obtained from 200 parents of the HS families (for training the GS prediction model), 400 selection candidates from the highest ranked HS families and 400 individuals from lowest ranked HS families (for applying to the GS prediction model). DNA was extracted from approximately 100 mg of leaf tissue per plant, as described by Anderson et al. (2018), and used to generate genotyping-by-sequencing (GBS) libraries, as described in Griffiths et al. (2019). In this case, the choice of restriction enzyme was ApeKI. Briefly, DNA from the 400 selection candidates, including the control samples and duplicates, were distributed across eleven
The performance of the model was assessed by Monte-Carlo crossvalidation. Here, the whole data set was divided into training (80%) and test sets (20%), where phenotypes of the test set were assumed to be unknown and predicted by the trained model (Erbe et al., 2010). Predictive ability was calculated as the Pearson correlation coefficient between the observed DMY BLUP and the predicted value after 100 iterations. The GEBVs for the 400 selection candidates from the highest ranked HS and the 400 individuals from the lowest ranked HS provided a mean GEBV for the base population, against which the selections could be compared.

**Selection strategies**

Two selection populations at a 10% among HS + 5% within HS selection pressure were made, to reflect typical HS family breeding as well as incorporating GS as shown in Figure 1. These were:

1. HS_p: random sampling of a single plant from within each of the highest ranked (based on phenotype) 20 HS families (n = 20 plants in total).
2. A_pWF_GS: selection of the plant with the highest GEBV from within each of the highest ranked (based on phenotype) 20 HS families (n = 20 plants in total).

In summary, the highest ranked (Top) (n = 20) HS families were chosen based on dry matter yield (DMY) BLUPs across years (2017/2018) and locations (Aorangi/Ruakura) providing a 10% among HS family phenotypic selection pressure. Twenty plants were grown from remnant seed from each selected HS family, genotyped and GEBVs derived. Two synthetic populations based on polycrosses of 20 plants each were generated, in which a single plant was selected either randomly (HS_p) or using GEBVs (A_pWF_GS) from each family and represented 5% within HS selection pressure (Figure 1).

**Synthetic population development**

The individuals selected for the HS_p and A_pWF_GS
synthetic populations were grown under glasshouse conditions at Palmerston North until mature and then transferred outside to vernalise over winter (2020). Over the summer of 2020/2021, each synthetic population was generated by random polycrossing of the selected 20 individuals in a bee-proof cage using bumble bees (Bombus spp.) which had been washed to remove wild pollen. After successful pollination, seed was harvested from each HS family in each synthetic population separately and then cleaned. An equal quantity of seed from each HS family within each synthetic population was combined to generate a balanced bulk representing that population.

Predicted response to selection
The relative efficiency of different selection strategies can be evaluated by estimating the response to selection per selection cycle (Hallauer and Filho 1981). Predicted response to selection, \( R \), was calculated according to Equation 2 (Lush 1937).

\[
R = h_n^2 S
\]

Where: \( R \) was the response to selection; \( h_n \) the narrow-sense heritability; and the selection differential defined as the difference between the mean of selected parents and the mean of the population from which the parents were selected.

Results
Phenotype data and variance components
Dry matter yield (DMY) best linear unbiased predictor (BLUP) values for the HS families, based on analysis of data across two years and two locations (Aorangi/Ruakura), were estimated. These BLUPs (Supplementary Table 1) were used to select the HS families. There was significant (\( P<0.01 \)) additive genetic and genotype-by-environment (GxE) interaction variation among the white clover genotypes for DMY across years and locations. Significant (\( P<0.05 \)) additive variance, genotype-by-location interaction was estimated (Table 1). Family mean narrow-sense heritability was estimated as \( 0.38 \pm 0.09 \).

Predictive ability of the genomic prediction model
GBS data were generated for the 200 HS family parents in the training population as well as the 400 selection individuals from the highest and lowest ranked HS families. This identified 361,220 SNPs which, after filtering, were reduced to 110,000 SNP genotypes. These were distributed across the genome and ranged from ~3000 to ~10,500 SNPs for chromosomes 14 and 1, respectively, with a mean of 6875 SNPs per chromosome (Supplementary Figure 1). This genotype information was used to generate a genomic relationship matrix (GRM). A GS prediction model was derived using genotype and field trial phenotype data from the training population individuals for DMY across years and locations. This enabled calculation of genomic estimated breeding values (GEBVs) for DMY for each selection individual in the GRM (Supplementary Table 2). The predictive ability of the model, calculated as the Pearson correlation coefficient between the observed DMY BLUP and the predicted value, was estimated at \( r = 0.30 \). This equated to a prediction accuracy of 0.48 when divided by the square root of the narrow-sense heritability.

