

Genetics of aliphatic glucosinolates. I. Side chain elongation in *Brassica napus* and *Arabidopsis thaliana*

R. MAGRATH*, F. BANO†, M. MORGNER*, I. PARKIN*, A. SHARPE*, C. LISTER‡, C. DEAN‡, J. TURNER†, D. LYDIATE* AND R. MITHEN*§

**Brassica* and Oilseeds Research Department, John Innes Centre, Institute of Plant Science Research, Colney, Norwich NR4 7TJ, UK, †School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK, ‡Molecular Genetics Department, John Innes Centre, Institute of Plant Science Research, Colney, Norwich NR4 7UJ, UK.

Aliphatic glucosinolates are thioglucosides synthesized by genera within the Capparales including *Arabidopsis* and *Brassica*. They have been shown to mediate pest or pathogen interactions and to reduce the feeding quality of rapeseed meal. Their biological activity is largely dependent upon the structure of the side chain which determines the nature of hydrolysis products produced following tissue damage. A generalized model of the biosynthesis of aliphatic glucosinolates is proposed and tested with reference to the genetic regulation of side chain length by analysing recombinant populations of *B. napus* and *A. thaliana* developed from parental lines which varied in their glucosinolate content. The results of the *B. napus* studies were consistent with a model in which alleles at a single locus (*Gsl-pro*) regulate the presence or absence of propyl glucosinolates, and those at two other loci (*Gsl-elong-C* and *Gsl-elong-A*) (which map onto a pair of homoeologous linkage groups, one from the C genome and one from the A genome) regulate side chain elongation of the amino acid derivative which results in the production of butyl and pentyl glucosinolates. As null alleles do not occur at the *Gsl-elong-A* locus, the proportion of propyl glucosinolates is limited to below 30 per cent of total aliphatic glucosinolates. Alleles at a single locus in *A. thaliana* (*Gsl-elong-Ar*), which maps 1.3 cM from the RFLP locus *m291* on chromosome 5, regulate side chain elongation of aliphatic glucosinolates in this species. It is suggested that the *Gsl-elong-Ar* gene is homologous to the *Gsl-elong-A* and *Gsl-elong-C* genes and, if cloned, could be used to down-regulate the endogenous *Gsl-elong* genes in *B. napus*.

Keywords: *Arabidopsis*, *Brassica*, gene mapping, glucosinolates.

Introduction

Aliphatic glucosinolates are secondary metabolites occurring in *Brassica*, *Arabidopsis* and other genera in the Capparales. They comprise a common glycone moiety with a variable aglycone side chain (Fig. 1) derived from methionine. Following tissue damage, aliphatic glucosinolates undergo hydrolysis to produce an array of products dependent upon the structure of the side chain. Several reviews on the biochemistry of glucosinolates have been published (e.g. Underhill *et al.*, 1973; Fenwick *et al.*, 1983; Poulton & Moller, 1993). However, despite the economic and biological importance of these secondary metabolites, there have

been few studies on the genetic regulation of their biosynthesis, particularly with regard to side chain elongation and modification.

Within *Brassica*, the aliphatic glucosinolates comprise alkenyl, hydroxyalkenyl, methylthioalkyl and methylsulphinylalkyl homologues of propyl, butyl and pentyl glucosinolates (Table 1). However, oilseed rape, and other forms of the amphidiploid species *B. napus*, have a restricted and uniform profile comprising butenyl and pentenyl glucosinolates and their hydroxylated homologues. In a specific tissue, such as mature leaves or seeds, the ratio of the different classes of aliphatic glucosinolates is constant. A previous study reported on the development of synthetic lines of *B. napus* from its diploid parents which contained aliphatic glucosinolates that do not occur in natural forms of *B. napus*,

§Correspondence.

and the use of these lines to investigate the genetic regulation of glucosinolate biosynthesis (Magrath *et al.*, 1993). A genetic model of aliphatic glucosinolate biosynthesis, which is partially tested within the present study, is shown in Fig. 2.

Aliphatic glucosinolate profiles vary between wild-type accessions of *A. thaliana* and in leaves comprise alkenyl, methylthioalkyl, methylsulphinylalkenyl and hydroxyalkyl homologues of propyl and butyl glucosinolates, with trace levels of other homologues with longer side chain lengths (Hogge *et al.*, 1988; Haughn *et al.*, 1991). By analogy with studies in *Brassica*, a model can be developed that explains the variation in aliphatic glucosinolates on the basis of variation in

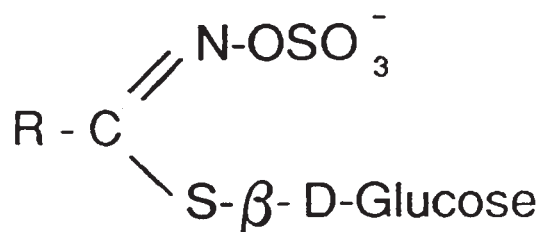


Fig. 1 Molecular structure of glucosinolates. For details of side chain ('R') see Fig. 2.

alleles at four loci which determine side chain length and subsequent side chain modifications (Fig. 2).

