An EU project on gene flow analysis between crop and wild forms of lettuce and chicory in the context of GMO biosafety: first results in lettuce

Clemens van de Wiel[1], Gerard van der Linden[1], Hans den Nijs[2], Andrew Flavell[3], Naeem Syed[3], Rikke Jorgensen[4], Francois Felber[5], Ivan Scotti[6], Jeroen Rouppe van der Voort[7] and Johan Peleman[7]

[1]Plant Research International B.V., P.O. Box 16, 6700 AA Wageningen, The Netherlands.
[2]University of Amsterdam, Institute for Biodiversity & Ecosystem Dynamics, P.O. Box 94062, 1090 GB Amsterdam, The Netherlands.
[3]University of Dundee, School of Life Sciences, Division of Biochemistry, MSI/WTB Complex, Dundee, Scotland DD1 4HN, United Kingdom.
[4]Risø National Laboratory, Plant Biology and Biogeochemistry Department, P.O.Box 49, 4000 Roskilde, Denmark.
[5] Université de Neuchâtel, Laboratoire de Botanique évolutive, P.O. Box 2, CH-2007 Neuchâtel, Switzerland.
[6]University of Udine, Department of Crop Science and Agricultural Engineering, Via delle Scienze 208, 33100 Udine, Italy.
[7]Keygene N.V., Molecular Marker Services, P.O. Box 216, 6700 AE Wageningen, The Netherlands.
Contact: clemens.vandewiel@wur.nl, (+31) 317 476981.

Abstract: An EU project is presented that aims at detecting gene flow between crop and wild forms of lettuce and chicory and possible consequences for the ecology of the wild forms in the light of GMO biosafety assessment. Two novel molecular marker systems, the retrotransposon-based SSAP and the disease resistance gene-based NBS-profiling, were successfully developed for testing their ability to trace introgression between crop and wild. Preliminary data on the testing of AFLP® for this purpose in lettuce are presented and discussed.

Keywords: *Lactuca sativa*, *Lactuca serriola*, lettuce, *Cichorium intybus*, chicory, molecular markers, AFLP, SSAP, NBS-profiling, gene flow, introgression, biosafety, GMO.

Introduction

The public fear for the impact of genetically modified crops on the natural environment triggered a steady stream of research. Among the possible impacts, the "escape" of the transgene, either through dispersal of the crop plant outside of agricultural area or through hybridization with wild relatives, attracted a lot of attention, in particular in relation to the possibility of increasing "weediness". Thus, a lot of studies appeared on gene flow and hybrid fitness, but they are for the larger part focussed on a few large arable crops, i.e. oilseed rape (*Brassica napus*), e.g. Chevre et al. 2000, Hauser et al. 1998, Rieger et al. 2002, and, to a lesser extent, on beet (*Beta vulgaris*), e.g. Desplanque et al. 1999, Bartsch et al. 2001. Few publications appeared on other crops, such as barley (Savova Bianchi et al. 2002) and carrot (Hauser & Bjorn 2001).

In several European countries, reports have been issued summarising knowledge on escape and hybridisation between crop species and their feral allies from herbaria and the floristic literature, the so-called botanical files. An example for The Netherlands was the study by De Vries et al. (1992), which indicated that, particularly in the leafy vegetables,

there are examples of crops for which detailed information on gene flow is lacking, for instance, the Asteracean species chicory (*Cichorium intybus*) and lettuce (*Lactuca sativa*). The latter species' closest wild relative, prickley lettuce (*L. serriola*) expanded its occurrence in North Western Europe during the last decades. An in-depth botanical files study indicated both species, *L. sativa* and *L. serriola*, to be conspecific, based on morphology and crossability (Frietema de Vries et al. 1994). They are regarded as self-pollinating crops. However, there is a variable amount of cross-pollination, as indicated by the occurrence of small percentages of heterozygotes (unpublished observations C. van de Wiel). *L. serriola*'s recent expansion together with its potential for hybridization with cultivated lettuce makes it an excellent model for studying "weediness" in relation to interactions with the crop.

