Dissection of quantitative traits in oilseed rape for marker development using *Brassica-Arabidopsis* comparative genomics

Rod Snowdon, Maen Hasan, Niclas Freitag, Florin Lipsa, Benjamin Wittkop, Anna Stein, Ana Gloria Badani, Wolfgang Friedt

Department of Plant Breeding, Research Centre for BioSystems, Land Use and Nutrition, Justus Liebig University, Heinrich-Buff-Ring 26-32, 35392 Giessen, Germany
Email: rod.snowdon@agrar.uni-giessen.de

Abstract

The growing resource of *Brassica* genome sequence data, along with the ever-improving annotation of *Brassica* sequences to the extremely well-characterised genome of the model crucifer *Arabidopsis thaliana*, represent an extremely valuable resource for molecular breeding and genetic characterisation of *B. napus*. In particular, the newest comparative genomics data enable a unique opportunity to navigate between and among the chromosomes of *A. thaliana* and *B. napus*, and to compare the map positions of quantitative trait loci (QTL) for complex traits of agronomic importance in the crop species with the positions of potential candidate genes in the model genome. In some cases such rough macrosynteny enables the detection of chromosome blocks in rapeseed corresponding to more or less unbroken syntenic genome regions in *Arabidopsis*. On the one hand this can allow orthologous *Brassica* sequences annotated to the corresponding *Arabidopsis* regions to be used for database-oriented identification of new markers for fine mapping, association studies or marker-assisted selection towards trait improvement. In some cases it is also possible to directly identify potentially relevant candidate genes for important traits in oilseed rape, based on their position in syntenic maps compared to relevant QTL. In other cases it can be feasible to navigate from potential candidate genes in *Arabidopsis* to homoeologous regions of the rapeseed genome, whereby not only the gene sequence itself but the complete surrounding region can be used to search for sequence polymorphisms that can be utilised as markers. In this paper we demonstrate the potential of comparative genomics data for candidate gene identification and marker development in traits related to oilseed rape seed quality. Different approaches are for dissection of traits that appear to be controlled by a relatively small number of major genes (e.g. seed colour and fibre content), and traits influenced by numerous homoeologous QTL containing an unknown number of active genes (e.g. glucosinolate content). In each case new information can be obtained that can lead to new, tightly-linked selection markers and the identification of new allelic diversity for breeding.

Key words: Oilseed rape, *Brassica napus*, seed quality, synteny-based marker development

Introduction

The growing resource of *Brassica-Arabidopsis* comparative genomics data (e.g. Parkin et al. 2005) is constantly improving our ability to navigate between and among the chromosomes of *Arabidopsis thaliana* and *Brassica napus*, and to compare the map positions of QTL for traits of agronomic importance in the crop species with the positions of potential candidate genes in the model genome. This can potentially allow *Brassica* sequences with homology to the corresponding *A. thaliana* regions to be used for database-oriented identification of new markers for fine mapping, association studies or marker-assisted selection towards trait improvement. Moreover, it is also potentially possible to identify relevant candidate genes for important traits in oilseed rape, based on their positions in syntenic maps compared to important QTL. We have used synteny-based comparative mapping to identify new markers with potential linkage to relevant candidate genes for important seed quality traits in oilseed rape, including glucosinolate content and antinutritive fibre compounds and flavonoids associated with seed colour. On the one hand the markers were used for association studies with glucosinolate content to identify useful genetic variation to expand the gene pool of double-low oilseed rape. On the other hand markers with potential linkage to relevant *transparent testa* (TT) genes were mapped in oilseed rape to compare their positions with major QTL involved in seed colour or related traits including contents of lignin and condensed tannins.

Material and Methods

Plant material

Association studies for seed glucosinolate content were performed in 94 genetically diverse genebank accessions, including winter and spring oilseed, fodder and vegetable rape varietie,s that were selected from a *B. napus* "core collection" to represent the genetic diversity resent in European gene bank collections (Lühs et al. 2003, Poulsen et al. 2004, Hasan et al. 2006). The accessions were grown in field trials in Rauischholzhausen, Germany, in 2003/2004. Seeds from 5-6 self-pollinated plants per genotype were harvested and mean total seed glucosinolate content was estimated by near infrared reflectance spectroscopy (NIRS).