In the present study, the model of the regulation of side chain length of aliphatic glucosinolate shown in Fig. 2 is tested by investigating the segregation of glucosinolate phenotypes within populations of recom-

Table 1 Major aliphatic glucosinolates occurring in *Brassica* and *Arabidopsis*

Systematic name	Trivial name
3-Methylthiopropyl	Glucobeverin
3-Methylsulphinylpropyl	Gluciberin
2-Propenyl	Sinigrin
3-Hydroxypropyl	
4-Methylthiobutyl	Glucoerucin
4-Methylsulphinylbutyl	Glucoraphanin
3-Butenyl	Glucanapin
2-Hydroxy-3-butenyl	Progoitrin
3-Hydroxybutyl	
5-Methylthiopentyl	Glucoberteroin
5-Methylsulphinylpentyl	Glucualyssin
4-Pentenyl	Glucobrassicinapin
2-Hydroxy-4-pentenyl	Glucanapoleiferin

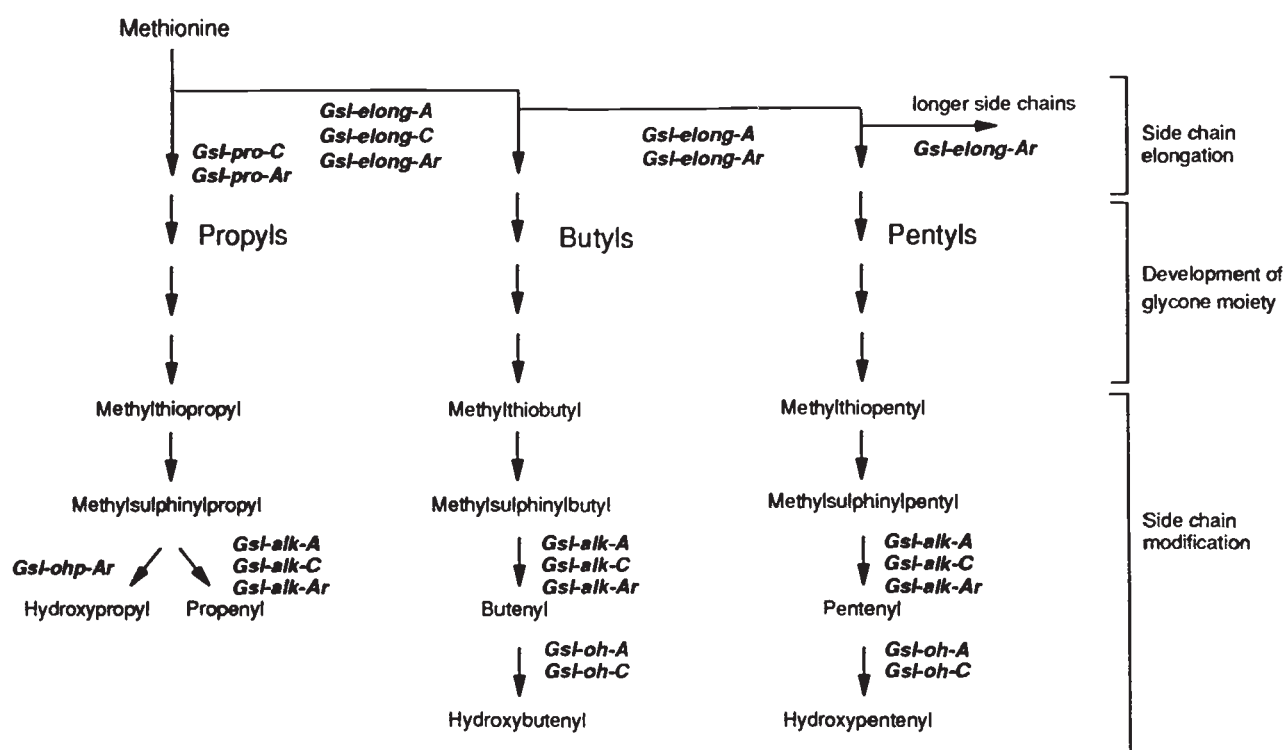


Fig. 2 Hypothetical model of aliphatic glucosinolates biosynthesis in *Brassica napus* and *Arabidopsis thaliana*. Alkenyl glucosinolates may be derived from methylthioalkyl glucosinolates (Chisholm & Matsuo, 1972).

Table 2 Genotypes of diploid parents, synthetic *Brassica napus* and *B. napus* cv Cobra at putative *Gsl* loci (see Fig. 2 for more details)

	C genome				A genome		
	<i>pro-C</i>	<i>elong-C</i>	<i>alk-C</i>	<i>oh-C</i>	<i>elong-A</i>	<i>alk-A</i>	<i>oh-A</i>
<i>B. atlantica</i>	++	--	++	--			
<i>B. rapa</i>					++*	++	--
Synthetic <i>B. napus</i>	++	--	++	--	++*	++	--
<i>B. napus</i> 'Cobra'	--	++*	++	++	++†	++	++

+, functional allele; -, null allele.

*Alleles result in production of butyl glucosinolates.

†Alleles result in production of butyl and pentyl glucosinolates.

