Looking at the past to prepare ahead: adapting Red Clover for future New Zealand climates

Angus D HESLOP*[1,2], Andrew G GRIFFITHS[3], Rainer W HOFMANN[1], M Zulfi Z JAHUFER[4], John L FORD[5], Anna C LARKING[3], Rachael ASHBY[6], Charles A HEFER[2] and Jessica O'CONNOR[6]

[1]*Faculty of Agriculture and Life Sciences, Lincoln University,7647, New Zealand*
[2]*BSI - AgResearch Group Lincoln, PB 4749, Christchurch, 8140 New Zealand*
[3]*BSI - AgResearch Group Grasslands, PB 11008, Palmerston North, 4442, New Zealand*
[4]*School of Agriculture and Food Sustainability, University of Queensland, Australia*
[5]*PGG Wrightson Seeds Ltd, C/ PB 11008, Palmerston North, 4442, New Zealand*
[6]*BSI - AgResearch Group Invermay, PO 50034, 9053, Mosgiel, New Zealand*
*Corresponding author: angus.heslop@agresearch.co.nz

Abstract

Developing adaptive, resilient pastural cultivars is crucial for maintaining New Zealand's (NZ) highly productive farming sector. With growing challenges from climate change, including heat and extreme rainfall, identifying genetic material that provides resilience is vital. Wild populations, shaped by their local environmental pressures and isolation, hold unique gene makeups that can help develop climate-adaptive cultivars. In this study, we examined the genetic response of 92 internationally geographically diverse red clover populations to their source bioclimatic environments using partial redundancy analysis. The aim was to identify bioclimatic variables driving environmental adaptation and the resulting DNA variants (outlier single nucleotide polymorphisms (SNPs)) associated with adaptation. Calculation of adaptive indices and genomic offset values enabled us to predict the suitability of these populations to future NZ environments. We found that Annual mean diurnal range, Isothermality (variance in daily temperature relative to annual variation), Mean temperature of the wettest quarter, and Precipitation seasonality underpinned adaptive genetic variation. Forty-two outlier SNPs strongly associated with key bioclimatic variables show potential as markers for climate-resilient breeding. Mapping adaptive indices and genomic offset values to NZ's current and predicted future climates showed the genetic diversity captured in these germplasm populations could help develop future-proofed adaptive cultivars.

Keywords: adaptation, redundancy analysis, genomic offset/maladaptation, genotype- environmental interaction, landscape genomics

Introduction

Red clover (*Trifolium pratense* L.) is an important legume that provides both quality forage and high summer production and has the potential to strengthen traditional temperate pastural mixes (Taylor and Quesenberry 1996). With climate change driving increased temperature and precipitation extremes, developing adaptive and resilient red clover cultivars is critical. As cultivar development takes 10 to 20 years, preparing for future extreme and erratic environmental conditions requires forward thinking now. To achieve this, capturing genetic diversity capable of providing adaptive traits in breeding programs is crucial (Taylor and Quesenberry 1996). One approach is to evaluate and harness the genetic diversity and adaptive potential of unexplored germplasm collected from diverse regions. Populations collected from the wild reflect their ecological past as they have been shaped by their environment which enriches for adaptive traits underpinning success in that region. Landscape genomics, which combines genetic and eco-geographic data, helps characterise key selective pressures driving adaptability within a species, and identifies adaptive genetic variation within these populations and the associated genes that may be harnessed to produce new adaptive cultivars (Capblancq and Forester 2021).

Redundancy analysis (RDA), a common landscape genomics method, explores relationships between response variables (DNA variation such as single nucleotide polymorphisms (SNPs); allele frequencies) and explanatory variables (environmental) (Capblancq and Forester 2021). This can identify SNPs strongly associated with environmental predictors, highlighting associations between genetic variation and environmental variables (Legendre and Legendre 2012; Capblancq and Forester 2021). Recent use of RDA in red clover identified symmetry between key bioclimatic variables and population structure and showed a relationship between plant phenotype in multi-year field trials and the underlying genetic drivers of local adaptation including 123 associated SNPs, some drought related (Heslop et al. 2025).

