Analysis of Synteny between Rapeseed (*B. napus* L.) and *Arabidopsis thaliana*

Wolfgang Ecke

Institute of Agronomy and Plant Breeding, University of Göttingen, Von-Siebold-Str. 8, 37075 Göttingen, Germany

Introduction

As part of the German plant genome project GABI we are developing two series of intervarietal substitution lines. These lines will be used to localize and fine map QTL for agronomically important traits of rapeseed. With this background the synteny (Fig. 1) between *Arabidopsis thaliana* and rapeseed was analysed to provide a means for the identification of candidate genes for QTL of rapeseed by utilizing the information available from *Arabidopsis* genome research. To test the feasibility of this approach it was analysed whether it would have been possible to identify a candidate gene for a known gene, the gene controlling erucic acid synthesis, a long chain fatty acid in the seed oil of many *Brassicaceae*.

Materials and Methods

Synteny analysis was based on an RFLP map of the rapeseed genome comprising 214 RFLP markers, 35 RAPD markers and one phenotypic marker (Uzunova et al. 1995). For synteny analysis 155 RFLP probes corresponding to 175 RFLP markers were sequenced (MWG-Biotech AG). Loci in the *Arabidopsis* genome homologous to the RFLP markers mapped in rapeseed were identified by comparing the probe DNA sequences with an *Arabidopsis* DNA sequence database extracted from the EMBL database. For databank comparisons the program BlastN was used.

Synteny:

Colinearity of homologous genetic loci on chromosomes of related species

Fig. 1 Definition of synteny

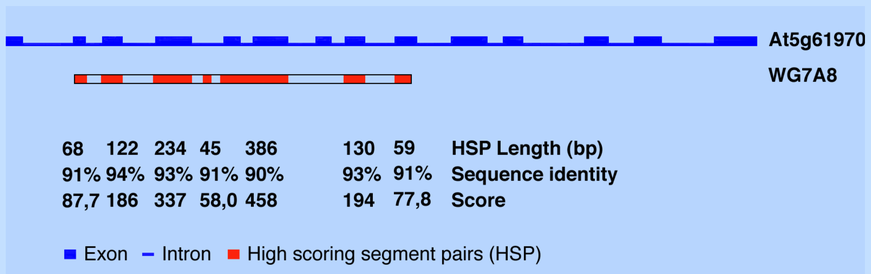

Fig. 2 Schematic sequence alignment of RFLP probe WG7A8 with At5g51970

The horizontal bars represent *Arabidopsis* gene At5g61970 and rapeseed RFLP probe WG7A8. Sequence identities in the high scoring segment pairs (HSP), that were the initial results of the BlastN search, range from 90% to 94%. The HSPs closely correspond to exons of the *Arabidopsis* gene. Sequence identities are lower in introns, resulting in a total sequence identity of only 80.1% over the full length of the probe (1952 bp).
The representation of At5g51970 was reprinted from MIPS *Arabidopsis thaliana* database.

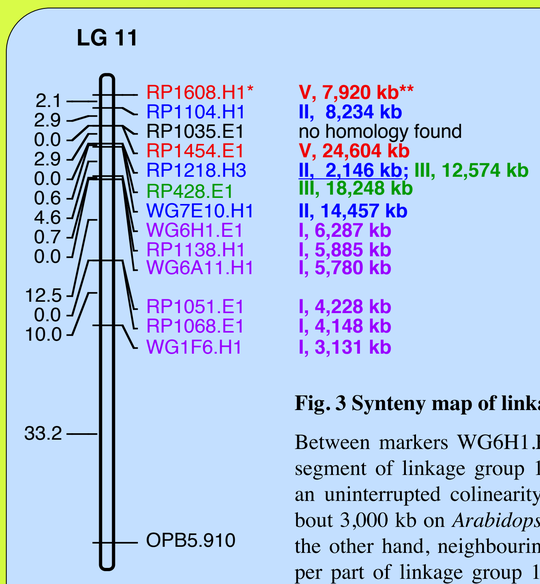

Fig. 3 Synteny map of linkage group 11

Between markers WG6H1.E1 and WG1F6.H1 a segment of linkage group 11 of 23.2 cM shows an uninterrupted colinearity with a region of about 3,000 kb on *Arabidopsis* chromosome I. On the other hand, neighbouring markers in the upper part of linkage group 11 are homologous to loci on different *Arabidopsis* chromosomes, indicating a small scale synteny due to a large number of rearrangements.

