UTILIZATION OF A RESTRICTION FRAGMENT LENGTH POLYMORPHISM LINKAGE MAP IN RAPESEED BREEDING

Y.-S. Chyi, M. Hoenecke, and L. Sernyk

Agrigenetics Company-Madison Laboratory
5649 E. Buckeye Rd., Madison, WI 53716, USA

INTRODUCTION

Among emerging molecular technologies, restriction fragment length polymorphism (RFLP) has imminent application to plant breeding (Murray et al. 1990; Tanksley et al. 1989) by enabling inspection of an individual plant's genotype over its entire genome using mapped RFLP markers. This method is environment-independent, economic in terms of time, space, and labor, and precise in genotypic information obtained. And, since RFLP monitors natural genetic recombination, it stirs little regulatory concern compared to genetic transformation.

The most immediate application of marker-based selection is to backcross breeding, where one or two desirable traits are introduced to further improve established genotypes. By assaying at the DNA level rather than by conventional phenotypic measurement, the breeder can select individuals with the highest recurrent genome content in each backcross generation, thus producing the final product in a substantially shorter amount of time and with greater precision (Paterson et al. 1988; Ma et al. 1990). Results from utilizing RFLP technology to facilitate conventional backcross breeding of rapeseed at Agrigenetics Company are reported here.

MATERIALS AND METHODS

Genomic DNA from leaf tissue of the canola cultivar "Westar" (*Brassica napus*) was digested with the restriction endonuclease PstI, and fragments of 500-3000 base pairs were cloned into the plasmid vector pGEM-7Zf(+) (Promega). The library was screened by colony hybridization (Sambrook et al. 1989) for low-copy-number-sequence clones.

The Agrigenetics *Brassica rapa* RFLP linkage map is being constructed from an F_2 population of two commercial cultivars (AGC code numbers BR0002 and BR0019) and has been used at various stages of completion. One F_2 plant selected for two desirable single-locus traits (one from each parental line) was backcrossed to the recurrent parent (BR0002). Fifty F_2BC_1 progeny plants selected for the same characters were subjected to RFLP analysis with 50 genomic clones (Fig. 1). The 50 genomic clones, used as probes, detected 66 RFLP loci covering 550 map units. Twelve F_2BC_1 individuals with varying recurrent parent genome content were chosen and backcrossed again. The F_2BC_2 progenies were selected for the same traits and 37 of them were subjected to RFLP fingerprinting with 96 genomic clones, including the initial 50. The 96 clones detected 116 marker loci with about 1150 map unit coverage of the *B. rapa* genome (Fig. 1).

The RFLP fingerprinting data were converted into graphic representations of the genome constitution of each progeny. The estimated percent genomic content of each parental line was computed with programs developed at Agrigenetics Company.

RESULTS AND DISCUSSION

The AGC B. rapa RFLP linkage map was constructed de novo based on segregation data obtained from an F_2 population. Within the time span of one breeding generation in the greenhouse, i.e., before the F_2BC_1 progeny produced from a selected F_2 individual bloomed, 80 RFLP marker loci had been placed on 10 independent linkage groups (Fig. 1). The genome coverage of these markers was about 550 map units. When this map was used to analyze the genomic constitution of 50 selected F_2BC_1 progeny, individuals with estimated recurrent parent genome content between 64.9% and 88.2% were observed. The estimated mean parent genome content for all 50 plants was 74.5%.

In a succeeding breeding generation, the total number of mapped RFLP marker loci increased to 200. These markers were distributed on 10 major linkage groups plus two unassigned "segments". Together they covered about 1150 map units of the B. rapa genome. When the newer version of the map was used to analyze the genomic constitution of the same 50 F_2BC_1 plants, the range of estimated recurrent parent genome content became 64.1% - 82.3%, with the mean at 73.4%. For each individual, the two values of recurrent parent genome content estimated from the earlier subset and the later set of marker loci correlated quite well (r^2 = 0.75, Fig.2). The individuals with high recurrent parent genome content selected for advancement in backcross breeding using either version of the map would therefore be the same. The two mean values of the estimated recurrent parent genome content matched quite well with the expected value of 75% for a random F_2BC_1 population, even though the population was subjected to selection for traits from both parents.

Twelve F_2BC_1 individuals were advanced to the next backcross generation. The recomputed mean recurrent parent genome content for these 12 individuals was 75.0%, with a wider spread (standard error was 4.8%, vs. 4.0% for the set of 50). Ten seedlings were raised from each of the 12 F_2BC_2 progenies and were re-selected for the same traits from the parental lines. Various numbers of individuals re-selected out of each progeny formed a sample for the F_2BC_2 generation (Table 1). The estimated recurrent parent genome content in these individuals ranged from 73.2% to 92.8%. The average recurrent parent genome content weighted to adjust for varying progeny sizes was 84.8%, close to the expected 87.5% for a random F_2BC_2 population. We wish to note that in this study the F_2BC_1 individuals selected came from both ends of the distribution (Fig. 3), whereas in applied breeding using RFLP, only the individuals with the highest recurrent parent genome content would be chosen.