An extension of RDA is partial redundancy analysis (pRDA), a modified version of RDA that accounts for

confounding factors such as population structure and geographical variables by removing their effect before performing RDA (Capblancq and Forester 2021). This facilitates a more nuanced understanding of how specific environmental variables influence genetic variation independent of these confounding factors. Research applying pRDA in forages is limited as the focus has been ecology-based studies such as detecting how genetic variation across plant populations is explained by environmental variables to predict adaptability within plant species for future environments. This approach can be used to estimate genomic offset, the predicted mismatch or maladaptation, between a population's current genetic makeup and the genetic composition required for adaptation to future environments. This provides a route to developing conservation management strategies as shown in European Beech (*Fagus sylvatica*) and Lodgepole Pine (*Pinus contorta*), and Channel Island Oak (*Quercus tomentella*) (Capblancq et al. 2020; Capblancq and Forester 2021; Mead et al., 2024) and may be of value in plant breeding.

The objective of this study was to build upon previous genetic diversity work and an initial RDA in a red clover diversity panel (Heslop et al. 2025) by applying pRDA for a more nuanced analysis to identify bioclimatic variables key to adaptation and facilitate calculation of an adaptive index and genomic offset for current and future NZ environments. This approach enables us to harness the history of wild red clover populations and identify adaptive alleles for increased cultivar resilience in future environments as a tool for selecting populations for breeding programs. The study was comprised of three stages: i) identify the bioclimatic variables key to explaining the genetic variation present within populations; ii) identify SNPs associated with the key bioclimatic variables that could be used as markers implicated in adaptive responses to environments; and iii) calculate and project adaptive gradients and genomic offset on current and future NZ environments.

Materials and Methods

Diverse germplasm panel

As described previously (Heslop et al. 2025), an international eco-geographically diverse panel of 92 novel red clover populations, with an increased weighting for material from low rainfall environments, was sourced from the Margot Forde Genebank (MFG) at AgResearch Grasslands, Palmerston North, NZ.

DNA isolation and genotyping of bulked samples

Genotype and allele frequency data for each of the 92 populations were generated in a previous publication (Heslop et al. 2025) and were derived from pools of 30 individuals per population. Briefly, DNA was extracted from seedlings as described (Anderson et al. 2018) and genotyping-by-sequencing (GBS) libraries were generated in 96-plex using a protocol adapted from Griffiths et al. (2019) with modifications such as digestion with *Pst*I restriction enzyme and inclusion of four technical replicates for each population as described (Faville et al., 2020). GBS libraries were single end sequenced (100 bp reads) using two lanes (~44 Gb data) on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA) at AgResearch Invermay, NZ. Sequence reads were quality-checked, demultiplexed and mapped to the ARS RC 1.1 (GCA_020283565.1) red clover reference genome (Bickhart et al. 2022). SNPs were detected and filtered accordingly (Heslop et al. 2025). Allele frequencies were combined across technical replicates and were an estimate of the allele frequency at a SNP for the pooled sample. These values were extracted, and alternative allele frequencies (AAF) were calculated relative to the reference alleles as described (Heslop et al. 2025).

Extraction of bioclimatic variables and future climate modelling

Bioclimatic information from the regions where the 92 populations were sampled was accessed as described (Heslop et al. 2025). Briefly, 19 bioclimatic variables (Bio1 to Bio19) extracted at 2.5 arcminutes (~4.5 km at the Equator) spatial resolution from the WorldClim database using collection site GPS coordinates from sample passport data obtained from the MFG database. In the current study, NZ climate variables were extracted using topographic shapefiles obtained from the Land Information New Zealand database to filter variables from within geographic borders using the "terra" R package. WorldClim future variables were extracted from the Worldclim database at 2.5-arcminute resolution using the HadGEM3-GC31-LL model after comparison of the twelve models available as part of the Coupled Model Intercomparison Project Phase 6 (CMIP6) framework. CMIP6 uses shared socioeconomic pathways (SSPs) which represent different levels of future emissions, climate policies and socioeconomic development. Of the five different available SSPs, we used: i) the SSP2-4.5 scenario, a "Middle of the road" scenario, with global warming increasing by ~ 2.7˚C by 2100; and ii) the SSP5-8.5 scenario, a "worst-case" scenario, with global warming increasing by ~ 4.5˚C by 2100. Future bioclimatic variables for each population for three timeframes "present condition", "2021-40" and "2041-2060" were extracted using GPS coordinates of their collection sites. For the NZ climatic variables, a topographic shapefile was used to filter variables for the sites of interest.

