DNA fingerprinting and genetic relationships among white clover cultivars

M.Z.Z. JAHUFER, B.A. BARRETT, A.G. GRIFFITHS and D.R. WOODFIELD
AgResearch, Grasslands Research Centre, P B 11008, Palmerston North
 zulfi.jahufer@agresearch.co.nz

Abstract

Genetic diversity provides the raw material for plant improvement programmes. The assessment of genetic diversity is commonly based on plant morphology; however, this indirect measure of genetic diversity is often confounded by environmental effects. Microsatellite (SSR) markers are a powerful tool that directly measures variation among plants at the DNA level. We assayed the pattern of relatedness among 32 white clover (*Trifolium repens* L.) cultivars using 39 SSR markers that are evenly distributed within the white clover genome. The 390 alleles detected were used to quantify the 496 pairwise genetic similarity estimates among the cultivars, detecting a mean value of 0.56 (range 0.44 – 0.68). Cluster analysis assigned the 32 cultivars to 14 clusters based on their genetic similarity. The resulting clusters revealed strong correlation with known genetic relationships based on pedigree and the geographic origin of the parental germplasm. The development of a DNA fingerprinting technique based on high-throughput SSR markers has potential applications in plant variety rights, seed certification and variety identification in grazed pastures.

Keywords: cluster analysis, DNA fingerprinting, genetic diversity, microsatellite, molecular markers, *Trifolium repens* L

Introduction

Over the past 20 years a diverse array of white clover cultivars have been grown in New Zealand, for either seed production or commercial use on farms (AgriQuality 2003). The majority of these cultivars have been bred by AgResearch, including a number of cultivars developed for overseas markets (Woodfield *et al.* 2001). Currently identification of, and differentiation between, cultivars is based predominantly on morphological characters, as is assessment of genetic diversity. While these morphological descriptors are commonly used for plant characterisation and identification by farmers, plant breeders and seed certification field officers, they are an indirect measure of genetic differences and may be confounded by environmental effects (Byth 1981).

Molecular markers provide a robust method for

direct assessment of relatedness among plants at the DNA level (Lee 1995), providing a complementary technology to augment morphological measures of plant identity and genetic diversity. A range of molecular markers have been used to study genetic variation in the major crop species such as rice (*Oryza sativa* L.) (Ni *et al.* 2002), sorghum (*Sorghum bicolor* L.) (Dahlberg *et al.* 2002), maize (*Zea mays* L.) (Warburton *et al.* 2002) and wheat (*Triticum* spp.) (Mukhtar *et al.* 2002). Initial molecular marker-based genetic diversity assays in white clover have used a variety of DNA marker systems on small subsets of germplasm (reviewed by Forster *et al.* 2001); however no comprehensive effort has been initiated to develop a robust fingerprinting technique. Importantly, a recent study revealed significant genetic diversity among cultivars that were morphologically similar, highlighting the importance of utilising comprehensive, direct measures of diversity within white clover germplasm (Griffiths *et al.* 1999). This is very relevant for PVR, because morphological characteristics alone may not be adequate to identify and characterise a new variety. Microsatellite or simple sequence repeat (SSR) markers were used in the current study because they are highly polymorphic, dispersed throughout the white clover genome, and are amenable to high-throughput automated analytical systems (Faville *et al.* 2003).

The development of variety and genotype-specific DNA markers that provide reproducible fingerprints would be a useful innovation for cultivar identification in PVR testing and seed certification, resulting in the supply of higher quality seed to farmers. It will also provide a better understanding of the genetic relationships among the white clover cultivars currently available to New Zealand farmers. Reliable DNA profiles would also enable farmers to assess cultivar authenticity, and test the perenniality and stability of new white clover cultivars under grazing.

In this study we examine the genetic relationships among 32 New Zealand and overseas bred white clover cultivars based on their SSR marker profiles and report initial steps to develop a repeatable DNA fingerprinting technique to differentiate among white clover cultivars.

Materials and methods

Plant material

Thirty-two white clover cultivars were selected for this study, including four pre-release cultivars that are currently in PVR trialling and should be commercially available within the next year (Table 1). Twenty-six of the entries were bred in New Zealand, while the five overseas cultivars chosen for inclusion in this experiment are industry standards in Australia (Haifa), Europe (Aran and Milkanova) and North America (Louisiana S.1 and Regal). The last entry, SRVR, is an important public Ladino germplasm accession that has contributed to a number of international cultivars. More detailed information and pedigrees for a number of these cultivars are available in Caradus & Woodfield (1997).