Mapping of gene-linked markers potentially involved in seed colour-related traits was performed in a population of 166 doubled-haploid lines derived from a cross between an inbred line of the black-seeded German winter oilseed rape cultivar

'Express' and the true-breeding, yellow-seeded line '1012/98', both with 00-seed quality. QTL involved in seed colour and in the contents of important seed fibre and flavonoid compounds were mapped based on analyses of seeds grown in field trials in Einbeck and Gross Gerau, Germany from 2003-2005.

SSR markers

A set of SSR markers with potential linkage to QTL or candidate genes for the respective traits were developed by comparative mapping between the genomes of *B. napus* and *A. thaliana* making use of genome navigation tools available from The Arabididopsis Information Resource (http://www.arabidopsis.org), Brassica ASTRA database at the Plant Genetics and Genomics platform of Primary Industries Research Victoria, Australia (http://hornbill.cspp.latrobe.edu.au/cgi-binpub/brassica/index.pl) and the *Arabidopsis thaliana* Information Database (http://atidb.org). Arabidopsis sequences upstream and downstream of potential candidate genes or from orthologous sequences to markers from trait QTL in B. napus were scanned for Brassica orthologs containing putative SSR sequences. All identified SSRs were screened in segregating *B. napus* genotypes for associations with the trait of interest or with QTL for corresponding traits.

Results

Glucosinolate-linked markers

Based on comparative mapping of QTL-linked markers four *A. thaliana* genome regions on chromosomes 4 and 5 were identified with potential synteny to major *B. napus* QTL for seed total glucosinolate content. By searching the biological process "glucosinolate biosynthesis" in the gene ontology database of the Arabidopsis Information Resource (TAIR: http://www.arabidopsis.org/) the genes *CYTOCHROME P450 MONOOXYGENASE 83B1* (CYP83B1: At4g31500), *CYTOCHROME P450 79A2* (CYP79A2: At5g05260), *METHYLTHIOALKYLMALATE SYNTHASE* (MAM1: At5g23010) and *ALTERED TRYPTOPHAN REGULATION* (ATR1: At5g60890) were identified as the physically closest potential candidates to the QTL-marker orthologs in the four relevant chromosome regions on *A. thaliana* chromosomes 4 and 5, respectively. For all of the four identified candidate genes we were able to identify putatively physically linked SSR markers with significant association to seed total glucosinolate content (Figure 1).

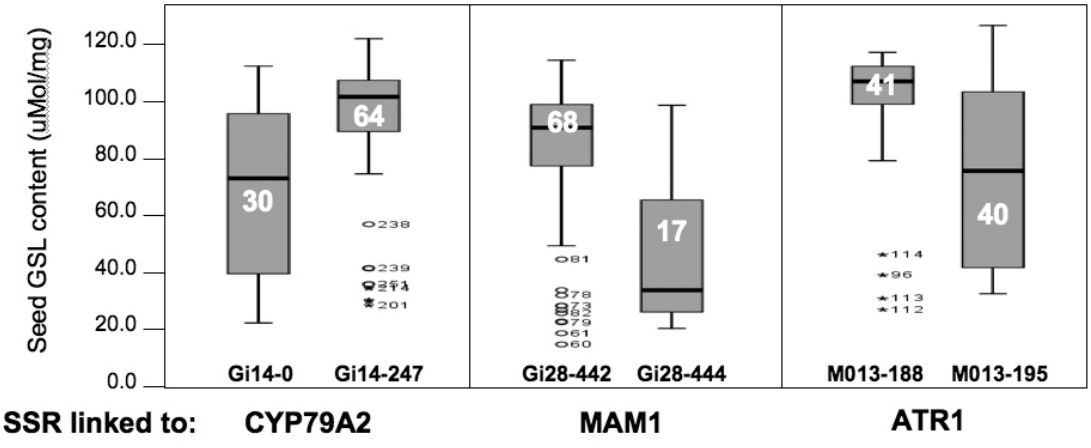

Figure 1: Examples of associations between SSR markers linked to three *A. thaliana* genes involved in glucosinolate biosynthesis with seed total glucosinolate content in a genetically diverse collection of oilseed rape genotypes. Allele frequencies in the set of 94 genotypes are shown in the boxes.

