'Breeding for Success: Diversity in Action' C.F. Mercer (ed).
Proceedings of the 13th Australasian Plant Breeding Conference,
Christchurch, New Zealand 18-21 April 2006. pp. 928-931.

Pastoral Genomics - a foray into the clover genome.

Chris S. Jones[1,2] , *Warren M. Williams*[1,2], *Kerry R. Hancock*[1,2], *Nick W. Ellison*[1,2], *Alicia G. Scott*[1,2], *Vern E. Collette*[1,2], *M. Zulfi Z. Jahufer*[1,2], *Kim A. Richardson*[1,2], *Mike J. M. Hay*[1,2], *Susanne Rasmussen*[1,2], *Craigh G. Jones*[1,2] *and Andrew G. Griffiths*[1,2].
[1]AgResearch Limited, Grasslands Research Centre, Private Bag 11008, Palmerston North, New Zealand
[2]Pastoral Genomics c/o ViaLactia Biosciences (NZ) Ltd, PO Box 109185, Newmarket, Auckland, New Zealand

Abstract. Pastoral Genomics is engaged in applying molecular mapping and functional genomics to characterise and manipulate key phenotypic traits associated with improved productivity in the commercially important forage crop white clover (*Trifolium repens* L.). We report on the development of a framework linkage map for white clover and the application of this map to the identification of quantitative trait loci (QTL) associated with vegetative and seed yield traits. Homologues of genes responsible for the traits of interest have been identified and cloned in order to identify 'perfect markers' for marker assisted breeding and to develop a cisgenic® clover cultivar.

We introduce western clover (*T. occidentale*), a diploid perennial clover species closely related to white clover, as a stoloniferous clover model to dissect low heritability traits initially intractable in white clover. We have identified extensive genetic variation in this species, which previously has been reported to be remarkably uniform. We have developed a framework map for *T. occidentale* and demonstrated conservation of marker order and spacing with white clover. An ethyl methane sulfonate (EMS) mutant population has been used to isolate three distinct groups of condensed tannin (CT) mutants. We have also investigated our model's transformation potential, achieving frequencies 2-3 fold higher than for white clover and suggest that this model is potentially useful for genomic analysis of white clover and other clonal species.

Introduction

Pastoral Genomics (PG) is the sole investment by the New Zealand pastoral industries into forage improvement through biotechnology. The PG white clover programme is funded by Meat and Wool NZ, Fonterra, AgResearch, Deer Industry NZ, FRST and Dairy InSight. The aim of the programme is to make step changes to forage improvement by applying molecular mapping and functional genomics to the characterisation of key phenotypic traits and the development of cisgenic® products. The molecular mapping and functional genomics programmes have been developed utilising gene sequences identified through the GeneThresher® (GT) technology (Rabinowicz *et al*. 1999), to which PG has exclusive pasture plant access, and Expressed Sequence Tag (EST) databases. The application of GT technology has enabled Pastoral Genomics to explore the clover genome for novel sequences within the gene space not commonly identified by EST-based gene discovery.

The mapping and functional genomics programmes are being applied to the identification of genes and genetic markers associated with enhanced persistence, enhanced dry matter production and increased condensed tannin (CT) content. CTs are responsible for protecting proteins whilst passing through the rumen which reduces pasture bloat, methane production while controlling gastrointestinal nematodes.

Ultimately, these discoveries will be used to support the production of new clover cultivars through the application of marker-assisted selection (MAS) to conventional breeding. Novel cultivars will also be evaluated primarily through the development of genetically modified clover containing DNA only derived from clover (termed cisgenic®).

Results and Discussion

Within Pastoral Genomics a database containing DNA sequences from a range of sources is a focal point for a fully integrated molecular mapping and functional genomics approach to exploit the clover genome. The proprietary database contains 445,116 sequences and is comprised of 364,537 white clover GT DNA sequences. These are enriched for regions of the gene space and are assembled into 86,643 contigs. The rest of the database is comprised of 14,527 sequences enriched for tissue and species-specific expression via Suppression Subtractive Hybridisation (SSH); 4,360 white clover genomic DNA sequences; 59,851 white clover EST sequences; and 1,841 *Trifolium occidentale* SSH sequences. In addition, we are developing a further 34,000 *T. occidentale* GT sequences to further enhance this substantial combined resource.