Emerging leaf tissue was individually sampled from 12 plants per cultivar, and DNA was purified using the Qiagen DNeasy plant DNA kit (Qiagen, USA). Equal amounts of DNA from each plant were combined into 32 cultivar-specific samples; and into

a single pooled sample including all 384 plants.

Estimation of genetic diversity in a cross-pollinating species such as white clover requires random sampling of individuals from within heterogeneous populations. Kölliker *et al.* (2001) indicated that a sample of 15 to 20 white clover genotypes would adequately represent most of the alleles within a population. In this study samples of 12 genotypes were used to maximise the number of cultivars that could be genotyped.

Molecular markers

Sixty-four white clover SSR markers were selected to provide a representative sampling across the eight homoeologous chromosome pair groups of the white clover genome (Faville *et al.* 2003). Initially, the 64 SSRs were screened against the single pooled (n=384 genotypes) DNA sample using an amplification protocol optimised for white clover (Griffiths *et al.* 2002). Thirty-nine markers displaying robust amplification in the single pool were chosen to fingerprint the 32 cultivar-specific

Table 1 The genetic background of 32 white clover cultivars used for DNA profiling (GC74, GC91, GC95 and GC122 are pre-release cultivars).

Cultivars	Pedigree
Barblanca	New Zealand, French, Portuguese and Spanish germplasm
Bounty	New Zealand, Mediterranean, and southern French germplasm
Challenge	New Zealand and Mediterranean germplasm
Colt	North Carolina, non-ladino, ecotype
Crusader	Syrian and French germplasm
Demand	New Zealand and Mediterranean germplasm
Destiny	Southern European germplasm
Durana	Georgia (USA) non-ladino ecotype
Huia	New Zealand ecotypes
Kopu	New Zealand, Spanish, USA (Ladino) germplasm
Kopu II	Hybrid between germplasm from Europe, New Zealand and USA
Milton	Hybrid between New Zealand and southern French germplasm
Nomad	New Zealand dryland ecotypes
NuSiral	Algerian germplasm
Patriot	Hybrid between a non-ladino Georgia (USA) ecotype and Ladino germplasm
Pitau	New Zealand and Spanish germplasm
Prestige	Northland, New Zealand ecotypes
Sustain	New Zealand, Mediterranean and USA ecotype germplasm
Tahora	New Zealand moist hill country ecotypes
Tillman II	Reselection from USA Ladino germplasm
Tribute	New Zealand, Syrian, and southern European germplasm
Triffid	New Zealand, Portuguese and USA ecotype germplasm
GC74	Syrian, non-ladino and Portuguese germplasm
GC91	Syrian, French and USA (ladino) germplasm
GC95	New Zealand, southern European, USA and Italian Ladino.
GC122	Mediterranean, Southern European, and USA germplasm
<i>Standards</i>	
Aran	Irish cultivar based on French and Israel germplasm
Haifa	Australian cultivar based solely on Israel germplasm
Louisiana S1	Bred in the USA from ecotype germplasm
Milkanova	Danish cultivar bred from northern European germplasm
Regal	USA Ladino cultivar
SRVR	USA Ladino germplasm

DNA pools. The numbers of markers on each of the eight white clover homoeologous groups (A to H) were 5, 5, 4, 5, 6, 5, 5 and 4, respectively. All SSR profiles were resolved in a ABI3100 capillary electrophoresis array, and data were collected using ABI Genotyper 3.7 software (2001).

Data analysis

The number of alleles detected per marker, ranged from 2 to 13 among the cultivars. Allele sizes ranged from 112 to 386 base pairs. The range of polymorphism among the 32 cultivars based on the 39 SSR markers is indicated by the range of total number of peaks per cultivar, 129 to 216. However, there were some missing values because certain markers did not work in all cultivars.

The 39 SSR allele profiles of the 32 cultivars were combined into a two dimensional data matrix based on presence or absence of a particular allele in each sample. This matrix was used to calculate the 496 pairwise Jaccard's genetic similarity coefficients (Jaccard 1908) among the 32 cultivars in NTSYS-pc version 1.80 (Rohlf 1998). The 32 cultivars were grouped based on the genetic similarity of their marker profiles using an agglomerative hierarchical clustering analysis with squared Euclidean distance as a measure of dissimilarity, and incremental sums of squares as a grouping strategy (Ward 1963; Wishart 1969). The GEBE software package was used to conduct the clustering (Watson *et al.* 1995). Identification of the optimum level of truncation was based on the distribution of among and within cultivar group sums of squares.