Table 3 Genotypes of Columbia and Landsberg *erecta* at four putative *Gsl* loci (see Fig. 2 for more details)

	<i>pro-Ar</i>	<i>elong-Ar</i>	<i>alk-Ar</i>	<i>ohp-Ar</i>
Columbia	++	++	--	--
Landsberg <i>erecta</i>	++	--	--	++

binant lines of *B. napus* and *A. thaliana* developed from crosses between parental lines with contrasting glucosinolate phenotypes (Tables 2 and 3). For the *Brassica* studies, we consider recombinant homozygous lines developed by the culture of microspores from an F₁ hybrid between the cultivar Cobra and a synthetic *B. napus* (Table 2). In *A. thaliana* we consider a population of recombinant inbred (RI) lines derived from an F₁ hybrid between the ecotypes Columbia and Landsberg *erecta* (Lister & Dean, 1993) (see Table 3).

Materials and methods

Development of *B. napus* recombinant homozygous lines

The synthetic *B. napus* line was developed from the interspecific hybridization of *B. atlantica* (a wild member of the *B. oleracea* *n*=9 complex) and a wild form of *B. rapa*. The synthetic line was crossed to the oilseed rape cultivar Cobra (Magrath *et al.*, 1993). Microspores were isolated by macerating surface sterilized buds collected from an F₁ hybrid plant with a glass rod in a few drops of Gamborgs B5-13 medium and filtered through a nylon membrane. The microspores were washed twice by centrifugation at 900 rpm for 3 min and resuspended in fresh B5-13. Following further centrifugation, the microspores were resuspended in Nitsch and Nitsch's NLN-13 medium (Sigma Ltd), and placed in the dark for 24 h at 32°C. This was

followed by further centrifugation and resuspension in fresh NLN-13. The microspore suspension was diluted to the equivalent of 6 buds per 5 ml of medium to provide a concentration of approximately 2×10^4 microspores ml⁻¹. The suspension was aliquoted into small petri dishes, sealed and returned to 32°C in the dark for a further 48 h. When the embryos had grown enough to be seen by the naked eye, they were placed on a slow shaker at 25°C in the dark for 4–6 weeks and then transferred to solid B5 medium at 20°C in a 16 h photoperiod. The embryos were repeatedly sub-cultured until they had regenerated into rooted plantlets, when they were transferred to the glasshouse and grown in John Innes No. 1 compost.

When flower buds could be seen, the apical meristem was removed and a 1 per cent colchicine solution was applied to the leaf axils and allowed to dry. When inflorescences developed that appeared to be fertile because of fully developed stamens and viable pollen (determined by staining with a 4 per cent acetocarmine solution), open flowers were self-pollinated and the inflorescence enclosed within a polythene bag into which CO₂ was injected to raise the internal concentration to above 5 per cent to overcome self-incompatibility. The bag was removed after 72 h. Haploid plants were vegetatively propagated.

B. napus DNA extraction and Southern hybridization

Approximately 500 mg of freeze-dried milled leaf tissue was added to 15 ml of Kirby mix [1 per cent sodium tri-isopropyl naphthalene sulphonate (Eastman Kodak), 6 per cent sodium 4-amino salicylate (BDH), 50 mM Tris-HCl pH 8.0, 6 per cent buffer saturated phenol], vortexed for several seconds and gently mixed on a shaking platform for 30 min. The preparation was extracted with 10 ml of phenol/chloroform solution and vortexed for 5–10 s before mixing on a platform for 5 min. The samples were

centrifuged at 3000 rpm for 10 min at room temperature to partition the aqueous phase from the organic phase and debris. Nucleic acids in the aqueous phase were precipitated by adding 0.1 volume of 3 M sodium acetate (pH 6.0) followed by 0.6 of the resulting volume of isopropanol, and stood at ambient temperature overnight. The precipitate was centrifuged as before and the pellet allowed to air-dry for 30 min. The dried pellet was resuspended in 2 ml of 1 × TE (10 mM Tris, 1 mM EDTA pH 8.0) plus 40 µg ml⁻¹ RNase, and incubated in a shaking incubator at 37°C. The preparation was phenol/chloroform extracted and precipitated once more. The pellet was air-dried and resuspended in 0.25 ml of 1 × TE, and transferred to an Eppendorf tube. DNA was digested with *EcoRI* and the concentration of DNA estimated by fluorometric analysis. DNA restriction fragments were separated by electrophoresis and capillary-blotted onto a nylon filter (Hybond N+).

Filters containing the transferred DNA were pre-hybridized by incubating for 4 h at 65°C in a Techné hybridization oven in 30 ml of pre-hybridization solution [50 × Denhardt's (1 per cent Ficoll 400, 1 per cent polyvinylpyrrolidone, 1 per cent BSA fraction V), 20 × SETS (3 M NaCl, 20 mM EDTA pH 8.0, 0.6 M Tris-HCl pH 8.0, 11 mM tetra-sodium pyrophosphate), 10 per cent SDS, 1 per cent Dextran sulphate plus 50 µg ml⁻¹ of freshly boiled herring testes DNA (Sigma D-6898)]. The prepared probe was boiled for 10 min and added to 5 ml of hybridization solution (50 × Denhardt's, 20 × SETS, 0.1 per cent SDS, 10 per cent Dextran sulphate plus 50 µg ml⁻¹ of freshly boiled herring testes DNA) to hybridize the DNA. The hybridization was carried out overnight at 65°C with gentle agitation in the hybridization oven. After hybridization the filters were washed twice at 65°C for a total of 1 h. The first wash contained 2 × SSC, 0.1 per cent SDS and the second 0.2 × SSC, 0.1 per cent SDS. After washing, the filters were wrapped in cling film and exposed to X-ray film at -70°C for 7–12 days.