In order to obtain more insight into gene flow in chicory and lettuce, an EU project started in 2001, named "Analysis of gene flow from crop to wild forms in lettuce and chicory and its consequences in the context of GM-crop biosafety" (acronym "ANGEL", QLK3_CT-2001-01657). Its objectives are: a) To trace evidence of introgression from cultivated to wild forms using several molecular marker techniques comprising both neutral markers and markers linked to traits that may affect fitness, e.g. disease resistance genes; b) To establish the degree of outcrossing under field conditions in wild forms using molecular marker techniques; c) To study the consequences of gene flow from cultivated to wild forms by field trials, by demographical monitoring and by modelling of both natural and experimental crop-wild hybrid populations; d) To obtain insight in the recent invasiveness into NW Europe of *L. serriola* by characterising populations along a north-south and an east-west transect through Europe using molecular markers. In the following, preliminary results from the first year are presented on tracing of introgression in lettuce.

Results and discussion

Molecular marker development for introgression tracing

Three molecular marker methods are tested for their ability to trace introgression between crop and wild forms. They comprise the established multi-locus fingerprinting technique AFLP® (Vos et al. 1995), and novel methods, one based on retrotransposons (SSAP: Sequence-Specific Amplified Polymorphism) and the other on conserved parts of disease resistance genes (NBS-profiling: Nucleotide Binding Site-directed profiling). The adaptation of the latter techniques for lettuce is first described, then preliminary results from AFLP work are presented.

The first of the novel marker methods is based on natural transposable elements, which are occurring ubiquitously in plant genomes. What makes the transposable element-based marker technology most suitable for introgression studies is the underlying mechanism of directional genomic changes over time. A retrotransposon stays at its place and at the same time can create a new insertion site somewhere else in the genome through an RNA intermediate. New insertion sites are thus unique (derivative) characters in phylogenetic terms that can be used reliably to establish relationships (e.g. Tatout et al. 1999). By the same token, these characters can be used as an indication of introgression. This variation can be accessed by various methods, from which Sequence-Specific Amplified Polymorphism (SSAP) was adopted for the species in this project. In SSAP, sites in the genome where a transposable element has inserted are detected by PCR using one primer derived from the transposable element, facing outward. The other primer is an adapter primer (similar to those used for AFLP), thus a primer based on a linker ligated to genomic DNA that has been digested by a specific restriction enzyme (Waugh et al. 1997). SSAP, like AFLP, is capable of generating efficiently a large amount of markers distributed over the whole genome.

The limiting step in the development of the marker system is the availability of terminal sequences from retrotransposons (Long Terminal Repeats: LTRs), which are used to design primers for the SSAP PCR reactions. To identify the LTR sequences, the following approaches were employed: a) chromosome walking from the well conserved RnaseH motif found close to the LTRs of Ty1-*copia* group of retrotransposons, by using degenerate primers; b) chromosome walking from retrotransposon sequences mined from lettuce EST libraries (information obtained from R. Michelmore, UC Davis, USA). Both approaches led to the identification of LTR sequences that were all tested for their performance in SSAP. The target genome was cut with a rare restriction enzyme cutter (in this case *Pst*I) and a frequent cutter enzyme (*Mse*I) and corresponding *Pst*I and *Mse*I oligonucleotide adapters were ligated. After the pre-amplification step, using adapter specific primers, the SSAP selective PCR was performed using one of the adapter primers and ^{33}P labeled LTR primer to generate the SSAP marker profiles on a sequencing gel.

