



Clover ASTRA: a Web-Based Resource for Trifolium EST Analysis

G. C. Spangenberg

La Trobe University, Australia

T. Sawbridge

La Trobe University, Australia

E. K. Ong

La Trobe University, Australia

C. G. Love

Victorian Bioinformatics Consortium, Australia

T. A. Erwin

La Trobe University, Australia

See next page for additional authors

Follow this and additional works at: <https://uknowledge.uky.edu/igc>



Part of the [Agricultural Science Commons](#), [Agronomy and Crop Sciences Commons](#), [Plant Biology Commons](#), [Plant Pathology Commons](#), [Soil Science Commons](#), and the [Weed Science Commons](#)

This document is available at <https://uknowledge.uky.edu/igc/20/satellitesymposium5/53>

The XX International Grassland Congress took place in Ireland and the UK in June-July 2005.

The main congress took place in Dublin from 26 June to 1 July and was followed by post congress satellite workshops in Aberystwyth, Belfast, Cork, Glasgow and Oxford. The meeting was hosted by the Irish Grassland Association and the British Grassland Society.

Proceedings Editor: D. A. McGilloway

Publisher: Wageningen Academic Publishers, The Netherlands

© Wageningen Academic Publishers, The Netherlands, 2005

The copyright holder has granted the permission for posting the proceedings here.

Presenter Information

G. C. Spangenberg, T. Sawbridge, E. K. Ong, C. G. Love, T. A. Erwin, E. G. Logan, and D. Edwards

Clover ASTRA: a web-based resource for *Trifolium* EST analysis

G.C. Spangenberg^{1,2,3}, T. Sawbridge^{1,2}, E.K. Ong¹, C.G. Love^{1,2}, T.A. Erwin^{1,2}, E.G. Logan^{1,2} and D. Edwards^{1,2,3}
¹Primary Industries Research Victoria, Plant Biotechnology Centre, La Trobe University, Bundoora, Victoria 3086, Australia ²Victorian Bioinformatics Consortium, Australia ³Molecular Plant Breeding Cooperative Research Centre, Australia Email: german.spangenberg@dpi.vic.gov.au

Keywords: white clover, expressed sequence tags (ESTs), gene annotation, bioinformatics ASTRA database, simple sequence repeats (SSRs), single nucleotide polymorphisms (SNPs)

Introduction White clover (*Trifolium repens* L.) is a major temperate forage legume.

Materials and methods A resource of 42,017 white clover expressed sequence tags (ESTs) from single pass DNA sequencing of randomly selected clones from 16 cDNA libraries representing a range of plant organs, developmental stages and environmental treatments was generated (Sawbridge *et al.*, 2003). Each of these sequences has been annotated by comparison to the GenBank and SwissProt public sequence databases. Automated intermediate Gene Ontology (GO) annotation has also been determined.

Results and conclusions Sequence annotation was determined for 74%, 58% and 54% of sequences using GenBank, SwissProt and GO annotation respectively, with only 20% of sequences determined to have no comparative annotation and thus appear to be unique to white clover (Fig 1). EST redundancy was resolved through assembly with CAP3, identifying at total of 15,989 unigenes within this dataset. All sequences and annotation are maintained within an ASTRA format MySQL database, with web based access for text searching, BLAST sequence comparison and Gene Ontology hierarchical tree browsing. Each white clover sequence was mapped onto an EnSEMBL genome viewer for comparison with the complete genome sequence of *Arabidopsis thaliana* and expressed sequences from related legumes.

In total 5,407 white clover simple sequence repeat (SSR) molecular markers were identified using SSRPrimer and specific PCR amplification primers were designed. 18,517 candidate single nucleotide polymorphism (SNP) molecular markers and 1,706 insertion deletions (InDels) were also identified across 1,409 loci using AutoSNP. The integration of different analysis tools and the resulting data within a central resource allow users to identify novel candidate genes and associated molecular genetic markers that can be applied to functional genomics and germplasm enhancement in this agronomically important forage crop.

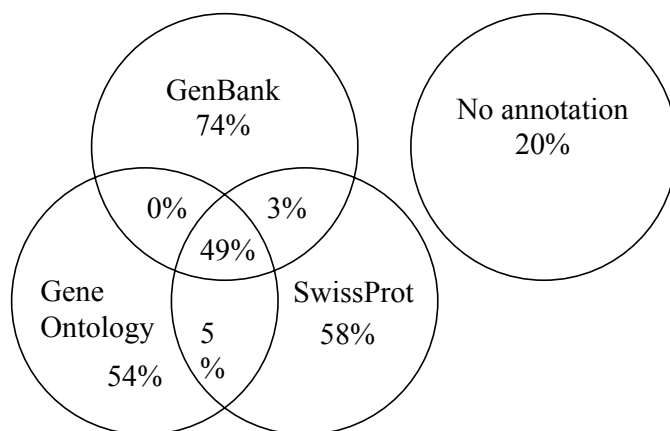


Figure 1 Annotation of white clover genes within the ASTRA database

Reference

Sawbridge, T., Ong, E.K., Binnion, C., Emmerling, M., Meath, K., Nunan, K., O'Neill, M., O'Toole, F., Simmonds, J., Wearne, K., Winkworth, A. and Spangenberg, G. (2003). Generation and analysis of expressed sequence tags in white clover (*Trifolium repens* L.). *Plant Science*, 165, 1077-1087.