Redundancy analysis

Redundancy analysis (RDA) was performed using the "*Vegan*" package in RStudio, with GBS-derived population AAF data as response variables and bioclimatic variables as the predictor variables (Capblancq and Forester 2021). In contrast to the previous study (Heslop et al. 2025), no phenotype data were used as a variable as the current study focused solely on the environmental influence on genetic variation. To maximize the total amount of genetic variance explained by the set of bioclimatic variables, and to reduce collinearity and minimize overfitting, forward model building was implemented using the "*ordiR2step*" function from the "*Vegan*" package as described (Capblancq and Forester 2021). In short, a global model that included all the variables was executed and significance tested. If the global model was significant, forward selection was implemented with a null or empty model run where multivariate genetic variation was explained by a fixed intercept. Each variable was then added individually using two stopping criteria to prevent overestimation of the explained variance. These were: i) a variable significance of $p < 0.01$ using 1,000 permutations; and ii) the adjusted R^2 of the global model. If a variable met either stopping threshold (greater *p*-value or a decrease in R^2) the variable was rejected, otherwise it was retained and added to the final model. For all analyses the variables were standardized to ensure compatibility (Legendre and Legendre 2012).

Partial redundancy analysis (pRDA) was then performed as described (Legendre and Legendre 2012; Capblancq and Forester 2021) using the AAF data as the response variable to investigate the influence and proportion of variance that environmental (bioclimatic variables), geographic (GPS coordinates) and demographic (neutral genetic structure) explanatory variables had on the distribution of genetic variation across the species range. We used bioclimatic variables ('climate'), two proxies of neutral genetic structure (population scores along PC1 and PC2 axes of an intergenic PCA; 'structure') and population GPS information (latitude and longitude) to characterize geographical variation ('geography') as predictor variables. Intergenic SNPs were identified by matching SNPs from our panel of populations with the annotated database of the ARS RC 1.1 (GCA_020283565.1) red clover reference genome (Bickhart et al. 2022). SNPs were classified as genic or intergenic if located inside or outside genes, respectively. Four pRDA models were run and total inertia (variance), constrained inertia, proportion of variance explained by constraints, model R^2 and *p*-values were extracted for each model. The four models were:

i) Full model: F ~ climate + structure + geography
ii) Pure climate: F ~ climate | (structure + geography)
iii) Pure structure: F ~ structure | (climate + geography)
iv) Pure geography: F ~ geography | (climate + structure)

Candidate adaptive markers were identified by pRDA using the "*Vegan*" and "*rdadapt*" packages in RStudio, with the population AAF data as the response and the bioclimatic variables, geographical and the first two PCs loadings of intergenic SNPs as the predictors. The procedure described by Capblancq et al. (2018) was used where outliers are identified based on their extremeness along a distribution of Mahalanobis distances estimated between each locus and the centre of the RDA space using a certain number of axes (K). Additionally, to determine if by accounting for population structure the number of outliers identified was conservative, an RDA was run where allele frequencies were the response, and the bioclimatic variables were the predictors. For both pRDA and RDA a *p*-value threshold of $p < 0.01$ was enforced using the False Discovery Rate approach within the "*rdadapt*" function. Outliers identified from both pRDA and simple RDA analysis were compared and outliers in common were used. Relationships between outliers and the underlying climatic variables were visualised on an RDA biplot.

Adaptive index and genomic offset

An RDA was performed with just the allele frequencies of the outliers identified as the response and the bioclimatic variables as the predictors. The scores of the climatic variables along the RDA axes were used to calculate a genetic-based index of adaptation for individual environmental pixels for a landscape image, in this case NZ. The climatic variable scores used were estimated based on how each bioclimatic variable affects adaptive genetic variation along the RDA axes (Capblancq and Forester 2021). For each RDA axis of interest this index is estimated independently using the formula:

$$\text{Adaptive Index} = \sum_{i=1}^{n} a_i b_i$$

Where: *a* is the climatic variable score (loading) along the RDA axis; *b* is the standardized value for this variable at the focal pixel; and *i* refers to one of the *n* different variables used in the RDA model (Capblancq and Forester 2021).