Fig. 4a and 4b showed the genomic constitution of C192, one of the F_2BC_1 individuals with high recurrent parent genome content, and R015, the F_2BC_2 progeny with highest recurrent parent genome content. R015 was derived from C192. Some of the crossing over points can be deduced from comparing the two genomic constitution graphs.

These results demonstrate that RFLP provides a means to accurately assess actual percent recovery of the recurrent genome in backcross progenies. A simple comparison will illustrate the advantage of utilizing RFLP fingerprinting in facilitating recurrent genome recovery in a backcross breeding project. Assuming no selection and random crossing-over, each

backcross will bring the recurrent parent genome content of the previous generation half way towards 100%, on the average. When an F_2BC_1 plant with high recurrent parent genome content (C192, Table 1) was chosen for advancement, the selected progeny showed above-average recovery. On the other hand, when an F_2BC_1 individual with low recurrent parent genome content (C178) was used for backcrossing, only two of the eight selected progeny achieved even the expected recovery which was approximately that of the high F_2BC_1 individual. Without the assistance of RFLP fingerprinting, there was a 50% chance that the individual with low recurrent parent genome content (C178 in this example) would have been used. The fact that trait selection was practiced in each generation might interfere with the recovery of recurrent parent genome at later generations, but does not affect the fairness of the comparison cited above.

The application of RFLP fingerprinting to select for recurrent parent genome content in this study resulted in BC_2 progeny that were at a BC_3 level of recovery. This amounts to a 25% saving in breeding time and expense, using only a small population in each generation.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Drs. M. Murray and Y. Ma for protocol development of Agrigenetics RFLP techniques and PCR of genomic library clones, respectively. Also, we extend our gratitude to Agrigenetics RFLP team members for general assistance in probe labelling and blot handling. Finally, sincere appreciation is offered to the several reviewers, especially Drs. D. West and J. Andrews, for their suggestions.

REFERENCES

MA Y., WEST D., CRAMER J.H., ROMERO-SEVERSON J., KIRSCHMAN J., PITAS J., DeMARS S., VILBRANDT L., and MURRAY M.G. 1990. "Near-Isogenic Line" revisited. (manuscript submitted)

MURRAY M.G., CHYI Y.S., CRAMER J.H., DeMARS S., KIRSCHMAN J., MA Y., PITAS J., ROMERO-SEVERSONJ., SHOEMAKER J., WEST D.P., and ZAITLIN D. 1990. Application of restriction fragment length polymorphism to maize breeding. NATO ASI on Plant Molec Biol 1990 (in press).

OSBORN T.C., SONG K.M., KENNARD W.C., SLOCUM M.K., FIGDORE S., SUZUKI J., and WILLIAMS P.H. 1990. Genome analysis in Brassica using RFLPs. NATO ASI on Plant Molec Biol 1990 (in press).

PATERSON A.H., LANDER E.S., HEWITT J.D., PETERSON S., LINCOLN S.E., and TANKSLEY S.D. 1988. Resolution of quantitative traits into Mandelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721-726.

TANKSLEY S.D., YOUNG N.D., PATERSON A.H., and BONIERBALE M.W. 1989. RFLP mapping in plant breeding: New tools for an old science. Bio/Tech 7: 257-264.

SAMBROOK J., MANIATIS T., and FRITSCH E.F. 1989. Molecular Cloning. Cold Spring Harbor Laboratory. pp. 1.90-1.104.

