Developing genomic selection for dry matter yield in white clover

Grace O. EHOCHE[1,2], Sai Krishna AROJJU[1], Anna LARKING[1], Ruy JAUREGUI[1], Greig COUSINS[2],
Jessica O'CONNOR[1], M. Z. Zulfi JAHUFER[1], and Andrew G. GRIFFITHS[1,*]
[1]*AgResearch, Grasslands Research Centre, Private Bag 11008, Palmerston North, New Zealand*
[2]*PGG 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 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 (HS_P); or where GEBVs were used to select the best individuals within the highest ranked HS families (A_PWF_{GS}). 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 (HS_P), incorporating genomic selection using $A_PW_{GS}HS$ 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; Jahufer *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 (ΔG) 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,

ISSN 2463-2872 (Print) ISSN 2463-2880 (Online) https://doi.org/10.33584/jnzg.2021.83.3502

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 ($A_PW_{GS}HS$). 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_2 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_3 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 25002800 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 (GEBVs; 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 *Ape*KI. Briefly, DNA from the 400 selection candidates, including the control samples and duplicates, were distributed across eleven

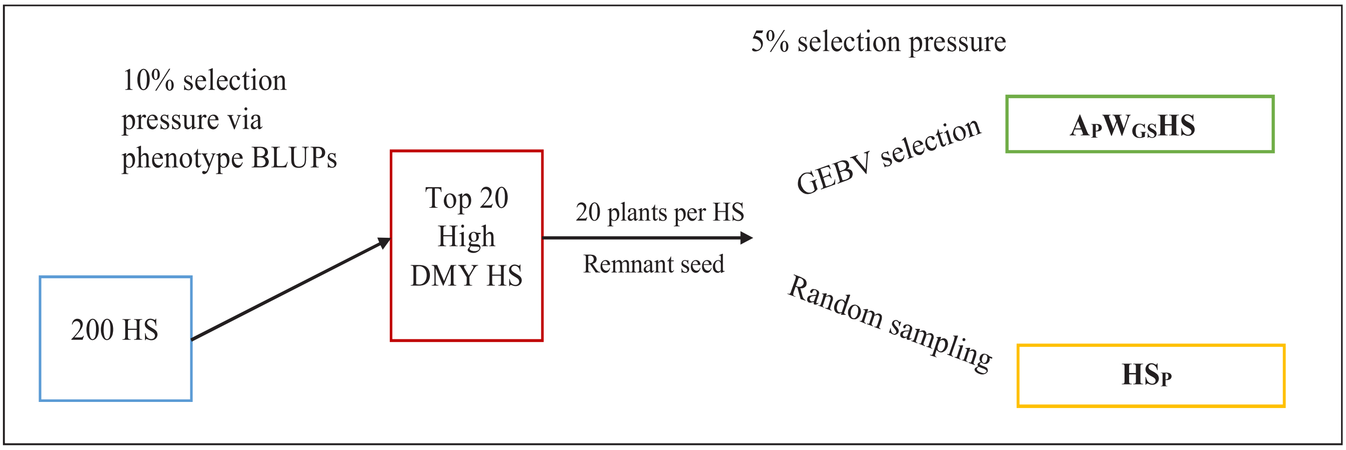

Figure 1 Schematic representation of the selection strategies applied to 200 white clover half-sibling (HS) families. Two synthetic populations based on polycrosses of 20 plants each were generated: HS_P based on phenotype; and A_PWF_{GS} combining phenotype and GEBVs.

96-plex GBS libraries and each library was sequenced on a single lane of an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA) at AgResearch Invermay, New Zealand. Single nucleotide polymorphism (SNP) genotype calling was performed using TASSEL5 (Glaubitz *et al.,* 2014) by aligning to the *Trifolium repens* genome (version five; Griffiths *et al.,* 2019) and filtering for minor allele frequency (MAF) $\geq$ 0.001, missing rate > 50%, read depth >1. These reference and alternative allele counts were exported to KGD, where the resulting SNP set was filtered for Hardy-Weinberg disequilibrium (HWdiseq >-0.05). A genomic relationship matrix (GRM), composed of the selection candidates (highest and lowest ranked HS families) and the parents of the phenotyped training and field trial population, was developed, as described in Dodds *et al.* (2015).

