

Towards a Framework PCR-Based Map of Onion (*Allium cepa* L.)

J. McCallum^{1*}, M. Pither-Joyce¹, A. Clarke², C. Kik³, S. Van Heusden³, J. Jakse⁴ and M.J. Havey⁴

¹New Zealand Institute for Crop and Food Research Limited, Private Bag 4704, Christchurch, New Zealand

²Allan Wilson Centre, Massey University, Palmerston North, New Zealand

³Plant Research International, P.O. Box 16, 6700 AA Wageningen, The Netherlands

⁴USDA-ARS, Dept. of Horticulture, Univ. of Wisconsin, Madison, WI 53706, USA

*Corresponding Author

Keywords: onion, genetic mapping, simple sequence repeat markers

Abstract

Genetic analysis of onion has been hampered by a lack of portable co-dominant markers based on the polymerase chain reaction (PCR). The public release of a relatively large set of non-redundant onion expressed sequence tags (ESTs) in 2003 has provided the opportunity to develop such markers for use in *Allium* research and industry. We have mined this collection for simple sequence repeats (SSRs) and screened over 200 primer sets designed to flank SSRs for their ability to amplify low-copy polymorphic products in *Allium cepa* and *A. roylei*. Screening to date confirms that these are a rich source of polymorphic markers in *Allium*. Half of the sets reveal polymorphism between onion and *A. roylei*, ~40% reveal putative allelic variation in a small set of onion germplasm and 20% reveal polymorphism in a single F₂ family in the cross 'W202A' x 'Texas Grano 438'. Mapping common markers has permitted alignment with the RFLP map developed in the 'Brigham Yellow Globe 15-23' x 'Alisa Craig 43' population. Preliminary studies show these markers are effective in distinguishing variation in allelic diversity among a set of inbred and open-pollinated onion varieties. Our results suggest that these markers will enable effective genome scanning and analysis of genetic diversity and identity in cultivated onion.

INTRODUCTION

Genetic understanding of onion (*Allium cepa* L.) is significantly less developed than other crops due to the practical difficulties of developing and maintaining genetic stocks and a lack of DNA sequence and marker resources. Existing *Allium* genetic maps based on RFLP (King et al., 1998) and AFLP markers (van Heusden et al., 2000b) have permitted the tagging of genome regions associated with economic traits but have not provided portable genetic markers based on the polymerase chain reaction (PCR) that can be widely applied to exploit or extend these findings. Until 2003, sequence resources in *Allium* were limited to small sets of onion expressed sequence tags (ESTs) generated by projects at USDA-ARS (King et al., 1998), Plant Research International (Unpublished) and Crop & Food Research (McCallum et al., 2002). We previously showed that standard mutation-scanning methods such as single-stranded conformation polymorphism (SSCP; McCallum et al., 2001) or cleaved amplified polymorphic sequence (CAPS; van Heusden et al., 2000a) analyses could be used to convert some EST sequences into useful PCR-based genetic markers. The USDA Asparagales genome project has recently released a set of over 18,000 onion ESTs (Kuhl et al., 2004). We describe recent progress in exploiting this resource for simple sequence repeat (SSR) marker development and preliminary results in using such markers in genetic map development and diversity studies.

RESULTS AND DISCUSSION

Genetic Resources

The key genetic resource to date has been the set of 58 F₃ families from the cross

'Brigham Yellow Globe 15-23' x 'Alisa Craig 43' (BYG15 x AC43) used to develop the RFLP map (King et al., 1998). We have more recently developed a population of 59 F₃ families obtained by self-pollinating 82 individual F₂ plants from the cross between 'W202A' (Goldman, 1996) and 'Texas Grano 438' (Asgrow Vegetable Seeds). This population has proved well suited for gene mapping due to its wide adaptation, reasonable storage and segregation for economic traits including pungency, sweetness and fertility restoration. We have also conducted limited marker analysis in an F₂ family of 91 plants from the cross 'Colossal PVP' x 'Early Longkeeper P12' (McCallum et al., 2002). The population of 65 F₂ individuals from the interspecific cross *Allium cepa* ('Jumbo') x *A. roylei* used to develop the AFLP map (van Heusden et al., 2000b) shows great promise for mapping multiple genes due to the ease of detecting polymorphism between *cepa* and *roylei* alleles with simple mutation-scanning techniques.