Results and discussion

Genetic analysis of the 32 white clover cultivars based on 39 polymorphic SSR markers detected 390 alleles across all loci. Twelve of the 32 cultivar

samples had SSR alleles that were unique to that cultivar alone (Table 2). The Huia and Louisiana S.1 samples tested had the most unique alleles (four each). These SSR markers that generate unique alleles will play a key role in creating a cost effective clover cultivar fingerprinting technique.

Estimates of genetic similarity between cultivars ranged from 0.44 (between Durana and Huia) to 0.68 (between Regal and Tillman II). Indicating that among the set of 32 cultivars, Durana and Huia were the most dissimilar and Regal and Tillman were the most similar. The mean genetic similarity coefficient among all 32 cultivars was 0.56.

The clustering of cultivars based on the genetic similarity estimates corresponds with known pedigrees (Figure 1). The two major branches in the dendrogram separate those varieties with a high proportion of Huia, northern and central European germplasm (groups 1-6), from those with a high proportion of southern European and Ladino germplasm (groups 7-14).

The first branch (groups 1-3) contains Huia and other cultivars with a strong Huia contribution to their pedigrees. The exception to this is Aran which would fit better into group 10, based on its reported French and Israeli parentage (Caradus & Woodfield 1997). Groups 4 and 5 each contain a single reference cultivar Louisiana S.1 and Haifa, respectively. Louisiana S.1 was bred in southern USA, but probably resulted from natural selection within Dutch white clover populations sown before Ladino white clovers were introduced into the USA (Caradus & Woodfield 1997). Group 6 contains cultivar Milkanova, a Danish cultivar based predominantly on Northern European ecotypes, and Grasslands Pitau and Bounty. Pitau and Bounty both have a significant Huia contribution in their pedigree.

In the second major branch (groups 7-14), there are also several notable clusters that reflect known pedigree. Groups 7, 8, 11 and 14 do not contain any Huia parentage. All of the ladino cultivars (Regal, SRVR and Tillman II) clustered in group 7. These ladino cultivars are closely related as three of the five parents of Regal were also parents of Tillman (Caradus & Woodfield 1997) from which Tillman II was reselected for improved seed yield. Tillman also contributed to the development of SRVR (Caradus & Woodfield 1997). Destiny is not a ladino because approximately 60% of Destiny genotypes are cyanogenic. However, Destiny was bred from material collected in Northern Italy close to where the Ladino introductions into the USA originated. Group 9 contains Kopu which has a substantial

Table 2 Members of the set of 32 white clover cultivars that had unique SSR alleles. Number of unique alleles and allele size indicated.

Cultivars	No. unique alleles	Allele size (bp)
Huia	4	164, 126, 186, 190
Louisiana S1	4	159, 279, 240, 303
SRVR	2	255, 222
Milkanova	1	179
Haifa	1	255
Kopu	2	218, 128
Demand	1	223
Prestige	1	267
NuSiral	1	270
Destiny	1	129
Nomad	1	272
Durana	1	124

Figure 1 Dendrogram based on cluster analysis of cultivar by SSR marker matrix. Fusion points, at the 14 group (Gr) level, and cultivars (Hui, Huia; Ara, Aran; Pre, Prestige; Cha, Challenge; Kp2, Kopu II; Dem, Demand; Sus, Sustain; Lou, Louisiana S1; Hai, Haifa; Mlk, Milkanova; Pit, Pitau; Bou, Bounty; Srv, SRVR; Reg, Regal; Til, Tillman II; Des, Destiny; Tah, Tahora; Kop, Kopu; Cru, Crusader; Mil, Milton; Tri, Tribute; Col, Colt; Pat, Patriot; Dur, Durana; Trf, Triffid; Nom, Nomad; Bar, Barblanca; NuS, NuSiral; GC74; GC91; GC95; GC122) are shown.

ladino parentage (including Regal) in its pedigree. Groups 10, 12, 13 and 14 contain cultivars that have a high Southern European and Mediterranean parentage. The cultivars in group 10 all contain a high proportion of French germplasm, predominantly from cultivars Crau and Major. Similarly three of the four cultivars (Barblanca, GC74 and GC122) in group 13 and 14 have a significant Portuguese and Southern European content. The exception, NuSiral, derives solely from an Algerian introduction into Australia.