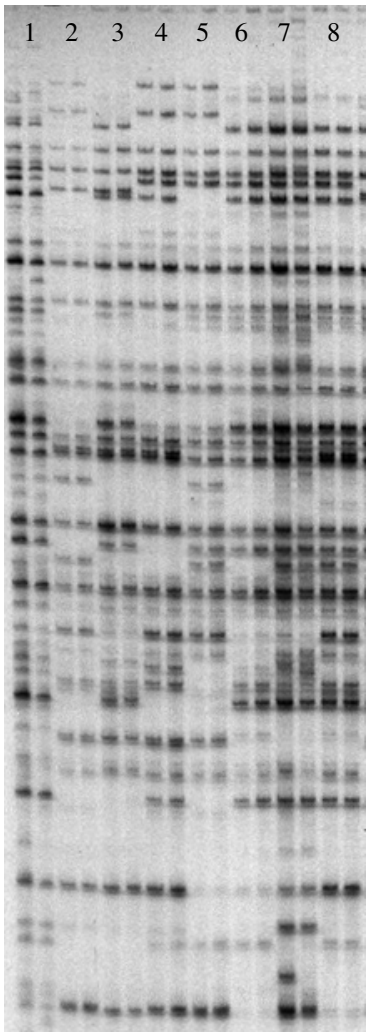
The SSAP protocol was further optimised using a few samples from wild as well as cultivated accessions of lettuce. A combination of two selective bases attached to the *Mse*I adapter primer and one base added to the LTR primer in the selective amplification step gave clear and, for the great majority of bands, reproducible SSAP marker profiles for duplicate DNA samples (see Figure 1a). The relatively low number of selective bases (as opposed to a usual number of three each) indicated a comparatively low copy number of these retro-elements (*gypsy* and *copia*) as compared to wheat and maize. Interestingly, the CO9-*gypsy* primer isolated from lettuce also gave excellent SSAP marker profiles for chicory.

The NBS profiling approach screens for variation within functional regions of the genome, i.e. disease resistance genes (Van Tienderen et al 2002). These genes are exploited intensively in the breeding process. For this, primers are used that are derived from published sequences of NBS (nucleotide binding site)-containing disease resistance genes, the most important group of resistance genes described so far. One primer is based on the NBS, the other is an adapter-based primer. A PCR using such primer combination will result in a multilocus marker profile with a high content of RGA (Resistance Gene Analogues)-like sequences. A set of ten *Lactuca sativa* cultivars and four accessions of other *Lactuca* species (two *L. serriola*, one *L. saligna* and one *L. virosa* accession) were used to develop and validate the NBS-profiling approach for lettuce. NBS-profiling as a technique has been developed to work in a number of crops with similar primers and enzymes, without modifications. Based on known *Lactuca* resistance gene analogue sequences, several of the universal primers were likely to work with *Lactuca* and recognize most of the known RGAs. Therefore, two of these universal primers were tested in the test set. NBS-profiling was performed on lettuce genomic DNA digested with *Rsa*I. This resulted in a banding pattern that was enriched for RGA-like sequences. The samples were run in duplicate (DNA split and the procedure performed twice) to check for reproducibility. For all primers, the duplicate samples were identical, indicating that the procedure is highly reproducible in lettuce. As expected, the banding patterns of *L. serriola* were most similar to those of the lettuce cultivars, whereas the other wild species showed quite distinct patterns. Each of the profiles generated between 60-90 bands, of which about half were polymorphic in the *L. sativa* cultivar set, and even more were polymorphic between species (Figure 1b).

To get an indication of the RGA content, a number of bands from each profile were sequenced. Several bands could be positively identified as putative RGAs. These included *RGC2* candidates of which one has been identified as the *Dm3* gene (Meyers et al.1998). A substantial number of bands (more than 50%) had no significant similarity to any of the sequences in the database (using standard settings for the XBLAST and the NBLAST programs). This may be due to the relative scarcity of RGA sequences from members of the Asteraceae in the databases. Thus, these sequences may actually be RGAs, but their sequence similarity to the relatively few known RGAs is too low to allow positive identification. However, it cannot be excluded that the

frequency of RGAs in the lettuce NBS5-profiles is indeed not higher than somewhere between 20-50%.