Potential gene-linked markers for TT candidate genes

By synteny-based *in silico* mapping we were able to identify a number of potential *Brassica* markers whose orthologs in *A. thaliana* show close physical linkage to the endothelium-specific genes TT1, TT10 and TT16. Re-mapping of some of these markers in *B. napus* enabled us to establish putative linkage between some of these markers with QTL involved in seed colour and related traits, including lignin content and relevant flavonoid compounds (Figure 2). This provides a strong indication that these genes may be involved in the expression of yellow seed colour and the corresponding reduction in antinutritive fibre and condensed tannins.

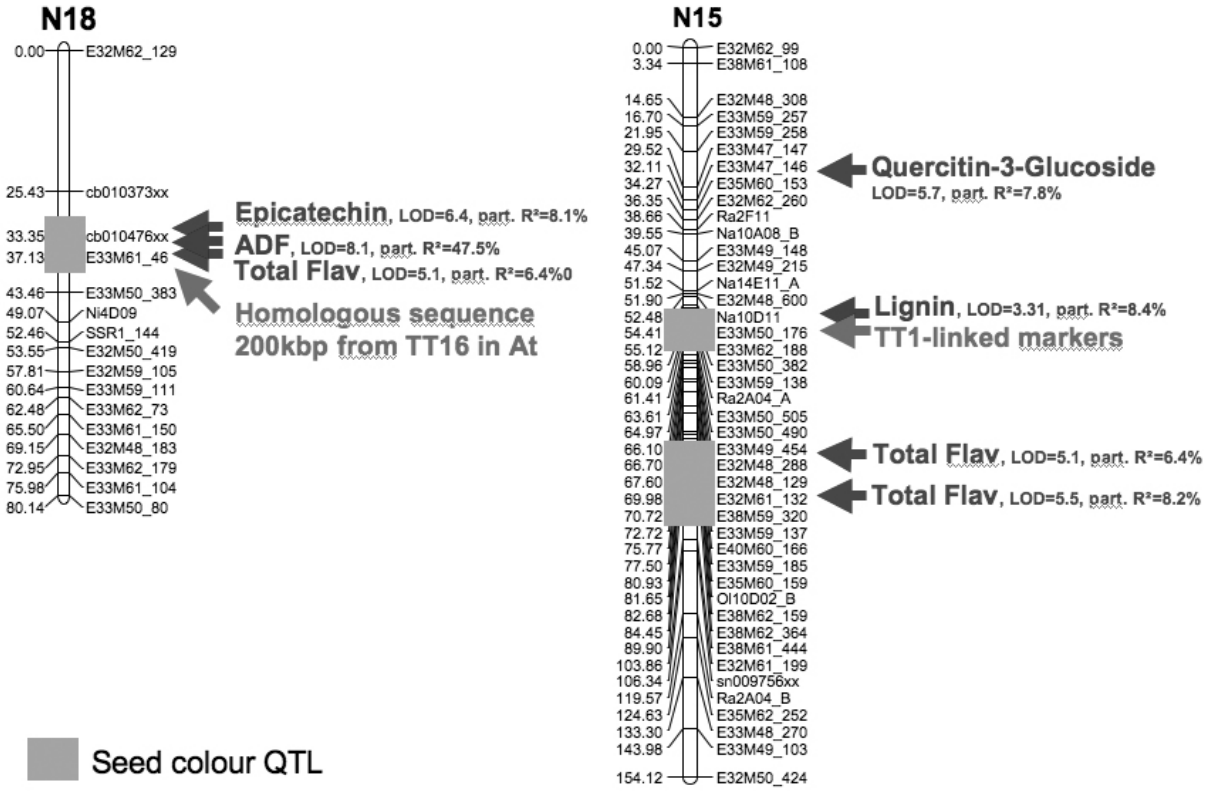

Figure 2: Localisation of potentially gene-linked markers for the candidate genes TT1 and TT16 in comparison to major QTL involved in seed colour, lignin and flavonoid content.