Although they have proved adequate for mapping studies to date, these populations are very small by current gene mapping standards. Considerably larger populations are required to enable finer scale mapping and detection of smaller genetic effects. This will pose many risks and challenges due to the expense of onion population development and maintenance, particularly where traits affecting fitness are segregating. The development of effective PCR-based markers in *Allium*, possibly coupled with new technologies such as whole genome amplification (Lasken and Egholm, 2003), will offer alternative strategies for more rapid and economical development of onion mapping populations without inbreeding. Mapping studies in F₁ populations may be desirable to improve the coverage of more heterozygous genome regions, notably those associated with heterosis. Where phenotypes can be determined on single plants or bulbs with acceptable control over environmental variation (e.g. hydroponic propagation for studies of pungency) a more economical approach will be to evaluate large F₂ or backcross populations and verify findings in replicated families derived from them. Doubled haploid lines (M. Mutschler, *loc. cit*) will offer superior material for developing mapping populations as well as other molecular and functional studies where genetic homogeneity is desirable.

DNA Sequence Resources

Until 2003, onion DNA sequence resources were limited to small sets (200-1000) of ESTs generated from libraries prepared from leaf and seedlings (USDA-ARS), sulfur-starved root (CFR) and bulb tissues (CFR; PRI). Several groups have used homology-based approaches to clone an increasing number of genes associated with economic traits such as carbohydrate metabolism (Vijn et al., 1998), sulfur assimilation (McCallum et al., 2002) and bulb colour (Kramer et al., 2003); (Kim et al., 2004). In 2004 the USDA Asparagales Genome project released a set of 18,484 ESTs from a normalised library prepared from root, bulb and callus tissues of four cultivars (Kuhl et al., 2004). Our recent marker development efforts have focussed on exploiting single nucleotide polymorphisms (SNPs) and simple sequence repeats (SSRs) detected in this data set. A searchable and functionally annotated interface to all public onion sequence data is provided on-line by the Institute for Genome Research (TIGR) in the Onion Gene Index (ONGI; <http://www.tigr.org/tdb/tgi/ongi/>). Marker information is available through our websites (haveylab.hort.wisc.edu and www.crop.cri.nz).

Marker Development

We have successfully developed PCR-based markers from sequences of some cDNAs used to reveal RFLP markers on the BYG15 x AC43 map (King et al., 1998) and several candidate genes associated with carbohydrate and sulfur metabolism, using conventional mutation-scanning techniques including heteroduplex analysis (HAD) and SSCP (McCallum et al., 2001). To date 46% (43/93) of primer sets surveyed have revealed allelic variation in onion. Typically we have designed primer sets that flank conserved intron positions within coding regions or that amplify from coding regions into 3' untranslated regions (3'UTR) of genes. The latter strategy has proven more reliable for

generating single-locus amplification products. SSCP and HDA are technically simple and permit resolution of products of multiple loci but are difficult to reproduce.

We have developed SSR markers by searching for repeats in ONGI EST sequences using MISA scripts (Thiel et al., 2003), followed by automated primer design. This has proven a highly productive means of generating informative markers in onion, similar to the experience in many other crops. To date we have evaluated over 200 primer sets and observed a similar rate of polymorphism within cultivated onion (~40%) to that observed by SSCP. We detect polymorphisms by electrophoresis on denaturing polyacrylamide gels followed by silver staining or by capillary electrophoresis of fluorescently labelled PCR products on automated sequencing platforms. In some cases polymorphisms may be revealed by agarose gel electrophoresis or non-denaturing acrylamide mini-gels. The majority (80%) of primer sets also amplify in *Allium roylei*, confirming the high transferability of EST-SSR markers to other species (Gao et al., 2003).