Comparison of GEBVs from different selection strategies
The influence of different selection strategies on the estimated DMY phenotypes of the plants selected for polycrossing was determined by comparison of GEBVs. Mean GEBVs were calculated for the 400 plants each from the highest and lowest ranked half-sibling (HS) families representing the base population, and the 20 individuals each for the HS\(_P\) and the \( A_pWF_{GS} \) selections.

The plants selected using HS\(_P\) combined with the use of GS for within HS family selection (\( A_pW_{GS}HS \)) had the highest predicted DMY BLUP phenotypes (mean GEBV = 2.85), showing a 6.9% increase over the base population (2.66), and 3.7% over the among family (based on phenotype) HS\(_P\) strategy (2.75) (Figure 2). HS\(_P\) selection caused a 3.1% increase in DMY BLUP GEBVs over the base population. Hence, the increase in mean GEBV relative to the base population was approximately twice that for the \( A_pW_{GS}HS \) selection compared that of the HS\(_P\) selection.

Predicted response to selection (\( R \))
The relative efficiency of different selection strategies can be compared by estimating the response to selection

### Table 1
Estimated additive genetic (\( \sigma^2_f \)) and pooled error (\( \sigma^2_{\epsilon} \)) variance components with their standard errors (± SE), associated interactions and family mean narrow-sense heritability (\( h^2_n \)) estimated across locations (L) (Aorangi and Ruakura), and across years (Y) for seasonal growth scores in 200 half-sib white clover families.

<table>
<thead>
<tr>
<th>Source</th>
<th>( \sigma^2_f )</th>
<th>( \sigma^2_{l_f} )</th>
<th>( \sigma^2_{l_y} )</th>
<th>( \sigma^2_{l_yl} )</th>
<th>( \sigma^2_{\epsilon} )</th>
<th>( h^2_n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi location</td>
<td>0.10 ± 0.03*</td>
<td>0.10 ± 0.04*</td>
<td>0 ± 0</td>
<td>0 ± 0</td>
<td>1.45 ± 0.05</td>
<td>0.38 ± 0.09</td>
</tr>
</tbody>
</table>

*Significant at \( P<0.05 \)
realised per cycle of selection. The predicted responses to selection for DMY by \( A_pW_{GS}HS \) and \( HS_p \) strategies, relative to the base population from which the selections were made and compared (Table 2). Using the \( A_pW_{GS}HS \) approach, there was a predicted 2.63% increase in DMY BLUPs over the base unsel ected population, which was approximately double that of the \( HS_p \) approach. This reflected the results described above in the comparison of mean DMY BLUP GEBVs (Figure 2).

Mean DMY GEBV of base population = 2.66; Narrow-sense heritability for DMY = 0.38; Selection differential \( (S) \) was the difference between the mean of selected parents and the mean of the base population from which the parents were selected.

**Synthetic population development**

The two groups of 20 plants selected by the \( HS_p \) and \( A_pW_{GS} \) strategies were polycrossed over the 2020/2021 summer period to generate synthetic populations. Balanced bulks of seed derived from each polycross are currently being prepared for empirical evaluation in multi-environment proof-of-concept field trials.

**Discussion**

Long breeding cycles and the inability to access within family variation present considerable challenges to white clover breeding. New strategies that improve breeding efficiency are critical to accomplishing breeding objectives. Genomic selection (GS) has been demonstrated to deliver a higher rate of genetic gain and accelerate cultivar development (Muranty et al., 2015; Lin et al., 2016; Jighly et al., 2019). This study used data from a multi-year and location field trial to develop a genomic prediction model following incorporation of phenotype and genotype data from 200 HS families and their parents, respectively. This supported GEBV prediction of DMY BLUPs in white clover individuals from within the HS families, enabling the comparison of DMY GEBVs under \( HS_p \) and \( A_pW_{GS} \) selection strategies. For DMY, a complex quantitative trait, significant additive genetic variance and G×E interaction effects were estimated in this study. Additionally, this trait showed moderate narrow-sense heritability of 0.38 ± 0.09. The DMY GS prediction model showed moderate predictive ability \( (r = 0.30) \), which was comparable to estimates by Faville et al. (2018) which reported predictive abilities of 0.07 to 0.43 for DMY in perennial ryegrass. In alfalfa (Medicago sativa), predictive accuracy of 0.13 for DMY was reported by Jia et al. (2018) while Annicchiarico et al. (2015) and Li et al. (2015) reported higher accuracies, in the range of 0.30 to 0.66, respectively. Although predictive ability plays an important role in determining the feasibility of GS, it does not describe the amount of genetic gain or response to selection achievable (Herter et al., 2019). For example, studies by Heffner et al. (2010) and Belamkar et al. (2018) in winter wheat derived GS models with low to moderate accuracy \( (0.17 – 0.30) \), but these outperformed both marker-assisted and phenotypic selection methods, respectively, in terms of genetic gain for yield.