The adaptive index estimates the genetic similarities or differences of all pixels across a landscape image based on the environmental predictor values at each location. When mapped, it visualizes adaptive gradients

Table 1 The four bioclimatic variables sourced from the WorldClim database (worldclim.org).

Bioclimatic Variable	Units	Interpretation
Annual Mean Diurnal Range (Bio2)	°C	The mean of monthly temperatures which shows temperature fluctuations.
Isothermality (Bio3)	%	The level of variance in daily temperature relative to annual variation
Mean Temperature of Wettest Quarter (Bio8)	°C	The mean temperatures during the wettest season.
Precipitation Seasonality (Bio15)	%	The measure of the variation in monthly precipitation totals over the course of the year

across the species' range. The estimation of adaptive indexes can then be used to predict future maladaptation/genomic offset values. Genomic offset was calculated from the difference of adaptive indexes calculated for both current and future predicted environmental conditions. This value quantifies the degree of genetic change required for a population to continue to be adapted as the environment changes. Higher genomic offset values indicate greater potential maladaptation, suggesting that a population may struggle to persist without genetic adaptation or migration. When mapped, it visualizes genomic offset gradients across the species' range indicating adaptability.

Results and Discussion

Variable selection and model building

As outlined in Heslop et al. (2025), GBS of pooled samples for each of the 92 red clover populations identified 12,168 SNPs which reduced to 4,509 SNPs after filtering. Of these SNPs, 1,289 were identified in the current study as intergenic, therefore, considered neutral markers for determining neutral population structure. A principal component analysis was completed with these markers and the loadings of the first two axis (PC1 and PC2) were retained and included in subsequent models as a proxy for neutral genetic structure. Using forward selection, a superior model was fitted that retained four bioclimatic variables: Annual mean diurnal range (Bio2); Isothermality (Bio3); Mean temperature of driest quarter (Bio8); and Precipitation seasonality (Bio15) (Table 1). Isothermality was highly correlated with longitude and may reflect the geographical spread of the red clover collection sites which ranged from the Iberian Peninsula through Western Europe to Central Asia (Tajikistan) (Heslop et al. 2025). Consequently, longitude was removed and only latitude used for building the models. The resulting full model was:

$$pRDAfull <- Bio2+Bio3+Bio8+Bio15 +PC1+PC2+latitude$$

Where the bioclimatic variables (Bio2, Bio3, Bio8, Bio15) were the explanatory variables of interest, and PC1, PC2 and latitude were included as covariates to control for neutral genetic structure and geographic influence.

In summary, forward selection identified the most relevant predictor variables explaining variation in the response variables while ensuring only the most impactful variables were included in the model. This increases model efficiency, reduces overfitting and avoids the inclusion of collinear or redundant variables, making the results more interpretable and robust (Blanchet et al. 2008; Capblancq and Forester 2021). This is an important step in building an RDA model, as poor selection of variables can potentially bias biological conclusions and downstream analyses (Capblancq and Forester 2021).

Partitioning variance: unravelling factors influencing genetic variation

The pRDA analysis identified the contributions of climate, neutral population structure and geography in explaining the genetic variation amongst the red clover populations. Combined, these three factors accounted for 47.7% of the total variance present in the genetic data, leaving 52.3% of the variance unexplained (Table 2). Similar levels of unexplained variance were found in studies investigating drivers of adaptive genetic variation and predicting future response to climate change in European Beech, Lodgepole Pine and a range of Eggplant species (Capblancq et al. 2020; Capblancq and Forester 2021; Omondi et al. 2024). The level of unexplained variance is possibly a combination of unmeasured environmental variables, genetic drift, sampling bias (sampling sites not being representative of the entire range of genetic diversity), gene by environment interactions or model limitations. However, by investigating the variance partitioning, we identified the factors influencing the 47.7% of explained genetic variance. These were climate, structure and geography which had highly significant ($p < 0.001$) contributions to total explained variance (Table 2). The climate and geography models each explained 19.7% and 17.3% of total variance, respectively, while the greatest portion of variance was explained by the neutral

Table 2 Variance portioning from pRDA (partial redundancy analysis) showing influence of climate, neutral population and geography on the genetic variation of the red clover populations.