a. SSAP



b. NBS profiling

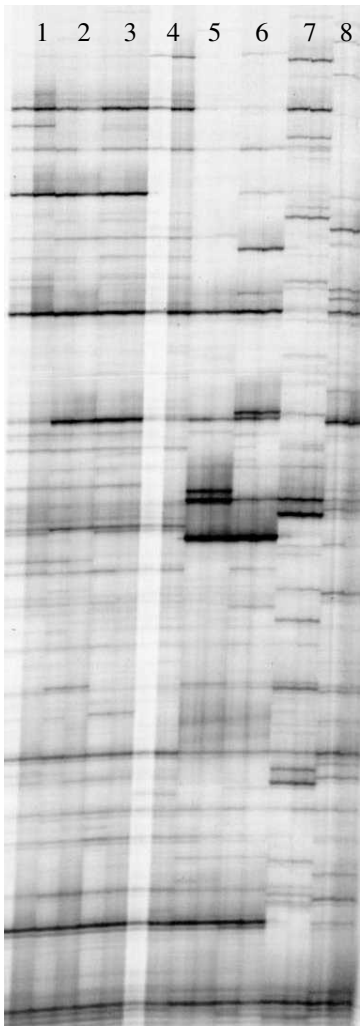


Figure 1. Novel molecular marker methods in lettuce: a) SSAP, lanes 1 and 2 are *L. sativa* and *L. serriola*, respectively, and the other lanes are RILs derived from them; b) NBS profiling, lane 1 is cv. Monet, 2 cv. Great Lakes, 3 cv. Karif, 4 Tianjin Big Stem, 5 *L. serriola* CGN04667, 6 idem CGN15684, 7 *L. saligna* CGN15697, 8 *L. virosa* CGN13339. All lanes are in duplicate to test reproducibility of individual bands; in 1a, lanes 7 lower part, just one significantly intense band appears not to be reproducible between the duplicates.

Introgression tracing by AFLP

The SSAP and NBS profiling will be implemented in trying to trace introgression between crop and wild forms. For this purpose, AFLP® was already explored in more detail on three wild populations, from northern Germany (LS17), southern Germany (LBM1) and southern Italy (LBS1), respectively. A comparative analysis was carried out between molecular marker genotypes generated from 32 *L. serriola* individuals from these three populations and 10 *L. sativa* individuals representing a broad range of variation within the cultivated species. They were fingerprinted by use of 11 AFLP® primer combinations (PCs). Subsequently, the

samples were genotyped on the basis of presence/absence polymorphisms of AFLP markers (bands) and the data were aligned with the Keygene integrated lettuce map. This integrated map consists of four independently generated linkage maps and encompasses approximately 1000 AFLP markers. On the integrated map, markers were identified as 'anchor' markers when they approximated, in the same order, a similar chromosomal position on two or more genetic linkage maps. Depending on the marker density of the alignment with the integrated map, this analysis should reveal the genomic locations where the genetic differences between *L. sativa* and *L. serriola* reside. In addition, 'strings' of identical markers closely linked on the integrated map ("haplotypes") could be identified among the samples, using Keygene haplotyping software called MACP (Marker Assisted Chromosome Painter, for details see the contribution of Peleman & Rouppe van der Voort in these proceedings). In this way, *L. sativa*-specific chromosome segments should be identified, which upon recovery in *L. serriola* would be a stronger indication for introgression from the crop than stray individual markers.

A total of 299 out of 749 markers scored in the three *L. serriola* populations and the reference *L. sativa* samples were recovered in the integrated map. The markers were ordered according to their position on the chromosome and a hypothetical map was composed based on the distances from the integrated map. The total distance covered by all markers was 1093.3 cM, 95% of the cM distance of the integrated map. Based on genetic distances, the lettuce samples could be classified roughly in three groups: a) one group of *L. sativa* samples (first 8 samples); b) one group of *L. sativa / L. serriola* intermediates (subsequent 2 samples); c) one group of *L. serriola* samples of which the three different populations could be clearly distinguished (dendrogramme not shown). A first haplotyping analysis showed one segment that could be found both in *L. sativa* and in one of the *L. serriola* populations. It is at present hard to tell whether this segment originated from a recent introgression from the cultivated form to the wild population. This is partly due to the relatively low marker density attained on the integrated map for such analysis: on a total number of 1153 markers comprising the integrated map, only 299, with an average cM interval of 3.7 cM, could be recovered in the populations analysed. An extension of the integrated lettuce map is planned for extending analyses of the data.