*Markers with prefixes RP and WG are RFLP markers, markers with Prefix OP are RAPD markers. **The roman numerical indicates the *Arabidopsis* chromosome, the figure the position of the locus homologous to the corresponding RFLP marker on the sequence map of the chromosome. When probes showed significant homology to two loci the locus with higher sequence identity is underlined.

Fig. 4 Genetics of erucic acid synthesis (C22:1) in *Arabidopsis* and rapeseed

- In *Arabidopsis* erucic acid synthesis is controlled by the *FAE1* locus on chromosome IV

  Lemieux B, Miquel M, Somerville C, Browes J (1990)
  Mutants of *Arabidopsis* with alterations in seed lipid fatty acid compositions.
  Theor Appl Genet 80:234-240

- The *FAE1* locus in *Arabidopsis* encodes a β-keto-acyl-CoA synthetase

  James DW, Lim E, Keller J, Plooy I, Ralston E, Dooner HK (1995)
  Directed tagging of the *Arabidopsis* fatty acid elongation 1 (*FAE1*) gene with the maize transposon activator.
  Plant Cell 7:309-319

- In the amphidiploid rapeseed erucic acid synthesis is controlled by two loci, *eru1* and *eru2*, that have been mapped on linkage groups 6 and 12

  Ecke W, Uzunova M, Weißleder K (1995)
  Mapping the genome of rapeseed (*Brassica napus* L.). II. Localization of genes controlling erucic acid synthesis and seed oil content.
  Theor Appl Genet 91:972-977

- The two erucic acid genes of rapeseed are homologous to the *FAE1* gene of *Arabidopsis*

  Fourmann M, Barret P, Renard M, Pelletier G, Delourme R, Brunel D (1998)
  The two genes homologous to *Arabidopsis FAE1* co-segregate with the two loci governing erucic acid content in *Brassica napus*.
  Theor Appl Genet 96:852-858

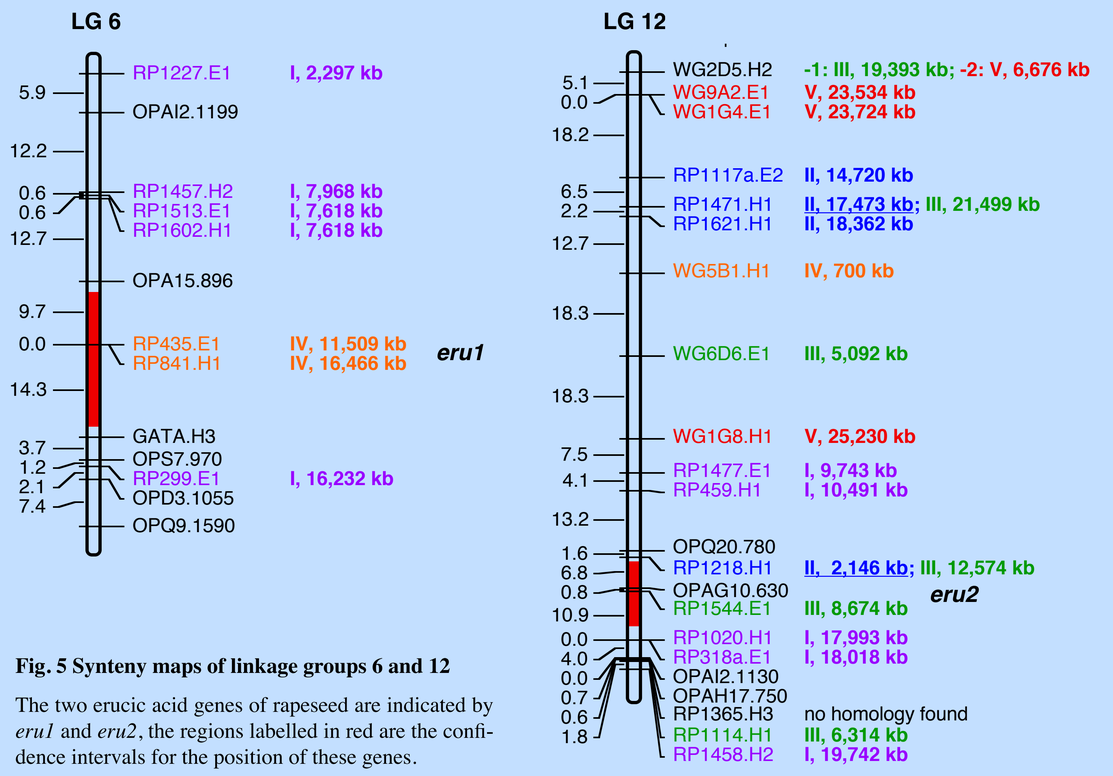

Fig. 5 Synteny maps of linkage groups 6 and 12

The two erucic acid genes of rapeseed are indicated by *eru1* and *eru2*, the regions labelled in red are the confidence intervals for the position of these genes.