Genetic Map Alignment

Our primary goal during map development in the 'W202A' x 'Texas Grano 438' population has been to anchor this via co-dominant, portable, PCR-based markers to the BYG15 x AC43 map (King et al., 1998). To date we have scored 40 EST-SSR and 13 SSCP markers, which form 10 linkage groups at LOD 5. By mapping common EST-SSR markers in the BYG15 x AC43 population and using SSCP markers derived from sequences of cDNAs used to reveal RFLP in this population, we have been able to identify linkage groups corresponding to the groups B, C, D, E, H and I on the BYG15 x AC43 map (King et al., 1998). A preliminary consensus map of linkage group B based on these two populations is shown in Fig. 1. The high level of useful polymorphism revealed within these two populations, which were derived from single F₁ families, suggests that EST-derived SSRs will provide an excellent source of co-dominant markers to anchor high-density maps generated by AFLP or similar technologies.

We have scored polymorphisms in the interspecific population in over 20 genes by SSCP or CAPS but to date have insufficient co-dominant markers to reliably place these in the dominant marker maps. A set of alien monosomic addition lines of shallot chromosomes in *A. fistulosum* (Shigyo et al., 1996) were used to assign AFLP markers from the interspecific map to specific *A. cepa* chromosomes (van Heusden et al., 2000a). We anticipate that use of these same genetic stocks will enable assignment of our SSR anchor loci to specific chromosomes and greatly facilitate comparisons between inter- and intra-specific maps.

Genetic Diversity

To date studies of genetic diversity in *Allium* have been conducted by RFLP (Bark, 1995), genetically uncharacterised genomic SSRs (Fischer and Bachmann, 2000) or RAPDs (Bradeen and Havey, 1995). SSRs are the marker of choice in population genetic studies because they represent a defined, single-copy, co-dominant nuclear marker that is amenable to automation and easily distributable. We have commenced an evaluation of EST-derived SSRs to determine their suitability for evaluating genetic diversity within and between cultivated onion populations using a set of 29 New Zealand-adapted open-pollinated varieties and inbred lines. Our observations to date confirm that many SSRs detect three or more alleles in this sample, which is strongly biased toward intermediate and long-day germplasm (Fig. 2). We are currently converting multi-allelic assays to run on high-throughput capillary-based genetic analysis systems to permit efficient, large-scale surveys of within- and between-population genetic variation in germplasm collections.

CONCLUSION

The USDA Asparagales EST sequencing project has provided the *Allium* community with the sequence resources to generate sufficient molecular markers to

provide a strong foundation for rapid advancement of *Allium* genetics. EST-derived SSRs have proved to be a highly productive source of markers in onion, as in many other crops. SNP markers derived from this project will provide an additional, complementary source of co-dominant markers well suited to automated analysis and practical applications such as identity and quality testing. We hope that the availability of portable PCR-based markers will encourage a new wave of exploration of *Allium* genomes and genetic resources. The next challenge facing the community is to develop the populations required to take *Allium* mapping to a higher level of detail and to share genome information for the benefit of the research community and industry.

ACKNOWLEDGEMENTS

This research was funded by the New Zealand Foundation for Research, Science and Technology. We gratefully acknowledge the assistance of Doug Grant (Hybrid Seed Co.) and Paul McCartney (CFR) in developing of the 'W202A' x 'Texas Grano 438' population.