Filters comprising DNA from 54 recombinant homozygous and parental diploid and amphidiploid lines were probed with 58 anonymous gDNA probes, derived from libraries of small (0.8–1.6 kb) *PstI* fragments from *B. napus*, *B. oleracea* and *B. rapa*. Full details of the development of the probes will be published separately. The filters were additionally probed with 10 anonymous cDNA *Brassica* probes. Linkage data were analysed with the aid of the MAPMAKER programs (Lander *et al.*, 1987). Recombination frequencies were converted to map distances using the Kosambi mapping functions (Kosambi, 1944).

Development of A. thaliana lines and RFLP mapping

The production and characterization of the 100 recombinant inbred lines used in this study is described by Lister and Dean (1993). Briefly, F₂ progeny from a cross between Landsberg *erecta* (as female) and Columbia were selfed and taken by single seed descent to the F₈ generation. Bulk F₉ seed from individual F₈ plants were collected and sown to provide leaf material for glucosinolate analyses. The segregation profile was compared to that of 67 loci mapped using 64 RFLP markers. Linkage analyses were undertaken as described previously (Lister & Dean, 1993). Glucosinolates were also extracted from individual F₂ plants grown under similar conditions.

Glucosinolate analysis

Leaf tissues from parental and recombinant lines were grown under standard glasshouse conditions, harvested and the glucosinolates extracted and analysed as described by Magrath *et al.* (1993). For the *A. thaliana* lines, glucosinolates were extracted from leaves of a bulk sample of between 10 and 20 plants per line. Glucosinolates were also extracted from *Brassica* microspore-derived plants grown under field conditions in September 1992 and from diploid tissue from plants grown from seed obtained from genotypes which were successfully doubled. HPLC analyses were as described previously. Glucosinolates from *B. napus* plants were eluted with 50 per cent acetonitrile and those from *Arabidopsis* with 20 per cent acetonitrile. The percentages of aliphatic glucosinolates with propyl, butyl and pentyl side chains were calculated as described by Magrath *et al.* (1993).

Results

Characterization of parental and recombinant homozygous lines of B. napus

The synthetic *B. napus* line contained propenyl and butenyl glucosinolates in contrast to the *B. napus* cultivar which contained butenyl, pentenyl, hydroxybutenyl and hydroxypentenyl glucosinolates. The F₁ hybrid possessed alkenyl and hydroxyalkenyl glucosinolates of each of the three side chain lengths except hydroxypropenyl glucosinolates, which do not occur in *Brassica*. Trace amounts of methylsulphinylpentyl glucosinolate were detected in the *B. napus* cultivar and the F₁ hybrid.

The levels of propenyl glucosinolates (calculated as a percentage of total aliphatic glucosinolates) in the recombinant lines were in two classes corresponding to

the parental phenotypes. No other glucosinolates with propyl side chains were detected. Twenty-three lines had less than 3 per cent propenyl glucosinolate (i.e. trace levels), whereas 31 lines had between 7 and 30 per cent propenyl glucosinolate (Figs 3 and 4). The ratio of the two classes is not significantly different from 1:1 ($P > 0.1$), which is consistent with the presence of a major gene (designated *Gsl-pro*) determining the presence or absence of propenyl glucosinolate.

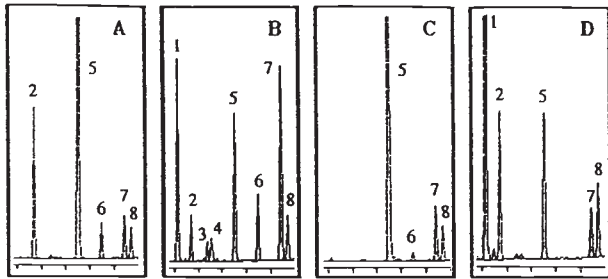


Fig. 3 HPLC profiles of desulphoglucosinolates from leaves of (A) synthetic *Brassica napus*, (B) F_1 hybrid between the synthetic *B. napus* and *B. napus* cv Cobra, (C) recombinant line with zero hydroxylation, zero propenyl glucosinolate and low levels of pentenyl glucosinolate, (D) recombinant line with high hydroxylation, high propenyl glucosinolate and low pentenyl glucosinolate. 1, hydroxybutenyl; 2, propenyl; 3, hydroxypentenyl; 4, methylsulphinylpentenyl; 5, butenyl; 6, 4-hydroxymethylindolyl; 7, pentenyl; 8, benzyl (retention time is 9.7 min, internal standard). Other indolyl desulphoglucosinolates have later retention times.