Discussion

The results of this study demonstrate the potential of *Arabidopsis-Brassica* comparative genome analysis for synteny-based identification of gene-linked markers that can be used in marker-assisted selection for important traits in oilseed rape. On the one hand association studies with gene-linked SSR markers gave strong indications that genetically linked homologous copies of a small number of key biosynthetic and regulatory genes are the major factors involved in accumulation of aliphatic, aromatic and indole glucosinolates in *B. napus* seeds. By identifying gene-linked SSR markers with significant associations to total seed glucosinolate content in genetically diverse oilseed rape germplasm we hope to provide a simple molecular tool for marker-assisted combination of positive alleles in new, low-glucosinolate genotypes. On the other hand we were able to find markers linked to potential candidate genes for seed colour traits in Arabidopsis that appear to be linked to corresponding QTL in *B. napus*.

Conclusions

With an ever-growing resource of *Brassica* sequence data it is today becoming increasingly possible – despite the complex rearrangements among *Brassica* genomes – to annotate, align and compare chromosomal and genomic data between *Brassica* and *Arabidopsis*, and to use this new information for genomic studies in the comparatively large *B. napus* genome. Navigation between *Brassica* and *Arabidopsis* physical maps using published genome annotation and synteny data uncovers an enormous wealth of tools for fine-mapping, synteny-based gene cloning and marker development for marker-assisted selection. If such markers are in linkage disequilibrium with the gene of interest, this strategy can be extremely useful for indirect mapping of candidate genes on *Brassica* chromosomes. The approach also has the potential to enable map localisation of duplicated copies of a given candidate gene for identification of major genes underlying important trait QTL in oilseed rape.

Acknowledgements

Part of the work described here was performed with funding from the Federal Ministry of Education and Research (BMBF) within the collaborative research projects GABI-CompGen and GABI-CGAT: YelLowSin with support from the commercial partners KWS Saat, Deutsche Saatveredelung, Norddeutsche Pflanzenzucht and Saatenunion Resistanzlabor.

References

Hasan M., Seyis F., Badani A.G., Pons-Kuhnemann J., Lühs W., Friedt W. and Snowdon R.J. 2006. Analysis of Genetic Diversity in the *Brassica napus* L. Gene Pool Using SSR Markers Genet. Resour. Crop Evol. 53: 793–802.

Lühs W., Seyis F., Frauen M., Busch H., Frese L., Willner E., Friedt W., Gustafsson M. and Poulsen G. 2003. Development and evaluation of a *Brassica napus* core collection. In: Knüpffer H. and Ochsmann J. (eds.), Rudolf Mansfeld and Plant Genetic Resources. Proceedings of a symposium dedicated to the 100th birthday of Rudolf Mansfeld, Gatersleben, Germany, 8–9 October 2001. Schriften zu Genetischen Ressourcen 19. ZADI/IBV, Bonn, pp. 284–289. (http://www.genres.de/infos/rei-bd22.htm).

Parkin I.A., Gulden S.M., Sharpe A.G., Lukens L., Trick M., Osborn T.C. and Lydiate D.J. 2005. Segmental structure of the *Brassica napus* genome based on comparative analysis with *Arabidopsis thaliana*. Genetics 171: 765-781.

Poulsen G., Busch H., Frauen M., Frese L., Friedt W., Gustafsson M., Ottosson F., Seyis F., Stemann G., Ulber B., Willner E., and Lühs W. 2004. The European *Brassica napus* core collection – Characterisation, evaluation and establishment. Cruciferae Newsl. 25: 115-116.