Partial RDA models	Inertia	R^2 (adjusted)	p- value (>F)	Proportion of explainable variance	Proportion of total variance
Full model: F ~ climate + structure + geography	0.045	0.222	0.001 ***	0.477	0.477
Pure climate: F ~ climate I (structure + geography)	0.018	0.062	0.001 ***	0.414	0.197
Pure structure: F ~ structure I (climate + geography)	0.028	0.069	0.001 ***	0.631	0.301
Pure geography: F ~ geography I (climate + structure)	0.016	0.033	0.001 ***	0.363	0.173
Total unexplained	0.049	-	-	0.523	0.523
Total inertia	0.093	-	-	-	1.149

Significance levels of model (*p < 0.05; **p < 0.01; ***p < 0.001).

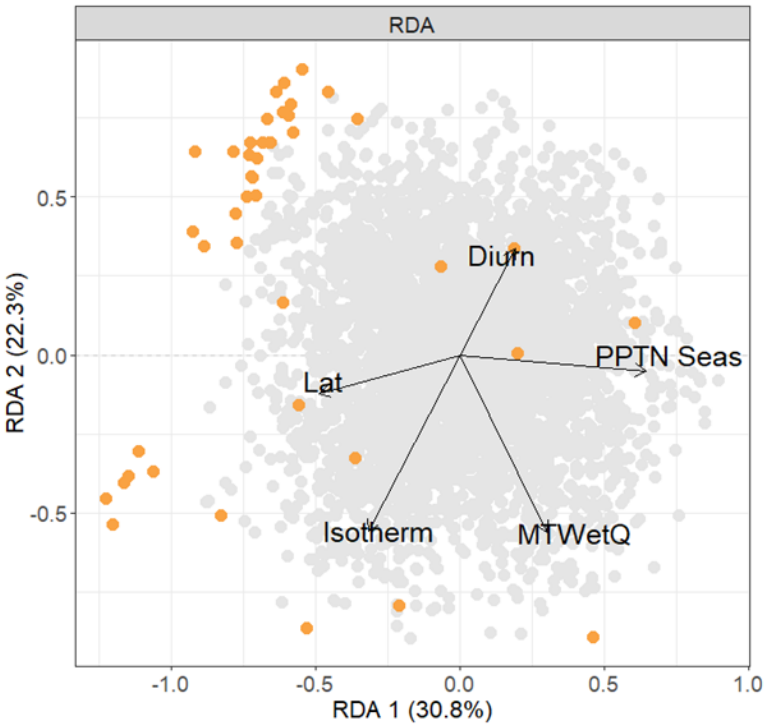

Figure 1 A Redundancy Analysis (RDA) biplot of all 4,509 SNPs and their association with bioclimatic variables which were: Annual mean diurnal range (Diurn); Isothermality (Isotherm); Mean temperature of the wettest quarter (MTWetQ); Precipitation seasonality (PPTN Seas); and Latitude (Lat) identified through redundancy analysis along RDA1 and RDA2. The 42 outlier SNPs identified as driving adaptation through strong association to bioclimatic variables are highlighted. Non-outlier SNPs are the remaining SNPs without strong association to bioclimatic variables.

population structure model (30.1%). However, the sum of pure effects (climate, structure, and geography) contributed 67.1% of total variance which exceeded the total variance of 47.7% explained by the full model (Table 2). This suggests a large amount of explained variance is shared among variables. Such an overlap highlights the interdependence of climate, structure, and geography variables, implying that their effects on the genetic data were not entirely independent, thus influencing each other in ways that cannot always be separated (Capblancq and Forester 2021).

Identifying outlier SNPs through genotype-environmental associations

From RDA and pRDA analysis, 42 adaptive outlier SNPs in common were identified using a significant