Table 2 Predicted response to selection \( (R) \), selection differential \( (S) \) and predicted percentage increase relative to the base population for increasing predicted dry matter yield BLUPs (DMY BLUP GEBV) in white clover using different selection strategies.

<table>
<thead>
<tr>
<th>Selection strategy</th>
<th>Mean DMY BLUP GEBV of selected plants</th>
<th>S</th>
<th>R</th>
<th>Predicted DMY BLUP GEBV post-selection</th>
<th>Percentage increase relative to base population</th>
</tr>
</thead>
<tbody>
<tr>
<td>( HS_p )</td>
<td>2.75</td>
<td>0.09</td>
<td>0.03</td>
<td>2.69</td>
<td>1.13%</td>
</tr>
<tr>
<td>( AF_pW_{GS} )</td>
<td>2.85</td>
<td>0.19</td>
<td>0.07</td>
<td>2.73</td>
<td>2.63%</td>
</tr>
</tbody>
</table>

**Figure 2** Dry matter yield (DMY GEBV) predicted for white clover individuals \( (n = 20) \) selected for synthetic population development under two different selection strategies \( (HSP = \text{half-sibling family selection}; APWGSHS = \text{among and within HS family selection}) \), compared with the base population of individuals \( (n = 800) \).
than conventional among HS family selection (HSₚ) methods. The differential of mean GEBVs relative to the base population was approximately double using the $A_p W_{GS}$ HS strategy compared with HSₚ. This was likely due to increased accuracy of selection and the ability to access the potential 75% additive genetic variation within HS families. It should be noted that HSₚ can be modified to allow for selection within families using phenotypic data, but this requires growing individuals within families to a specific growth stage where the trait under selection can be phenotypically assessed (Vogel and Pedersen 1993; Casler and Brummer 2008). This increases the length of the breeding cycle and incurs additional labour and phenotyping costs. Furthermore, within-family phenotypic evaluations are usually carried out on spaced plants which poorly represent typical mixed sward growing conditions (Hayward and Vivero 1984).

Predicting responses to selection is helpful to breeders as it provides an estimate of the magnitude of achievable genetic gain. The percentage increase in predicted DMY BLUPs in the selected plants relative to the base population was double in $A_p W_{GS}$ HS selection compared to HSₚ. This highlighted the value of using GS to identify the best individuals within HS families. In perennial ryegrass, combining among and within half-sib family selection using GS showed a predicted two-fold increase in genetic gain for DMY yield in a single selection cycle (Faville et al., 2018). This has been further evaluated empirically by Faville et al. (2021). Accordingly, it can be expected that $A_p W_{GS}$ HS, which leverages GEBV information to select individuals with higher predicted DMY, will result in a higher response to selection and consequently deliver greater realised genetic gain over the HSₚ approach in empirical assessments.

Conclusions and practical implications
This study showed that, compared with conventional approaches, integrating GS was predicted to double the increase in DMY BLUPs and responses to selection relative to the base population. Using two strategies, proof-of-concept synthetic populations have now been made and are being prepared for field trials to provide empirical evidence of the influence of GS. From a practical standpoint, long breeding cycles could be significantly shortened by estimation of GEBVs at the seedling stage without the need to phenotype. Additionally, application of GEBVs can increase selection pressure for each cycle. These features highlight the value of integrating available breeding strategies with SNP-based selection and bioinformatic tools to enhance breeding methods, resulting in increased rates of genetic gain.

ACKNOWLEDGEMENTS
The research was supported funded by Pastoral Genomics Plus (grant number PSTG1501), a joint venture co-funded by Dairy New Zealand, Beef + Lamb New Zealand, Dairy Australia, AgResearch Ltd., Barenbrug Agriseeds Ltd. Grasslands Innovation Ltd. and the Ministry of Business, Innovation and Employment, New Zealand. We thank Prue Taylor, Bridget Wise, Peter Moran, Jana Schmidt, Won Hong, Brain Haw, Doug Ryan and others for their assistance in field trial harvests and sample processing.

REFERENCES


Massman JM, Jung H-JG, Bernardo R. 2013. Genomewide Selection versus Marker-assisted Recurrent Selection to Improve Grain Yield and