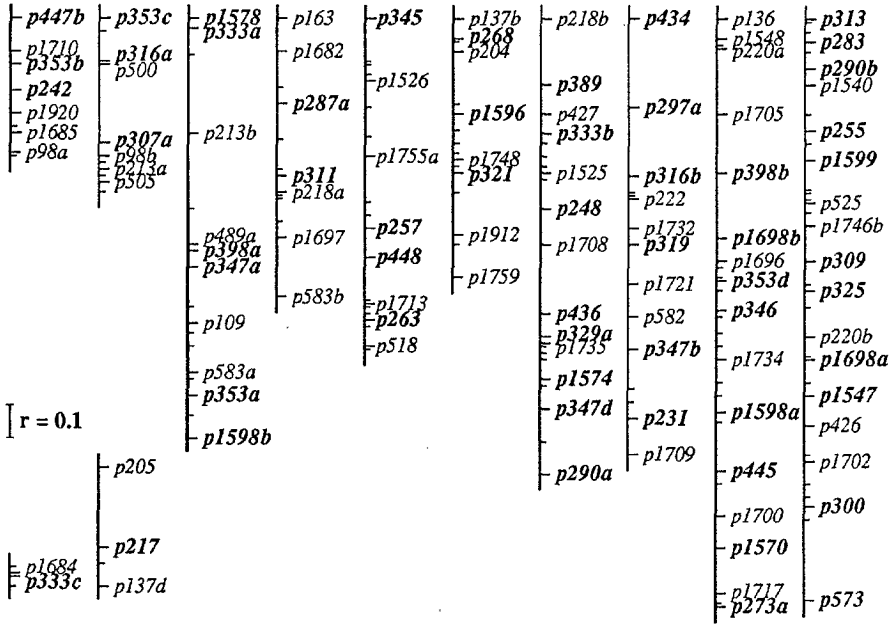

Fig. 1. Agrigenetics Brassica rapa RFLP Linkage Map
(Marker loci analyzed are shown. Bold-face markers were included in the later analysis, when the revised map covered an estimated 75% of the genome.)

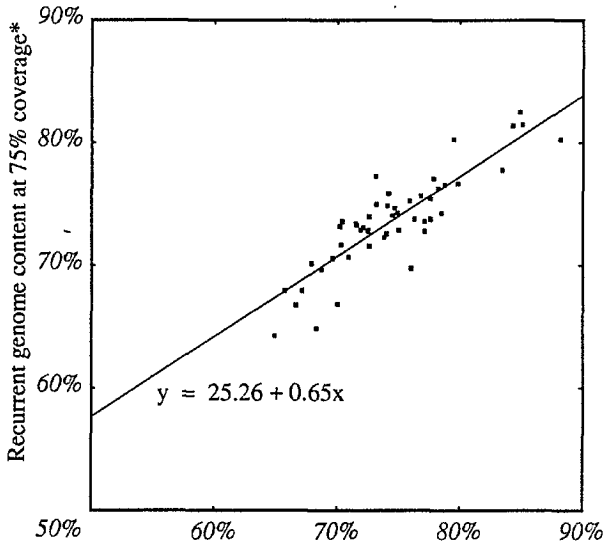

Fig. 2. Comparison of Estimated Recurrent Genome Content of F_2BC_1 Individuals Using Partial RFLP Linkage Maps
**Based on a hypothetical 1500-unit map.*

Table 1. Recurrent parent genome content of selected backcross progenies
(The F_2BC_2 (R) progeny follow their corresponding F_2BC_1 (C) parents.)

C178	64.4%	C197	75.8%	C247	73.6%
R131	84.0%	R021	91.6%	R083	82.2%
R132	73.2%	R025	85.9%	R089	79.0%
R133	80.4%	R026	88.5%		
R134	80.5%			C251	71.4%
R136	78.6%	C206	72.3%	R104	81.8%
R138	81.3%	R031	81.5%	R105	76.4%
R139	77.2%			R106	82.8%
R140	82.1%	C215	73.3%	R108	79.9%
		R043	88.7%	R110	82.4%
C191	78.8%	R044	84.1%		
R004	89.3%	R046	83.7%	C255	72.0%
R008	83.6%	R047	84.5%	R116	84.5%
R009	92.5%	R049	80.9%		
R010	89.1%	R050	89.2%	C256	72.8%
				R123	83.4%
C192	81.1%	C230	75.3%	R126	85.9%
R015	92.8%	R064	85.9%	R130	86.1%
R019	91.3%				
		C231	79.9%		
		R074	85.6%		

Weighted mean for the R progenies 84.8%

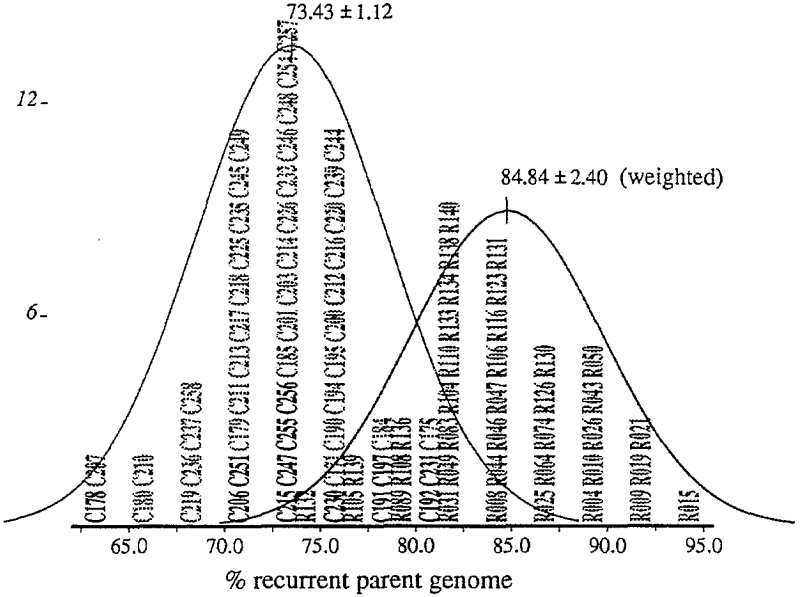

Fig. 3. Distribution of Genomic Constitution of F_2BC_1 (the C's) and F_2BC_2 (the R's) Populations
(Individuals used for advancement are indicated by bold-face type.)