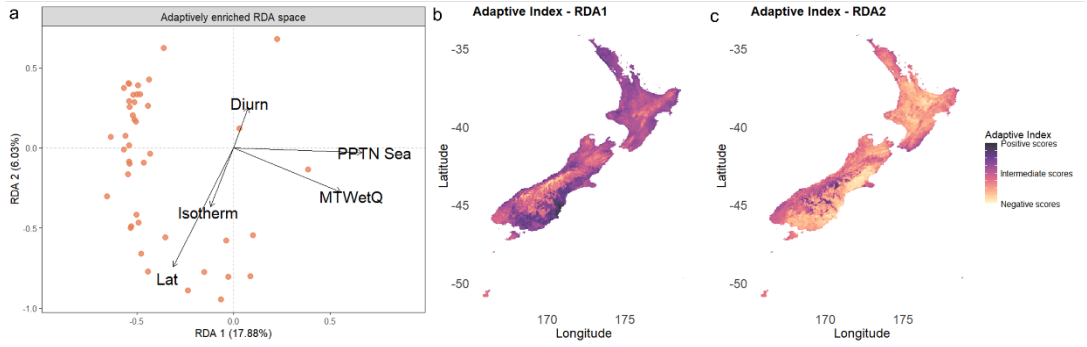

Figure 2 a) Adaptive RDA showing associations of the 42 outlier SNPs with bioclimatic variables (Annual mean diurnal range, Diurn; Isothermality, Isotherm; Mean temperature of the wettest quarter, MTWetQ; Precipitation seasonality, PPTN Seas; and Latitude, Lat) identified through redundancy analysis along RDA1 and RDA2. b) Mapped adaptive index scores showing genetic adaptation to current NZ conditions b) RDA1 and C) RDA2 and d) RDA2.

level threshold of $p < 0.01$. Of these SNPs, a large group had a negative selection or reduced adaptation to mean temperature of the wettest quarter (Figure 1; top left corner) which indicates these SNPs are associated with populations adapted to cooler, wet seasons. Populations at higher latitudes with cooler, more seasonal climates had distinct adaptive SNPs (Figure 1; bottom left corner).

Projecting adaptive gradients across space to predict local maladaptation

Using the identified 42 outlier SNPs only, a new RDA was performed that showed the correlation between genetic markers and different bioclimatic variables and amongst bioclimatic variables. The first two axes exhibited most of the adaptive genetic variance among the populations (Figure 2a). RDA1 (17.9%) was correlated to Precipitation seasonality and Mean temperature wettest quarter, while RDA2 (6.0%) was correlated to Isothermality, Mean diurnal range and Latitude. The distribution of these SNPs along the axes indicates how the populations are structured based on the environmental gradients. When mapped to current climatic conditions in NZ (Figure 2b, c) it visualizes how well populations with these outlier SNP alleles are genetically adapted across the landscape. In short, a positive (purple) or negative (yellow) score indicates regions where populations genetically associated with high or low scores, respectively, along a particular axis are best suited. This means if red clover populations that possess outlier SNPs associated to each RDA axis were selected for, they will likely be better adaptive to conditions in areas that have positive scores (purple) and less adaptive to areas with negative scores (yellow). For example, populations with high SNPs scores for Precipitation seasonality (Figure 2a RDA1; righthand side) were suited to much of NZ coastal regions (Figure 2b), particularly in the South, but less so Central Otago and North Canterbury/Marlborough which were more suited to populations with low SNP scores (Figure 2a RDA1; lefthand side). Similarly, populations with outlier SNPs negatively aligned with RDA2 (Figure 2a; bottom left), driven by Isothermality, Mean diurnal range and Latitude, are suited for much of NZ (Figure 2c). These may be target SNPs to incorporate into markers for breeding programmes.

Genomic offset quantifies the amount of genetic change required for a population to remain adapted to a region by using the difference between current and future climatic conditions (Capblancq et al. 2020; Mead et al. 2024; Jia et al. 2019). High genomic offset values (yellow/green) indicate areas where the current genetic makeup of the populations based on the outlier SNPs is poorly matched to an environment (maladaptation) and these population would require significant evolutionary change to perform well (Figure 3). By contrast, low genomic offset values (purple) indicate regions better matched with the current genetic make-up of the populations. Mapping genomic offset to NZ future conditions (Figure 3) identified regions to which the genetics of the 92 populations were already aligned or require further change. Genomic offset values were mapped for the following scenarios: i) "Middle of the road" 2021-2040, ii) "Middle of the road" 2041-2060, iii) "Worst-case" 2021-2041 and iv) "Worst-case" 2041-2060.

Across all four scenarios, low genomic offset values were shown (Figure 3), which indicates that alleles in the 92 diverse red clover populations align well with current and future environmental conditions and that these populations have genetic variation that will increase adaptation and lower the risk of maladaptation.