Literature Cited

- Bark, O.H. and Havey, M.J. 1995. Similarities and relationships among populations of the bulb onion as estimated by nuclear RFLPs. *Theor. Appl. Genet.* 90:407-414.
- Bradeen, J.M. and Havey, M.J. 1995. Randomly Amplified Polymorphic DNA in Bulb Onion and Its Use to Assess Inbred Integrity. *J. Am. Soc. Hort. Sci.* 120:752-758.
- Fischer, D. and Bachmann, K. 2000. Onion microsatellites for germplasm analysis and their use in assessing intra- and interspecific relatedness within the subgenus *Rhizirideum*. *Theor. Appl. Genet.* 101:153-164.
- Gao, L., Tang, J., Li, H. and Jia, J. 2003. Analysis of microsatellites in major crops assessed by computational and experimental approaches. *Molec. Breed.* 12:245-261.
- Goldman, I.L. 1996. A List of Germplasm Releases From the University of Wisconsin Onion Breeding Program, 1957-1993. *HortScience* 31:878-879.
- Kim, S., Binzel, M., Yoo, K., Park, S. and Pike, L. 2004. Inactivation of DFR (Dihydroflavonol 4-reductase) gene transcription results in blockage of anthocyanin production in yellow onions (*Allium cepa*). *Molec. Breed.* (in press).
- King, J.J., Bradeen, J.M., Bark, O., McCallum, J.A. and Havey, M.J. 1998. A low-density genetic map of onion reveals a role for tandem duplication in the evolution of an extremely large diploid genome. *Theor. Appl. Genet.* 96:52-62.
- Kramer, C.M., Prata, R.T., Willits, M.G., De Luca, V., Steffens, J.C. and Graser, G. 2003. Cloning and regiospecificity studies of two flavonoid glucosyltransferases from *Allium cepa*. *Phytochemistry* 64:1069-1076.
- Kuhl, J.C., Cheung, F., Yuan, Q., Martin, W., Zewdie, Y., McCallum, J., Catanach, A., Rutherford, P., Sink, K.C., Jenderek, M., Prince, J.P., Town, C.D. and Havey, M.J. 2004. A Unique Set of 11,008 Onion Expressed Sequence Tags Reveals Expressed Sequence and Genomic Differences between the Monocot Orders Asparagales and Poales. *Plant Cell* 16:114-125.
- Lasken, R.S. and Egholm, M. 2003. Whole genome amplification: abundant supplies of DNA from precious samples or clinical specimens. *Trends Biotechnol.* 21:531-535.
- McCallum, J.A., Leite, D., Pither-Joyce, M.D. and Havey, M.J. 2001. Expressed sequence markers for genetic analysis of bulb onion (*Allium cepa* L.). *Theor. Appl. Genet.* 103:979-991.
- McCallum, J.A., Pither-Joyce, M.D. and Shaw, M.L. 2002. Sulfur deprivation and genotype affect gene expression and metabolism of onion roots. *J. Am. Soc. Hort. Sci.* 127:583-589.
- Shigyo, M., Tashiro, Y., Isshiki, S. and Miyazaki, S. 1996. Establishment of a Series of Alien Monosomic Addition Lines of Japanese Bunching Onion (*Allium fistulosum* L.) With Extra Chromosomes From Shallot (*A. cepa* L. Aggregatum Group). *Genes & Genetic Systems* 71:363-371.
- Thiel, T., Michalek, W., Varshney, R.K. and Graner, A. 2003. Exploiting EST databases

- for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* 106:411-422.
- van Heusden, A.W., Shigyo, M., Tashiro, Y., Vrieling-van Ginkel, R. and Kik, C. 2000a. AFLP linkage group assignment to the chromosomes of *Allium cepa* L. via monosomic addition lines. *Theor. Appl. Genet.* 100:480-486.
- van Heusden, A.W., van Ooijen, J.W., Vrieling-van Ginkel, R., Verbeek, W.H.J., Wietsma, W.A. and Kik, C. 2000b. A genetic map of an interspecific cross in *Allium* based on amplified fragment length polymorphism (AFLPTM) markers. *Theor. Appl. Genet.* 100:118-126.
- Vijn, I., Vandijken, A., Luscher, M., Bos, A., Smeets, E., Weisbeek, P., Wiemken, A. and Smeekens, S. 1998. Cloning of Sucrose - Sucrose 1-Fructosyltransferase From Onion and Synthesis of Structurally Defined Fructan Molecules From Sucrose. *Plant Physiol.* 117:1507-1513.