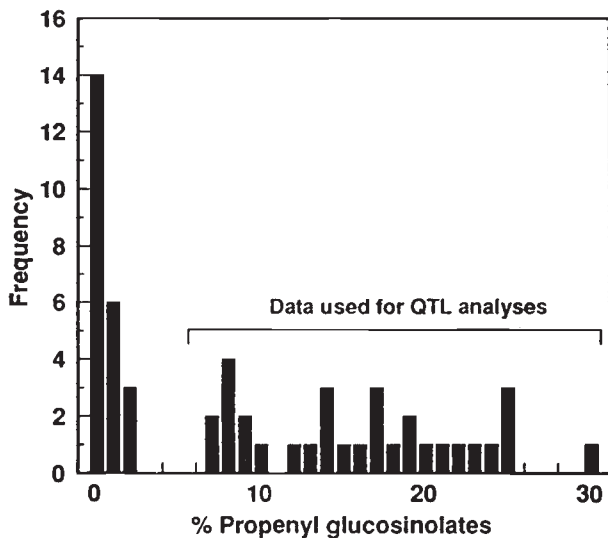


Fig. 4 Segregation of propenyl glucosinolate into two major classes in *Brassica napus* microspore-derived lines.

The levels of glucosinolates with pentyl side chains (i.e. pentenyl + hydroxypentenyl + methylsulphinylpentenyl glucosinolate) in the recombinant lines were also in two discrete classes, one class of 39 lines had between 1 and 12 per cent pentyl glucosinolate and a second class of 15 lines had between 30 and 50 per cent pentyl glucosinolate. These ranges correspond to the two parental phenotypes. The ratio of the two classes was significantly different from a 1:1 ratio ($P < 0.05$). However, as previous studies had demonstrated that a single locus, designated *Gsl-elong-A*, determined the amount of pentyl glucosinolate (Magrath *et al.*, 1993), the divergence from a 1:1 ratio is probably as a result of selection for a trait linked to the *Gsl-elong-A* locus during the development of the microspore-derived lines. Biased segregation of Mendelian traits in microspore-derived lines of *B. napus* has been previously reported (Foisset *et al.*, 1993).

The level of butyl glucosinolates (i.e. butenyl + hydroxybutenyl glucosinolates) within the recombinant lines exhibited continuous segregation. However, if one considers the subset of genotypes which have the Cobra alleles at the *Gsl-elong-A* locus (i.e. alleles which result in the production of pentyl glucosinolates) and functional alleles at the *Gsl-pro* loci, there is discontinuous segregation into two classes, five plants with less than 45 per cent butyl glucosinolate and five plants with greater than 52 per cent butyl glucosinolate, which is assumed to be due to segregation at the *Gsl-elong-C* locus. Segregation patterns for the proportion of butyl glucosinolates in the other combinations of genotypes at the *Gsl-pro* and *Gsl-elong-A* loci were difficult to interpret due to the small number of genotypes within each class.

Within each line, the ratios of the three different side chain lengths is constant when clones of the same genotype were grown under different conditions. Glucosinolate profiles of diploid tissues were identical to that of haploid tissue. Glucosinolate profiles of seeds were similar to that of leaf tissue, apart from a reduction in the proportion of pentyl glucosinolates as described by Magrath *et al.* (1993).

Linkage analyses

The 58 probes used in this study identified 110 RFLP loci. The *Gsl-elong-C* and *Gsl-elong-A* loci mapped onto two linkage groups, G3 and G13 (*sensu* D. J. Lydiate, pers. comm.), which possess pairs of collinear homoeologous RFLP loci (Fig. 5). For example, probe *pN22* identifies two loci, *pN22a* and *pN22b*, on G3 and G13, respectively; *pN22a* (and hence G3) being derived from the A genome parent and *pN22b* from the C genome parent (Fig. 6). Therefore, the two linkage