Customised bioinformatics tools have been developed to facilitate the identification of genetic markers, quantitative trait loci (QTL), genes and their regulatory elements. The bioinformatics tools that have been developed include: an interactive genome browser (GBrowse), which enables our proprietary clover sequences to be aligned with the *Medicago truncatula* genome in a secure environment and facilitates *in silico* comparative mapping studies; data mining pipelines, which enable sequence assembly, gene annotation, protein domain searches, microsatellite (SSR) and single nucleotide polymorphism (SNP) marker identification and primer design.

Utilisation of the genetic marker identification pipeline detected over 15,600 GT sequences harbouring SSRs and over 13,500 harbouring SNPs. From these, primer pairs were designed for over 4,000 non-redundant SSRs and a subset of these has been employed in developing a comprehensive white clover consensus map containing 691 loci over 16 linkage groups (map length 1244 cM; map density one locus per 0.6 cM; Griffiths *et al.* 2006b). This map has been licensed to AgResearch so that it may be aligned with the AgResearch linkage map (Barrett *et al*. 2005). The combined PG-AgResearch marker resource will facilitate fine-mapping of QTLs, validation of QTLs in independent populations and accelerate the application of marker-assisted selection (MAS) in white clover.

Two white clover mapping populations contrasting in a range of traits have been developed and the analysis of the first year's phenotypic data completed. In mapping population 1, eight QTL for seed yield traits were identified. QTL representing number of seeds per floret, seed yield per inflorescence, and seed yield per plant accounted for 36.2%, 24.2% and 22.4% of the phenotypic variation, respectively, and were co-located. QTL representing number of inflorescences and inflorescence diameter accounted for 19.3% and 14.8% phenotypic variation, respectively. In addition, three QTLs were identified associated with floral maturity accounting for 15.1%, 16.9% and 13.9% variation, respectively. In mapping population 2, replicated over three geographic locations in New Zealand, three QTL for vegetative traits: stolon thickness, leaf size and stolon branching, were identified from the first year's field data. These accounted for 14.6, 17.3 and 11.1% of the phenotypic variation, respectively. Furthermore, QTL of

highly correlated traits such as leaf size and stolon thickness were found to be co-located.

White clover is an allogamous allotetraploid, and is consequently a challenging species in which to develop MAS, particularly for low heritability traits. To provide a complementary approach to dissecting clover QTL for MAS, we have identified and investigated a suitable model for white clover. The model is *T. occidentale*, a diploid, autogamous, perennial clover species that is phylogenetically closely related to, and a candidate progenitor of, white clover (Ellison *et al*. 2006, Williams *et al*. 2006). This coastal species shares white clover's stoloniferous, clonal habit and although it has been reported to be remarkably uniform, recent collections of *T. occidentale* have found extensive genetic variation in morphology, breeding system and molecular traits (Williams *et al*. 2006). As a result of these findings, this species is potentially useful for genomic analysis of many white clover traits.

Similarity between the two species has been exemplified by development of a *T. occidentale* framework linkage map integrating information from two mapping populations (362 loci over 8 linkage groups with a total map length of 610 cM and 1 locus per 1.7 cM; Griffiths *et al.* 2006b) and its alignment with the white clover map which exhibited both conservation of marker order and spacing (Griffiths *et al.* 2006a). From the first year's data, initial QTL for stolon diameter, internode length, flower number, nodal rooting and plant vigour were identified. Our aim is now to amalgamate the two clover maps in order to make a consensus clover framework map.

Additional *T. occidentale* genetic resources include an EMS mutant population containing 1075 EMS M2 seed lines which was created for the map-based cloning of new and novel clover genes. Over 8000 plants were screened for mutations in CT production and distribution by the *p*-dimethylaminocinnamaldehyde (DMACA) CT-specific staining assay (Li *et al.* 1996), and 42 mutants were identified which fall into 3 distinct phenotype groups: no CTs; abnormal CT profiles; and no trichomes. Some of these mutants also showed a transparent testa phenotype (Koorneef, 1990) with no CTs in the seed coat while others had abnormal flowers. These phenotypes segregated with the CT mutant phenotype suggesting that the same gene, or a second gene in close proximity to the CT gene, had been mutated.

Finally in order to determine the applicability of *T. occidentale* as a diploid genomic model of white clover we investigated its transformation potential. Frequencies 2 to 3 fold higher than those for white clover (in the order of 7 to 10%) were achieved suggesting that this species has great potential as a genomic model.