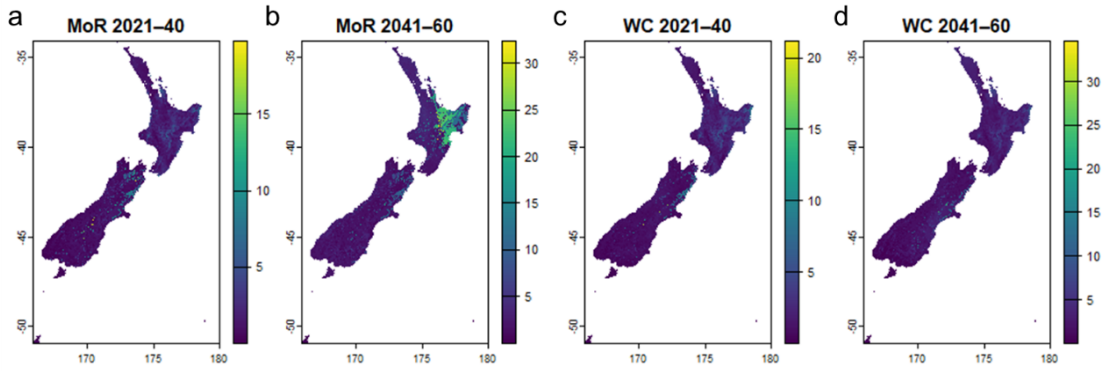

Figure 3 Mean predicted genomic offset for the red clover populations in future NZ environmental conditions in "MoR" (Middle of the road ~2.7°C by 2100) for a) 2021-2040 and b) 2041-2060, and "WC" (Worst-case ~4.5°C by 2100) for c) 2021-2041 and d) 2041-2060.

For both temperature regimes in the 2021-2041 period (Figure 3a, c), genomic offset values showed similar patterns, indicating these international populations having pre-existing genetic variation to adapt to these conditions. However, for the period between 2041-2060 the "Middle of the road" scenario showed greater genomic offset values, especially in the North-Eastern North Island compared to the "Worst-case" scenario. This could be explained by the differences in models, as the "Worst case" scenario predicts a rapid and extreme change to key bioclimatic variables which could push populations past a threshold where selection stabilizes. By contrast, the gradual specific changes to key bioclimatic variables under the "Middle of the road" creates complex shifts in the dynamics of selection pressures.

Fitting genomic offset models using only the 42 outlier SNPs instead of all 4,509 SNP frequencies had little impact on genomic offset predictions for future NZ environments (Supplementary Figure 1). This suggests the outlier SNPs identified are sufficient for estimating genomic offset and may help predict genomic maladaptation in red clover breeding programmes. Genomic offset prediction as a tool has been implemented in a range of scenarios. For example, the method has been used for conservation efforts such as with California Channel Island Oak (*Q. tomentella*) where genomic offset highlighted the need to intervene through assisted dispersal of new genetics to introduce adaptive alleles to existing populations predicted to be poorly adapted to future climate (Mead et al. 2024). Meanwhile, our genomic offset predictions showed the diverse red clover populations have sufficient genetic diversity present suitable for adapting to predicted NZ climates and will be a significant breeding resource, particularly if combined with the adaptive SNP marker data in the breeding process.

For pRDA to be a useful tool in breeding programs, ensuring there is a route to utilising the information is critical. Key takeaways from this study show the selection of populations with a low genomic offset, high adaptive index and that contain the 42 outlier SNPs associated with climate resilience for inclusion into breeding programs would ensure better adaptability to NZ's changing climate. Districts identified as having a high genomic offset for red clover, such as North-Eastern North Island under a "Middle of the Road" climate model, would require specific breeding programs or multi-site testing in current environments that reflect this future environment to identify populations that perform well there. An additional interpretation of the pRDA results highlights the importance of maintaining genetic diversity in breeding pools to retain genetic plasticity for changing climates. Future work could look at using genomic selection models to predict which would be the best populations based on future climates to include in breeding pools. Additional screening of germplasm populations for the SNPs identified or associations to key bioclimatic variables could increase the resilience of these breeding programs.

Practical implications

The results of this study demonstrate that within an eco-geographically and genetically diverse range of 92 international red clover populations, there is sufficient genomic plasticity to ensure adaptation to future NZ environments in the face of predicted climate change. Identifying and harnessing the identified adaptive alleles in breeding programmes will facilitate development of future-proofed red clover to support sustainable pastural production into the future.