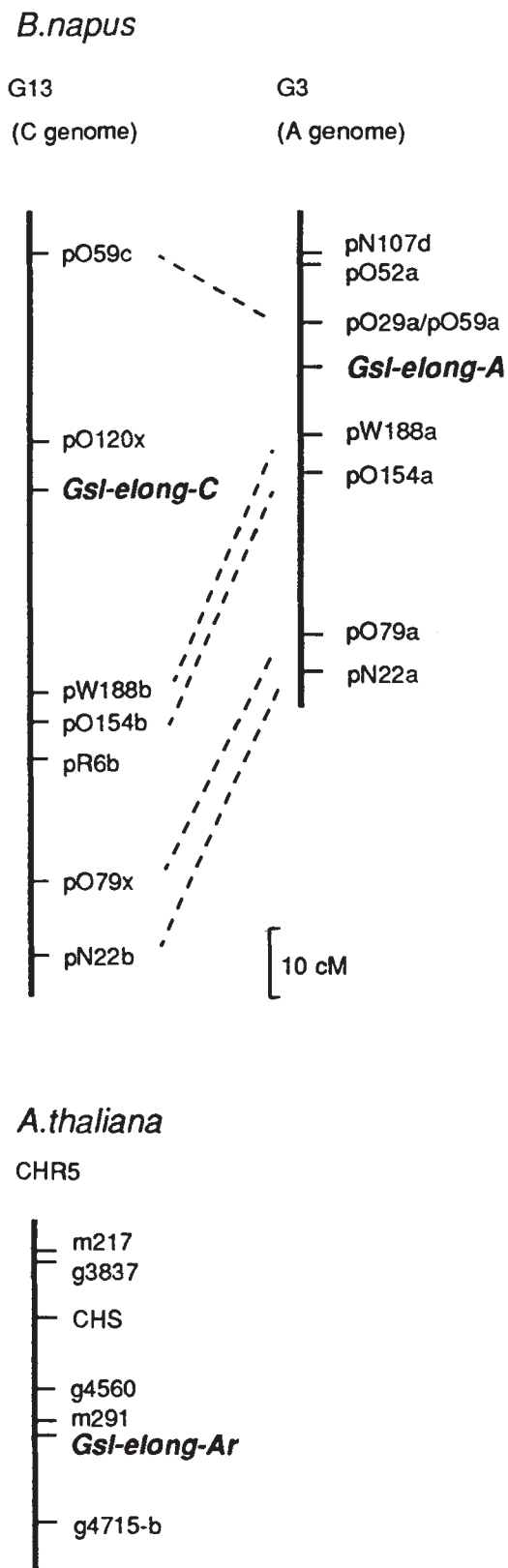


Fig. 5 RFLP linkage maps showing the relative positions of the *Gsl-elong-C*, *Gsl-elong-A* and *Gsl-elong-Ar* loci.

groups represent a pair of homoeologous linkage groups, G3 derived from the A genome and G13 from the C genome. *Gsl-elong-C* maps 8 cM from *pO120x* on G13 and *Gsl-elong-A* maps 7.5 cM from *pO59a* on G3. The *Gsl-pro* locus was 14 cM from the RFLP locus *pN144x*, but neither locus was significantly linked to any other RFLP loci and could not be positioned on a *B. napus* linkage group.

Confirmation of linkage analyses of Gsl-elong-C by QTL analysis and epistatic interactions

Since the data are presented in percentage form, alleles at each of the *Gsl-pro*, *Gsl-elong-C* and *Gsl-elong-A* loci may be expected to have an epistatic effect on the proportions of the three classes of glucosinolates with different side chain lengths (Fig. 2). Two forms of QTL analyses, a traditional single-marker approach and the more recently described interval-mapping method (Stuber *et al.*, 1992), provided evidence for these epistatic interactions, and in so doing confirmed the positions of all three loci including that of *Gsl-elong-C* which is based on a small data set. When the mean level of propenyl glucosinolates in the plants with a functional allele at the *Gsl-pro* locus (Fig. 4) was calculated for each of the two genotypes at every RFLP locus, and the two means compared with a Student's *t*-test, two regions of the genome were shown to have a significant effect on the level of propenyl glucosinolates: a region between RFLP loci *pO59c* and *pW188b* ($P < 0.001$); and a region between loci *pO52a* and *pW188a* ($P < 0.01$), which are coincident with the putative positions of the *Gsl-elong-C* and the *Gsl-elong-A* loci identified through two-point and multipoint analyses (Table 4). Also, analyses of the data with the MAPMAKER/QTL program (Lander *et al.*, 1987; Lander & Botstein, 1989) identified the same regions of linkage groups G3 and G13 as having a quantitative effect on the level of propenyl glucosinolates (Fig. 7). The similar results obtained with the single-marker and interval-mapping methods is consistent with that previously reported (Stuber *et al.*, 1992). Likewise, QTL analyses of the level of butyl and pentyl glucosinolates, by either method, confirmed the position of the *Gsl-pro*, *Gsl-elong-C* and *Gsl-elong-A* loci (data not shown).

Characterization of parental and recombinant lines of A. thaliana

The glucosinolate profiles of *Landsberg erecta* and *Columbia* varied in their side chain length and in their side chain structure; *Landsberg erecta* contained hydroxypropyl glucosinolates while *Columbia* con-

tained methylsulphonylbutyl glucosinolate and trace levels of methylsulphonylpropyl glucosinolate (Fig. 8) and other homologues with longer side chain lengths, as reported previously (Haughn *et al.*, 1991). Within

the F₂ and RI populations, four discrete classes of glucosinolate phenotypes were observed (Fig. 8, Table 5). The ratio of the different phenotypes in the F₂ is not significantly different from a 9:3:3:1 ratio, which is