A GS prediction model was derived using the genotype and field trial phenotype data of 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. A KGD-GBLUP prediction model, developed specifically for GBS data (Dodds *et al.,* 2015), was used to generate DMY GEBVs for each selection candidate, as described by Faville *et al.* (2018). This mixed model approach had the GRM included as a variance-covariance matrix (Equation 1).

$$y = 1\mu + Zb + \varepsilon \qquad (1)$$

Where y was the vector of phenotypic records; μ the grand mean; Z the incidence matrix for random effects; b the vector of random marker effects with a normal distribution $b \sim N(0,G\ )$ where G was the genomic relationship matrix (GRM) and the additive genetic variance and ε the vector of random residual effects.

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).

$$R = h_n^2 S \qquad (2)$$

Where: was the response to selection; 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 ± 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 (σ^2_f) and pooled error (σ^2_ε) 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.

Source	σ^2_f	$\sigma^2_{f.l}$	$\sigma^2_{f.y}$	$\sigma^2_{f.y.l}$	σ^2_ε	h^2_n
Multi location	0.10 ± 0.03*	0.10 ± 0.04*	0 ± 0	0 ± 0	1.45 ± 0.05	0.38 ± 0.09

*Significant at P<0.05

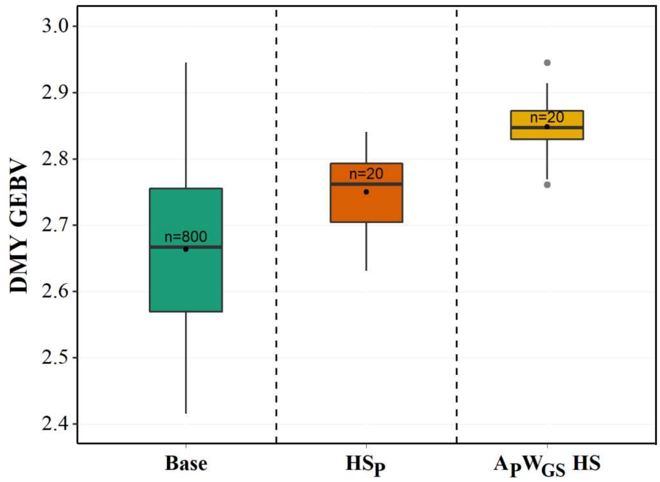

Figure 2 Dry matter yield (DMY GEBV) predicted for white clover individuals (n = 20) selected for synthetic population development under two different selection strategies (HSP = half-sibling family selection; APWGSHS = among and within HS family selection), compared with the base population of individuals (n = 800).

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 unselected 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_PWF_{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_PWF_{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.

Plants selected for crossing using a strategy integrating phenotypic and genomic selection (A_PW_{GS}HS) had a higher predicted mean DMY BLUP

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.

Selection strategy	Mean DMY BLUP GEBV of selected plants	*S*	*R*	Predicted DMY BLUP GEBV post-selection	Percentage increase relative to base population
HS_P	2.75	0.09	0.03	2.69	1.13%
AF_P-WF_{GS}	2.85	0.19	0.07	2.73	2.63%

than conventional among HS family selection (HS_P) methods. The differential of mean GEBVs relative to the base population was approximately double using the $A_PW_{GS}HS$ strategy compared with HS_P. 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 HS_P 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 $A_PW_{GS}HS$ selection compared to HS_P. 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 $A_PW_{GS}HS$, 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 HS_P 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

Anderson CB, Franzmayr BK, Hong SW, Larking AC, van Stijn TC, Tan R, Moraga R, Faville MJ, Griffiths AG. 2018. Protocol: a versatile, inexpensive, high-throughput plant genomic DNA extraction method suitable for genotyping-by-sequencing. *Plant Methods 14*: 75. https://doi.org/10.1186/s13007-018-0336-1

Annicchiarico P, Nazzicari N, Li X, Wei Y, Pecetti L, Brummer EC. 2015. Accuracy of genomic selection for alfalfa biomass yield in different reference populations. *BMC Genomics 16*: 1020. https://doi.org/10.1186/s12864-015-2212-y

Annicchiarico P, Nazzicari N, Pecetti L, Romani M, Russi L. 2019. Pea genomic selection for Italian environments. *BMC Genomics 20*: 603. https://doi.org/10.1186/s12864-019-5920-x