Application of an integrated approach, utilising both white clover and *T. occidentale* resources has facilitated the investigation of key target traits such as CT biosynthesis and persistence. We have utilised the GT database to design degenerate PCR primers to identify and isolate candidate genes for these key traits from white clover. White clover homologues representing most structural and regulatory genes involved in the CT biosynthetic pathway and candidate sequences targeting enhanced persistence have been identified. The CT candidate genes have been further validated by expression profiling which has demonstrated that they are expressed in CT rich tissue. Further profiling of selected genes has been carried out in *T. occidentale* EMS mutants that have mutations in the later stages of CT biosynthesis. Persistence candidates include: homologues of genes involved in lateral root and branch formation; novel sequences identified in the PG SSH libraries which have been demonstrated to be expressed solely in stolon tissue; ORF54 (identified originally by ViaLactia) a gene that

increases biomass through increased tillering in rice; and VP1 a vacuolar H^+-pyrophosphatase which confers increased resistance to high concentrations of NaCl and to water deprivation through increased solute accumulation and water retention (Gaxiola *et al*. 2001).

The GT database is a significant resource for plant promoters. A total of 496 promoter candidates have been identified through sequence homology searches with characterised genes, particularly from closely-related legumes. PG has cloned promoters associated with traits of interest and these are currently being analysed together with our candidate genes for the development of 'perfect markers'. The promoters identified by PG are also being developed as a resource for constitutive, tissue specific and temporally regulated expression of targeted cisgenic® white clover products (Scott *et al.* 2006).

In conclusion, PG has developed comprehensive molecular mapping and functional genomics capabilities to exploit the clover genome. We are applying these technologies to accelerate the development of new clover cultivars to benefit New Zealand pastoral industries.

References

Barrett B, Griffiths A, Schreiber M, Ellison N, Mercer C, Bouton JH, Ong B, Forster J, Sawbridge T, Spangenburg G, Bryan GJ, Woodfield DR (2004) A microsatellite map of white clover. *Theoretical and Applied Genetics* 109, 596–608.

Ellison NW, Liston A, Steiner JJ, Williams WM, Taylor NL (2006) Molecular phylogenetics of the clover genus (*Trifolium* - Leguminosae) *Molecular Phylogenetics and Evolution* 39, 688-705.

Gaxiola RA, Li J, Undurraga S, Dang LM, Allen GJ, Alper SL, Fink GR (2001) Drought- and salt-tolerant plants result from overexpression of the AVP1 H^+-pump. *Proceedings of the National Academy of Science of the USA* 98, 11444–11449.

Griffiths AG, Anderson CB, Williams WM, Hay MJM, Jones CG, Jones CS (2006a) Aligning clover genomes: synteny between white clover (*Trifolium repens*), *T. occidentale* and *Medicago truncatula*. Third International Conference on Legume Genomics and Genetics, Brisbane, Australia.

Griffiths AG, Bickerstaff P, Anderson CB, Franzmayr BK, Jones CS (2006b) Threshing the white clover genome for gene-associated molecular markers. In 'Breeding for success: diversity in action. Proceedings of the 13th Australasian Plant Breeding Conference, Christchurch, New Zealand'. (Ed. CF Mercer) pp. 817-821. www.apbc.org.nz.

Koorneef, M (1990). Mutations affecting the testa color in *Arabidopsis. Arabidopsis Information Service* 27, 1-4.

Li, YG, Tanner G, Larkin P (1996) The DMACA-HCl protocol and the threshold proanthocyanidin content for bloat safety in forage legumes. *Journal of the Science of Food and Agriculture* 70, 89–101.

Rabinowicz PD, Schutz K, Dedhia N, Yordan C, Parnell LD, Stein L, McCombie WR, Martienssen RA (1999) Differential methylation of genes and retrotransposons facilitates shotgun sequencing of the maize genome. *Nature Genetics* 23, 305-309

Scott A, Ellison N, Richardson K, Allan A, Maher D, Jones C, Griffiths AG (2006) Isolation of white clover promoters using a Gene Thresher® database. In 'Breeding for success: diversity in action. Proceedings of the 13th Australasian Plant Breeding Conference, Christchurch, New Zealand'. (Ed. CF Mercer) pp. 775-779, www.apbc.org.nz.