ACKNOWLEDGEMENTS

This research was supported through funding from the AgResearch Impact Prize 2018 (PRJ0036572) and the Grasslands Innovation Research Programme - AgResearch Clover and Herb Breeding Continuum (I22000-PRJ0449754). AH also had support from the Kathleen Spragg Agricultural Fellowship (G-202201-00511).

REFERENCES

Anderson CB, Franzmayr BK, Hong SW, Larking AC, van Stijn TC, Tan R, Moraga R,R, Faville M, Griffiths A. 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

Bickhart DM, Koch LM, Smith TP, Riday H, Sullivan M. 2022. Chromosome-scale assembly of the highly heterozygous genome of red clover (*Trifolium pratense* L.), an allogamous forage crop species. *Gigabyte* 2022, gigabyte42. https://doi.org/10.46471/gigabyte.42

Blanche, FG, Legendre P, Borcard D. 2008. Forward selection of explanatory variables. *Ecology 89*: 2623-2632. https://doi.org/10.1890/07-0986.1

Capblancq T, Luu K, Blum MGB, Bazin E, Umr L. 2018. Evaluation of redundancy analysis to identify signatures of local adaptation. *Molecular Ecology Resources 18*: 1223-1233. https://doi.org/10.1101/258988

Capblancq, Morin X, Gueguen M, Renaud J, Lobreaux S, Bazin E. 2020. Climate-associated genetic variation in *Fagus sylvatica* and potential responses to climate change in the French Alps. *Journal of Evolutionary Biology 33*: 783-796. https://doi.org/10.1111/jeb.13610

Capblancq T, Forester BR. 2021. Redundancy analysis: A Swiss Army Knife for landscape genomics. *Methods Ecology Evolution 12*: 2298-2309. https://doi.org/10.1111/2041- 210X.13722

Faville MJ, Griffiths AG, Baten A, Cao M, Ashby R, Ghamkhar K, Hong W, Larking A, Williamson M, Webber Z. 2020. Genomic assessment of white clover and perennial ryegrass genetic resources. *Journal of New Zealand Grasslands 82*: 27-34. https://doi.org/10.33584/jnzg.2020.82.437

Griffiths AG, Moraga R, Tausen M, Gupta V, Bilton TP, Campbell MA, … Andersen, S. 2019. Breaking free: the genomics of allopolyploidy-facilitated niche expansion in white clover. *Plant Cell 31*: 1466-1487. https://doi.org/10.1105/tpc.18.00606

Heslop AD, Arojju SA, Hofmann RW, Ford JL, Jahufer M, Zulfi Z, Larking AC, Ashby R, Hefer CA, Dodds KG, Saei A, O'Connor J, Griffiths AG. 2025. Local adaptation, genetic diversity and key environmental interactions in a collection of novel Red Clover germplasm. *Frontiers in Plant Science 16*: 1553094. https://doi.org/10.3389/fpls.2025.1553094

Jia KH, Zhao W, Maier PA, Hu XG, Jin Y, Zhou SS, Jiao SQ, El-Kassaby YA, Wang T, Wang XR, Mao JF. 2020. Landscape genomics predicts climate change-related genetic offset for the widespread *Platycladus orientalis* (Cupressaceae). *Evolutionary Applications 13*: 665-676.

Legendre P, Legendre L. 2012. *Numerical ecology* (Vol. 24, 3rd ed.). Elsevier.

Mead A, Fitz-Gibbon S, Knapp J, Sork VL. 2024. Comparison of conservation strategies for California Channel Island Oak (*Quercus tomentella*) using climate suitability predicted from genomic data. *Evolutionary Applications 17*: e70057. https://doi.org/10.1111/eva.70057

Omondi EO, Lin CY, Huang SM, Liao CA, Lin YP, Oliva R, van Zonneveld M. 2024. Landscape genomics reveals genetic signals of environmental adaptation of African wild eggplants. *Ecology and Evolution 14*: e11662. https://doi.org/10.1002/ece3.11662

Taylor NL, Quesenberry KH. 1996. Red Clover Science Vol. 28. Netherlands: Springer Science & Business Media. https://doi.org/10.1007/978-94-015-8692-4