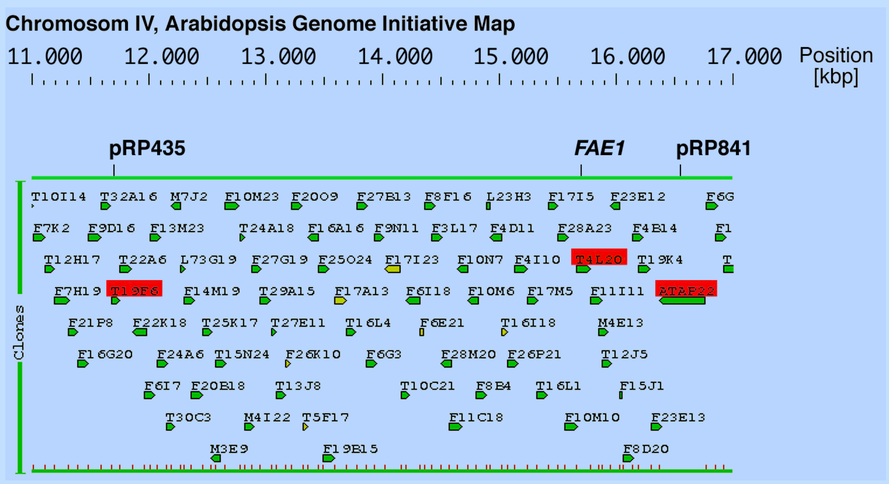

Fig. 6 Physical map of *A. thaliana* chromosome IV in the vicinity of the *FAE1* gene

A section of the physical map of *Arabidopsis* chromosome IV is represented as an ordered sequence of the clones that constitute the map. The clone T4L20 that contains the *FAE1* gene and flanking clones T19F6 and ATAP22 that contain loci homologous to RFLP probes pRP435 and pRP841 are labelled red. pRP435 and pRP841 have shown markers linked to the erucic acid gene on linkage group 6.
The map is a modified reprint from "The Arabidopsis Information Resource" (TAIR) databank

Results

In exons sequence identities between *Arabidopsis* and rapeseed DNA sequences were highly significant, ranging from 85% - 95% (Fig. 2). With few exceptions sequence identities were lower in introns and intergenic sequences. Since some of the probes showed homology to two loci in the *Arabidopsis* genome, reflecting internal duplications in *Arabidopsis*, a total of 162 homologous loci could be identified for 139 of the sequenced probes. The remaining 16 did not show any significant homology with *Arabidopsis* sequences.

Synteny between *Arabidopsis* and rapeseed ranges from a small scale synteny to extended synteny blocks encompassing large parts of a linkage group (Fig. 3). Most of the linkage groups are composed of a mosaic of segments with synteny to different parts of the *Arabidopsis* genome (Fig. 3 & 5), indicating a large number of chromosomal rearrangements since *Arabidopsis* and rapeseed diverged from a common ancestor.

The genetic control of erucic acid synthesis in *Arabidopsis* and rapeseed is outlined in Fig. 4. Fig. 5 shows the synteny maps of linkage groups 6 and 12 where the erucic acid genes of rapeseed were mapped. Only on linkage group 6 are markers linked to the erucic acid gene which have homologous loci on chromosome IV of *Arabidopsis* that carries the *FAE1* locus. On the physical map of chromosome IV these loci flank the *FAE1* gene (Fig. 6), indicating that it would have been possible to identify *FAE1* as a candidate gene for the erucic acid gene on linkage group 6 of rapeseed based on the synteny analysis.

Conclusions

- Due to the high sequence identities in coding regions of the *Arabidopsis* and rapeseed genomes a synteny analysis using sequenced RFLP probes is feasible.
- The synteny relationship of *Arabidopsis* and rapeseed is characterized by a high number of chromosomal rearrangements. Nevertheless, there are many segments where colinearity extends over several cM up to tens of cM.
- Based on the synteny analysis an identification of candidate genes for genes and QTL genetically mapped in rapeseed is possible. In individual cases the success of this approach will depend on whether the gene resides in a segment with extended synteny or in a region with small scale synteny. In the latter case the specific set of RFLP markers linked with the gene will determine success or failure of the approach as is evident in the example of the erucic acid genes.

Literature

Uzunova M, Ecke W, Weißleder K, and Röbbelen G (1995) Mapping the genome of rapeseed (*Brassica napus* L.). I. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theor Appl Genet 90:194–204