Figures

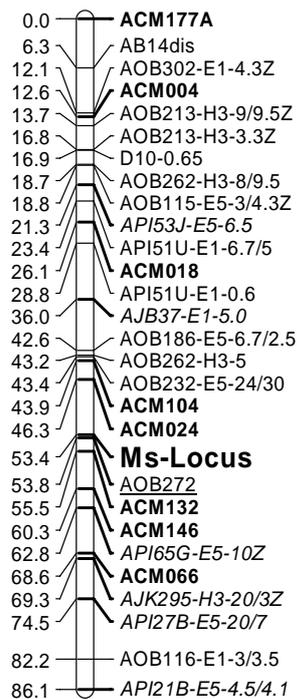


Fig. 1. Consensus linkage map for onion linkage group B (King et al., 1998) based on 'Brigham Yellow Globe 15-23' x 'Alisa Craig 43' and 'W202A' x 'Texas Grano 438' populations. SSR markers are denoted in bold, mutation-scanning markers (SSCP, HDA or CAPS) in italics and SNP markers are underlined.

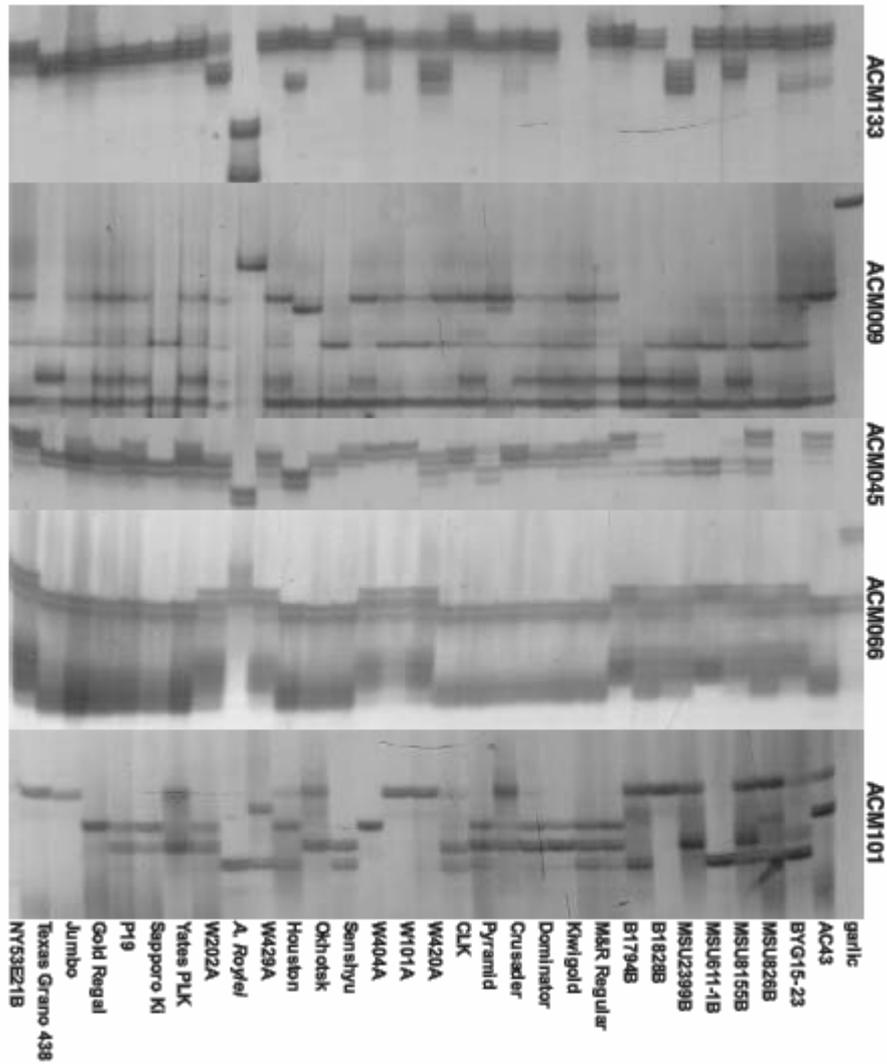


Fig. 2. Allelic variation among 29 onion lines, *Allium roylei* and garlic ('Thermodrome') revealed by 5 EST-SSR markers. PCR products were resolved on 6% denaturing polyacrylamide gels and visualised by silver staining.