Belamkar V, Guttieri MJ, Hussain W, Jarquín D, El-basyoni I, Poland J, Lorenz AJ, Baenziger PS. 2018. Genomic Selection in Preliminary Yield Trials in a Winter Wheat Breeding Program. *G3: Genes|Genomes|Genetics 8*: 2735. https://doi.org/10.1534/g3.118.200415

Caradus JR, Clifford PTP, Chapman DF, Cousins GR, Williams WM, Miller JE. 1997. Breeding and description of 'Grasslands Sustain', a medium-large-leaved white clover (*Trifolium repens* L.) cultivar. *New Zealand Journal of Agricultural Research 40*: 1-7. https://doi.org/10.1080/00288233.1997.9513224

Casler MD, Brummer EC. 2008. Theoretical Expected Genetic Gains for Among-and-Within-Family Selection Methods in Perennial Forage Crops. *Crop Science 48*: 890-902. https://doi.org/10.2135/cropsci2007.09.0499

Dodds KG, McEwan JC, Brauning R, Anderson RM, van Stijn TC, Kristjánsson T, Clarke SM. 2015. Construction of relatedness matrices using genotyping-by-sequencing data. *BMC Genomics 16*: 1047. https://doi.org/10.1186/s12864-015-2252-3

Erbe M, Pimentel ECG, Sharifi AR, Simianer H. 2010. Assessment of cross-validation strategies for genomic prediction in cattle. Gesellschaft für Tierzuchtwissenschaften e. V, Giessen. grassland.org.nz/publications/nzgrassland_publication_684.pdf

Esfandyari H, Fè D, Tessema BB, Janss LL, Jensen J. 2020. Effects of Different Strategies for Exploiting Genomic Selection in Perennial Ryegrass Breeding Programs. *G3 Genes|Genomes|Genetics 10*: 3783-3795. https://doi.org/10.1534/g3.120.401382

Faville MJ, Ganesh S, Cao M, Jahufer MZZ, Bilton TP, Easton HS, Ryan DL, Trethewey JAK, Rolston MP, Griffiths AG, Moraga R, Flay C, Schmidt J, Tan R, Barrett BA. 2018. Predictive ability of genomic selection models in a multi-population perennial ryegrass training set using genotyping-by-sequencing. *Theoretical and applied genetics 131*: 703-720. https://doi.org/10.1007/s00122-017-3030-1

Faville MJ, Schmidt J, Trolove M. Moran P, Hong W, Cao M, Ganesh S, George R, Barrett BA. 2021. Empirical assessment of a genomic breeding strategy in perennial ryegrass. *Journal of New Zealand Grasslands 83*. In press.

Galwey NW. 2006. *Introduction to Mixed Modelling: Beyond Regression and Analysis of Variance*. John Wiley & Sons, West Sussex, England. 376p. ISBN-10: 0470014962

Gibson PB, Cope WA. 1985. White Clover. In: Taylor NL Ed. *Clover Science and Technology*. Madison, Wisconsin, USA: American Society of Agronomy, pp 471-488. ISBN 10: 0891180834

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler E. 2014. TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PLoS One 9*. https://doi.org/10.1371/journal.pone.0090346

Griffiths AG, Moraga R, Tausen M, Gupta V, Bilton TP, Campbell MA, Ashby R, Nagy I, Khan A, Larking A, Anderson C, Franzmayr B, Hancock K, Scott A, Ellison NW, Cox MP, Asp T, Mailund T, Schierup MH, Andersen SU. 2019. Breaking Free: The Genomics of Allopolyploidy-Facilitated Niche Expansion in White Clover. *The Plant Cell 31*: 1466. http://doi.org/10.1105/TPC.18.00606

Hallauer AR, Filho JBM. 1981. *Quantitative Genetics in Maize Breeding*: Iowa State University Press, Ames, Iowa, USA. P.468. ISBN:08-138-15207

Harville DA. 1977. Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *Journal of the American Statistical Association 72*: 320-338. https://doi.org/10.2307/2286796

Hayes BJ, Cogan NOI, Pembleton LW, Goddard ME, Wang J, Spangenberg GC, Forster JW. 2013. Prospects for genomic selection in forage plant species. *Plant Breeding 132*: 133. https://doi.org/10.1111/pbr.12037

