

auto4x: a novel software to assess the distribution of genetic diversity in autotetraploid species demonstrated with white clover

S. Barth^{1*}, V. Connolly¹, M. Klaas², M. Febrer¹, N. Salamin³

¹Teagasc, Crops Research Centre, Oak Park, Carlow, Ireland; ²National University of Ireland Maynooth, Department of Biology, Cell Biology Laboratory, Ireland; ³University of Dublin, Trinity College, School of Botany, Ireland

Introduction

White clover (*Trifolium repens* L.) is an important forage species in Ireland. Because of its high economic and ecological importance, and the foreseeable need for improvement and adaptation of these crops to changing agricultural needs, charting the existing genetic diversity is an important long-term investment. The assessment of the population structure in autotetraploid species requires a different approach than for diploids, and software to perform the analysis is not readily accessible. The aim of this project was (1) to assess the molecular genetic diversity in the autotetraploid species white clover with co-dominant simple sequence repeat (SSR) markers and (2) to develop and test a new software with a large dataset.

Materials and Methods

Seventeen white clover accessions with 48 individual plants each were investigated. Eight populations were collected in the 1980's in Ireland, five were derived from the Oak Park breeding programme, and four were obtained from the gene bank collection in Gatersleben/Germany. DNA was isolated using the Sigma Plant DNA isolation kit. 38 SSR markers were designed from publicly available microsatellite motifs from GenBank and PCR conditions were optimised. Seven of those SSR markers were used for this study. Fluorescently labelled PCR products were analysed on the ABI3100 sequencer and alleles sized with the Genemapper™ v3.0 software.

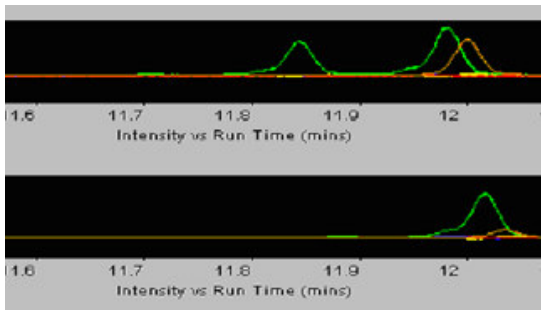


Figure 1: Polymorphism at marker locus TRagr1013 for two lines.

Absolute allele frequencies and relative allele frequencies were summarised to create an inventory of alleles. Relative allele frequency data were used as input data matrix to compute Nei's genetic distance

coefficients for genetic data with the NTSYS statistics package.

To assess the population structure the new software *auto4x* was developed. The internal algorithms were written in C, while the graphic user interface uses the GTK+2 libraries. The algorithms implemented were derived from Ronfort et al. (1998), and a bootstrap procedure was added to test the significance of the different population structure estimates.

To facilitate the conversion of the Genmapper output files to *auto4x* input files, a Visual Basic for applications (VBA)/Excel macro was written.

Results

38 SSR markers for white clover have been developed from *in silico* sequences for this study. Generally, Irish white clover populations displayed a large number of different alleles. Some loci showed evidence of a bottleneck in the Irish populations. Similar bottleneck effects were also found in the non-Irish populations, e.g. in the "Tadjikistan" population. Two exotic gene bank accessions "Tajikistan" and "Iran" were rather distantly related to all other accessions. Closely related to the Irish white clover populations was the land race "Dutch White Clover".

The *auto4x* programme was suitable to calculate the inbreeding co-efficients across populations and differentiation of subpopulations (F_{ST}), effect of the population structure on the average heterozygosity of individuals (F_{IT}), inbreeding within subpopulations (F_{IS}) and rho (Ronfort et al. 1998) across marker loci. F_{IS} values were lowest at marker locus TRagr1075. Restricted gene flow between Irish and non-Irish accessions was found. F_{ST} values ranged from -0.004 to -0.012 across the marker loci (Table 1).

Table 1: F_{ST} , F_{IT} and F_{IS} values across the marker loci

Marker locus	F_{ST}	F_{IT}	F_{IS}
TRagr179	-0.006	0.933	0.935
TRagr1001	-0.004	1.091	1.091
TRagr1013	-0.006	0.973	0.974
TRagr1023	-0.006	0.974	0.975
TRagr1075	-0.012	0.610	0.616
TragrA1.28	-0.005	1.022	1.022

Conclusions

Irish populations contained several unique alleles and are therefore valuable sources for rare alleles. F_{ST} values indicated restricted gene flow between Irish and non-Irish populations. The software developed *auto4x* will be very helpful for the scientific community interested in other autotetraploid species as well. The software will be made available at:

www.tcd.ie/Botany/dnabank/software.

References

Ronfort, J., Jenczewski, E., Bataillon, T., Rousset, F. (1998) *Genetics* 150: 921-930