Acknowledgements

The authors thank Dr. R. Michelmore (UCDavis, USA) for providing transposable element sequences from lettuce.

References

Bartsch, D. Brand, U. Morak, C. Pohl Orf, M. Schuphan, I. and Ellstrand, N. C. 2001. Biosafety of hybrids between transgenic virus-resistant sugar beet and swiss chard. Ecological Applications 11: 142-147.

Chevre, A. M. Eber, F. Darmency, H. Fleury, A. Picault, H. Letanneur, J. C. and Renard, M. 2000. Assessment of interspecific hybridization between transgenic oilseed rape and wild radish under normal agronomic conditions. Theoretical and Applied Genetics 100: 1233-1239.

Desplanque, B. Boudry, P. Broomberg, K. Saumitou Laprade, P. Cuguen, J. and Van Dijk, H. 1999. Genetic diversity and gene flow between wild, cultivated and weedy forms of *Beta vulgaris* L. (Chenopodiaceae), assessed by RFLP and microsatellite markers. Theoretical and Applied Genetics 98: 1194-1201.

Frietema de Vries, F.T. Van der Meijden, R. and Brandenburg, W.A. 1994. Botanical files on Lettuce *Lactuca sativa* - On the chance for gene flow between wild and cultivated lettuce *Lactuca sativa*

L. including *L. serriola* L. Compositae and the generalized implications for risk-assessments on genetically modified organisms. Gorteria Suppl 2: 44pp.
Hauser, T.P. Jorgensen, R.B. and Østergard, H. 1998. Fitness of backcross and F2 hybrids between weedy *Brassica rapa* and oilseed rape *B. napus*, Heredity 81: 436-443.
Hauser, T.P. and Bjørn G.K. 2001. Hybrids between wild and cultivated carrots in Danish carrot fields. Genetic Resources and Crop Evolution 48: 499-506.
Meyers, B.C. Chin, D.B. Shen, K.A. Sivaramakrishnan, S. Lavelle, D.O. Zhang, Z. and Michelmore, R.W. 1998. The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant cell 10: 1817-1832.
Peleman, J.D. and Rouppe van der Voort, J. 2003. Towards breeding by design. Pp. 125-130 *in* Hintum, Th.J.L. van, A. Lebeda, D.A. Pink and J. Schut (eds.) Eucarpia Leafy Vegetables 2003. Centre for Genetic Resources, The Netherlands (CGN), Wageningen, The Netherlands
Rieger, M.A. Lamond, M. Preston, C. Powles, S.B. and Roush, R.T. 2002. Pollen-mediated movement of herbicide resistance between commercial canola fields. Science 296: 2386-2388.
Savova Bianchi, D. Keller Senften, J. and Felber, F. 2002. Isozyme variation of *Hordeum murinum* in Switzerland and test of hybridization with cultivated barley. Weed Research 42: 325-333.
Tatout, C. Warwick, S. Lenoir, A. and Deragon, J.M. 1999. SINE insertions as clade markers for wild crucifer species. Molecular Biology and Evolution 16: 1614-1621.
Van Tienderen, P.H. De Haan, A.A. Van der Linden, C.G. and Vosman, B. 2002. Biodiversity assessment using markers for ecologically important traits. Trends in Ecology and Evolution 17: 577-582.
Vos, P. Hogers, R. Bleeker, M. Reijans, M. Van der Lee, T. Hornes, M. Frijters, A. Pot, J. Peleman, J. Kuiper, M. and Zabeau, M. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414.
De Vries, F.T. Van der Meijden, R. Brandenburg, W.A. 1992. Botanical files. A study of the real chances for spontaneous gene flow from cultivated plants to the wild flora of the Netherlands. Gorteria Suppl 1: 100pp.
Waugh, R. McLean, K. Flavell, A.J. Pearce, S.R. Kumar, A. Thomas, B.B.T. and Powell, W. 1997. Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms S-SAP. Molecular and General Genetics 253: 687-694.