Hayward MD, Vivero JL. 1984. Selection for yield in *Lolium perenne.* II. performance of spaced plant selections under competitive conditions. *Euphytica 33*: 787-800. https://doi.org/10.1007/BF00021905

Heffner EL, Lorenz AJ, Jannink JL, Sorrells ME. 2010. Plant breeding with Genomic selection: Gain per unit time and cost. *Crop Science 50*. https://doi.org/10.2135/cropsci2009.11.0662

Herter CP, Ebmeyer E, Kollers S, Korzun V, Miedaner T. 2019. An experimental approach for estimating the genomic selection advantage for Fusarium head blight and *Septoria tritici* blotch in winter wheat. *Theoretical and Applied Genetics 132*: 2425-2437. https://doi.org/10.1007/s00122-019-03364-7

Hoyos-Villegas V, O'Connor JR, Heslop AD, Hilditch A, Jahufer MZZ, Barrett BA. 2019. Rate of Genetic Gain for Persistence to Grazing and Dry Matter Yield in White Clover across 90 Years of Cultivar Development. *Crop Science 59*: 537-552. https://doi.org/10.2135/cropsci2018.07.0471

Jahufer MZZ, Cooper M, Ayres JF, Bray RA. 2002. Identification of research to improve the efficiency of breeding strategies for white clover in Australia - A review. *Australian Journal of Agricultural Research 53*: 239-257. https://doi.org/10.1071/AR01110

Jahufer MZZ, Luo D. 2018. DeltaGen: A Comprehensive Decision Support Tool for Plant Breeders. *Crop Science 58*: 1118-1131. https://doi.org/10.2135/cropsci2017.07.0456

Jia C, Zhao F, Wang X, Han J, Zhao H, Liu G, Wang Z. 2018. Genomic Prediction for 25 Agronomic and Quality Traits in Alfalfa (*Medicago sativa*). *Frontiers in plant science 9*: 1220-1220. https://doi.org/10.3389/fpls.2018.01220

Jighly A, Lin Z, Pembleton LW, Cogan NOI, Spangenberg GC, Hayes BJ, Daetwyler HD. 2019. Boosting Genetic Gain in Allogamous Crops via Speed Breeding and Genomic Selection. *Frontiers in Plant Science 10*. https://doi.org/10.3389/fpls.2019.01364

Li X, Wei Y, Acharya A, Hansen JL, Crawford JL, Viands DR, Michaud R, Claessens A, Brummer EC. 2015. Genomic Prediction of Biomass Yield in Two Selection Cycles of a Tetraploid Alfalfa Breeding Population. *The Plant Genome 8*. https://doi.org/10.3835/plantgenome2014.12.0090

Lin Z, Cogan NOI, Pembleton LW, Spangenberg GC, Forster JW, Hayes BJ, Daetwyler HD. 2016. Genetic Gain and Inbreeding from Genomic Selection in a Simulated Commercial Breeding Program for Perennial Ryegrass. *The Plant Genome 9*. https://doi.org/10.3835/plantgenome2015.06.0046

Lush JL. 1937. *Animal Breeding Plans*: Iowa State College Press.

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

Stover-quality Traits for Cellulosic Ethanol in Maize. *Crop Science 53*: 58-66. https://doi.org/10.2135/cropsci2012.02.0112

Muranty H, Troggio M, Sadok IB, Rifaï MA, Auwerkerken A, Banchi E, Velasco R, Stevanato P, van de Weg WE, Di Guardo M, Kumar S, Laurens F, Bink MCAM. 2015. Accuracy and responses of genomic selection on key traits in apple breeding. *Horticulture Research 2*: 15060. https://doi.org/10.1038/hortres.2015.60

Patterson HD, Thompson R. 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika 58*: 545-554. https://doi.org/10.1093/biomet/58.3.545

Vogel KP, Pedersen JF. 1993. Breeding systems for cross-pollinated perennial grasses. *Plant Breeding Reviews 11*: 251-274. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1948andcontext=agronomyfacpub

Woodfield DR, Caradus JR. 1996. Factors affecting white clover persistence in New Zealand pastures. *Proceedings of the New Zealand Grassland Association* 58: 229-235. https://doi.org/10.33584/jnzg.1996.58.2196