There are some clear examples of cultivars clustering into unexpected groups. Tahora (group 9) and Nomad (group 12) derive solely from New Zealand ecotypes but clustered with groups containing predominantly southern European parentage. This may reflect inadequate sampling of

these cultivars but it could also indicate that their parental ecotypes originated from southern European introductions rather than Huia as suspected.

Caradus *et al.* (1989) investigated the genetic relationships among a world collection of 109 white clover cultivars (including 11 cultivars used in the current study, *viz.* Aran, Haifa, Huia, Kopu, Louisiana S.1, Milkanova, Regal, Tahora, Demand, Prestige and Tillman) based on plant morphology, flowering, cyanogenesis, and leaf markings. Associations between geographic origins of the cultivars based on leaf size and cyanogenesis, were evident in this study, for example cultivars with a strong ladino background also clustered together. In contrast to the current study, cultivars Aran and Haifa (Israeli background) were in a similar group.

Figure 2 Genetic barcodes based on 39 SSR marker profiles across 32 white clover cultivars. The homoeologous groups (chromosomes) are indicated as A, B, C, D, E, F, G and H. Each horizontal bar represents an allele. The variation in density of the bars is caused by numbers of alleles being close together.

Kölliker *et al.* (2001) used AFLP markers to characterise genetic variation among 52 white clover cultivars and accessions. Significant genetic variation was detected, and cluster analysis of the AFLP data indicated only a partial association between cultivar groups and geographic origin. However, except for California Ladino, all the other ladino type cultivars fell into one major cluster

(Kölliker *et al.* 2001).

In our study, the clustering of cultivars across the 14 groups based on 39 SSR markers indicated a strong correlation with known pedigrees and geographic origin of the germplasm used in their breeding. This degree of discrimination supports the utilisation of these SSR markers for DNA fingerprinting applications.

DNA finger printing

Each of the cultivars had a unique DNA fingerprint based on the SSRs used. The SSR allele by cultivar matrix is presented in the form of a barcode chart (Figure 2). Each column can be considered as a genetic bar code or finger print of a specific cultivar. The horizontal bars that run down each cultivar (column), indicate the allele presence/absence pattern generated by the 39 SSR markers. Additionally, the 8 homoeologous chromosome pair groups of white clover are indicated by letters A-H at the left of the barcode, indicating the broad distribution within the clover genome of relatedness patterns. These DNA profiles are specific to each cultivar and unlike morphology are independent of environmental effects. Identification of SSR alleles specific to individual cultivars will enhance their identification. In this experiment, 25% of the cultivars had unique alleles, with the separation among other cultivars based on specific combinations of allele profiles.

The barcode DNA profiles of each cultivar should be interpreted with care. As white clover is a cross pollinating species, there is always the possibility of genetic shift or contamination of a cultivar (e.g. during seed production). This complicates construction of an accurate and repeatable DNA profiling system for species such as white clover. In order to overcome such constraints, the next phase of this research will need to:

- a) assay additional markers within the clover genome to increase the number of detectable unique alleles per cultivar;
- b) improve the accuracy of each cultivar fingerprint by adding information to the 'missing values' indicated by the grey regions in Figure 2;
- c) identify a subset of SSR marker alleles that provides maximum discrimination among cultivars; and
- d) assess barcode reproducibility by comparing DNA profiles of individual cultivars based on multiple samples and generations.

A major objective of this project was to develop accurate and reproducible cultivar identification genetic fingerprinting systems. The genetic barcodes generated for the 32 cultivars in this study is the first step towards development of such a technique for implementation in clover research, the seed industry, and for on-farm applications. Outcomes from this research will be used to develop DNA marker based assays that could be used under field or lab conditions for the broad array of applications described above.

ACKNOWLEDGEMENTS

We acknowledge the contribution of the Plant Biotechnology Centre, Agriculture Victoria in the development of the white clover SSR markers through the Plant Genomics Program with AgResearch. We also thank Tony Dunn and Deborah Knox for their technical assistance.