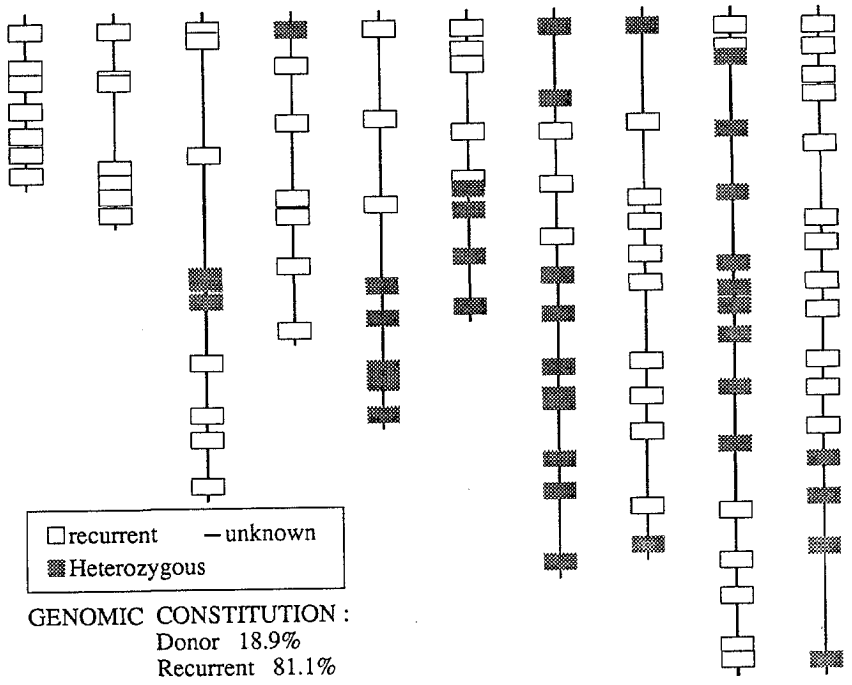

Fig. 4a. RFLP Genotype for F_2BC_1 Individual C192

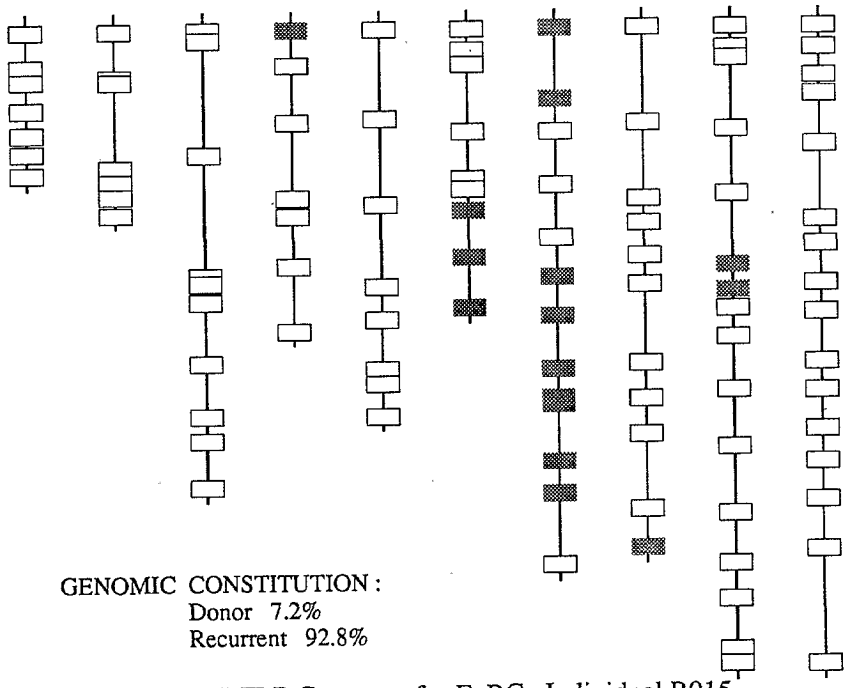

Fig. 4b. RFLP Genotype for F_2BC_2 Individual R015