REFERENCES

- AgriQuality New Zealand. 2003. Seed Certification Statistics 2002-2003. AgriQuality Seed Certification Bureau, Christchurch. 57 pp.
- Byth, D.E. 1981. Genotype x environment interaction and environmental adaptation in plant improvement – an overview. pp. 2-11. *In: Interpretation of plant response and adaptation to agricultural environments*. Eds. Byth, D.E.; Mungomery, V.E. Australian Institute of Agricultural Science, Queensland Branch, Australia.
- Caradus, J.R.; Woodfield, D.R. 1997. World checklist of white clover varieties II. *New Zealand Journal of Agricultural Research* 40: 115-206.
- Caradus, J.R.; MacKay, A.C.; Woodfield, D.R.; Bosch, van den J.; Wewala, S. 1989. Classification of a world collection of white clover cultivars. *Euphytica* 42: 183-196.
- Dahlberg, J.A.; Zang, X.; Hart, G.E. and Mullet, J.E. 2002. Comparative assessment of variation among Sorghum germplasm accessions using seed morphology and RAPD measurements. *Crop Science* 42: 292-296.
- Faville, M.; Barrett, B.; Griffiths, A.; Schreiber, M.; Mercer, C.; Baird, I.; Ellison, N.; Bryan, G.; Woodfield, D.; Forster, J.W.; Sawbridge, T.; Spangenberg, G.; Easton, E. 2003. Implementing molecular marker technology in forage improvement. *Proceedings of the New Zealand Grassland Association* 65: 229-238.
- Forster, J.W.; Jones, E.S.; Kölliker, R.; Dayton, M.C.; Dumsday, J.L.; Dupal, M.P.; Guthridge, K.M.; Mahoney, N.L.; Jong, E. van Zijl de; Smith, K.F. 2001. Development and implementation of molecular markers for forage crop improvement. pp. 101-133. *In: Molecular Breeding in Forage Crops*. Ed. Spangenberg, G. Kluwer Academic Publishers, Netherlands.
- Genotyper software version 3.7. 2001. Applied Biosystems. Foster City, Ca, USA.
- Griffiths, A.; Kidwell, K.; Barrett, B.; Williams, W.; Woodfield, D. 1999. Comparing AFLP and RFLP marker systems for diversity assessment in white clover. *Proceedings of the 11th Australian Plant Breeding Conference* (2): 210-211.
- Griffiths, A.; Barrett, B.; Ellison, N.; Sawbridge, T.;

- Spangenberg, G.; Bryan, G.; Schreiber, M. 2002. Getting out there: transportability of white clover EST and genomic DNA- derived microsatellites. *Proceedings of the Australasian Plant Breeding Conference 12*: 69-72.
- Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. *Bulletin de la Société Vaudoise des Sciences Naturelles 44*: 223-270.
- Kölliker, R.; Jones E.S.; Jahufer M.Z.Z. and Forster J.W. 2001. Bulked AFLP analysis for the assessment of genetic diversity in white clover (*Trifolium repens* L.). *Euphytica 121*: 305-315.
- Lee, M. 1995. DNA markers and plant breeding programs. *Advances in Agronomy 55*: 265-344.
- Mukhtar, M.S.; Rahmanw, M. and Zafar, Y. 2002. Assessment of genetic diversity among wheat (*Triticum aestivum* L.) cultivars from a range of localities across Pakistan using random amplified polymorphic DNA (RAPD) analysis. *Crop Science 128*: 417-425.
- Ni, J.; Colowit, P.M. and Mackill, J. 2002. Evaluation of genetic diversity in Rice subspecies using Microsatellite markers. *Crop Science 42*: 601-607.
- Rohlf, F.J. 1998. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System. Version 2.02I, Exeter Software, New York.
- Ward, J.H. 1963. Hierarchical grouping to optimise an objective function. *Journal of the American Statistics Association 58*: 236-244.
- Warburton, M.L.; Xianchun, X.; Crossa, J.; Franco, J.; Melchinger, A.E.; Frisch, M.; Bohn, M.; Hoisington, D. 2002. Genetic characterization of CIMMYT inbred maize lines and open pollinated populations using large scale fingerprinting methods. *Crop Science 42*: 1832-1840.
- Watson, S.L.; DeLacy, I.H.; Podlich, D.W.; Basford, K.E. 1995. GEBE1 An analysis package using agglomerative hierarchical classificatory and SVD ordination procedures for genotype x environment data. Centre for Statistics Research Report, Department of Agriculture, The University of Queensland, Australia.
- Wishart, D. 1969. Algorithm for hierarchical classifications. *Biometrics 25*: 165-170.
- Woodfield, D.R.; Clifford, P.T.P.; Cousins, G.R.; Ford, J.L.; Baird, I.J.; Miller, J.E; Woodward, S.L.; Caradus, J.R. 2001. Grasslands Kopu II and Crusader: New generation white clovers. *Proceedings of the New Zealand Grassland Association 63*: 103-108.