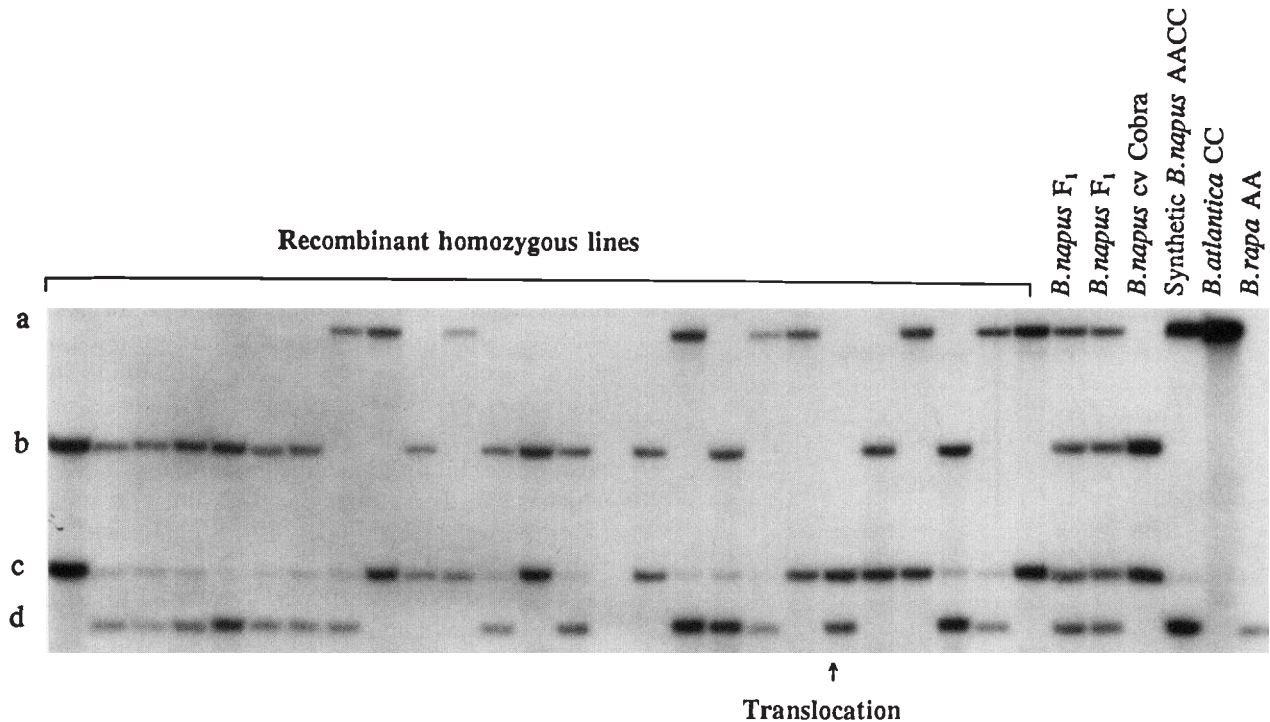


Fig. 6 Autoradiograph of Southern hybridization between 25 recombinant and parental lines and probe *pN22*. Bands a and b are alleles at locus *pN22b* and bands c and d are alleles at locus *pN22a* (Fig. 5). In one of the recombinant lines, there has been a translocation so that both alleles are found at locus *pN22a* and neither at locus *pN22b*.

Table 4 *t*-values obtained from Student's *t*-tests between the mean level of propenyl glucosinolates for groups of plants of each genotype of the RFLP loci on G13 and G3; the data suggest that a locus between RFLP loci *pO59c* and *pW188b* has a major effect on the level of propenyl glucosinolates, which is coincident with the putative position of the *Gsl-elong-C* locus; a second locus between loci *pO52a* and *pW188a* has a smaller, but significant effect, which is coincident with the position of the *Gsl-elong-A* locus

G13			G3		
Locus	Difference in means	<i>t</i> -value	Locus	Difference in means	<i>t</i> -value
<i>pO59c</i>	5.25	1.46	<i>pN107d</i>	3.60	1.57
<i>pO120x</i>	9.05	5.49***	<i>pO52a</i>	4.38	1.93
<i>pW188b</i>	11.30	4.33***	<i>pO29a/pO59a</i>	6.83	3.32**
<i>pO154b</i>	5.28	2.42*	<i>pW188a</i>	8.90	2.13*
<i>pR6b</i>	6.63	2.04	<i>pW154a</i>	5.09	2.35*
<i>pO79x</i>	2.30	0.91	<i>pO79a</i>	4.60	2.03
<i>pN22b</i>	0.61	0.22	<i>pN22a</i>	4.43	1.84

*** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

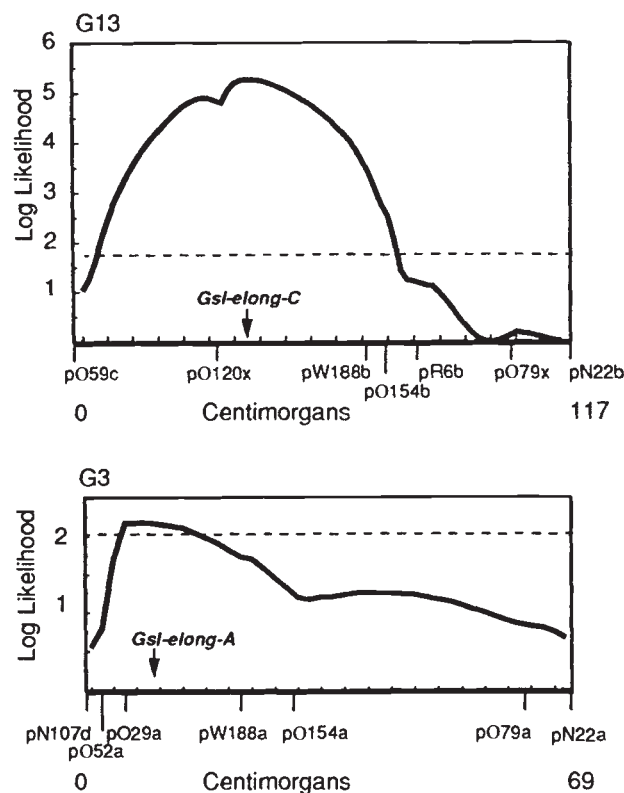


Fig. 7 Log-likelihood probabilities of an association between 2 cM regions of linkage groups G13 and G3 with the level of propenyl glucosinolate in *Brassica napus* lines which have a functional allele at the *Gsl-pro* locus (Fig. 4). A LOD score of two is taken to indicate a significant QTL. Two regions of G13 and G3 are shown to have a significant effect, coincident with the positions of the *Gsl-elong-A* and *Gsl-elong-C* loci identified through two-point and multi-point analyses. The data are consistent with a single-marker approach (Table 3).

consistent with the proposed model (Fig. 2), in which a single locus (designated *Gsl-elong-Ar*) regulates side chain elongation and a single locus (designated *Gsl-ohp-Ar*) regulates the conversion of methylsulphinylpropyl glucosinolate to hydroxypropyl glucosinolate. The ratio in the RI lines is significantly different from the expected 1:1:1:1 ratio with a predominance of phenotypes which possess only propyl side chains, suggesting uneven segregation at the *Gsl-elong-Ar* locus.

Linkage analysis

The *Gsl-elong-Ar* locus maps onto chromosome 5, 1.3 cM from RFLP locus *m291* (Fig. 5). The alleles from Columbia and Landsberg *erecta* at this RFLP locus showed a segregation ratio of 1:1.84, which is significantly different from the expected ratio of 1:1 ($\chi^2 = 8.34$, $P < 0.01$). This suggests that this region of

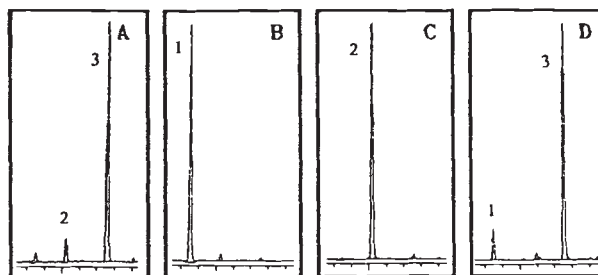


Fig. 8 HPLC profiles of desulphoglucosinolate from (A) leaves of Columbia, (B) Landsberg *erecta* and (C and D) two recombinant lines. 1, Hydroxypropyl (3.4 min); 2, methylsulphinylpropyl (5.1 min); 3, methylsulphinylbutyl (7.6 min). Benzyl desulphoglucosinolate (internal standard), indolyl desulphoglucosinolate and aliphatic desulphoglucosinolates with longer side chains have later retention times; see Haughn *et al.* (1991) for further details.

Table 5 Segregation of F_2 plants and RI lines into four phenotypic classes (see Fig. 8).

	Phenotype (Fig. 8)	F_2	RI
D	Methylsulphinylbutyl Methylsulphinylpropyl (tr) Hydroxypropyl (tr)	84	17
A	Methylsulphinylbutyl Methylsulphinylpropyl (tr)	31	14
B	Hydroxypropyl	17	36
C	Methylsulphinylpropyl	12	23
	Total	144	90

The ratio of the genotypes within the F_2 populations is not significantly different from a 9:3:3:1 ratio ($P > 0.05$), supporting the model proposed in Fig. 2.

Segregation of the RI lines is significantly different from a 1:1:1:1 ($P < 0.01$).

tr = trace levels only.

the chromosome from Landsberg *erecta* has been preferentially selected during the development of the RI lines, which explains the departure in the segregation ratio of glucosinolate phenotypes in the RI lines from the expected ratio. Other genome regions also showed segregation ratios which differed significantly from 1:1 (Lister & Dean, 1993).

Discussion

Allelic variation at the *Gsl-elong* loci

It is likely that the *Gsl-elong* genes produce, or regulate the production of, an elongase that results in the step-

wise addition of the methyl carbon of acetate to homomethionine and 2-amino-6-methylthiohexanoic acid, to produce the precursors of butyl and pentyl glucosinolates (Fig. 9; Underhill *et al.*, 1973). These genes may also regulate the addition of the methyl carbon to methionine, the first step in aliphatic glucosinolate biosynthesis (Fig. 9). The variation in the length of glucosinolate side chains in different species and genotypes is most simply explained by variation in the type of alleles at the *Gsl-elong* loci. This may be due to the different alleles producing elongases with structural variations that result in different affinities for chain-elongated derivatives of methionine, which can be used as substrates for the addition of a further carbon.

In the C genome donor of the synthetic *B. napus*, no chain-elongated glucosinolates are produced. Hence, it is likely that the alleles at the *Gsl-elong-C* locus in this

genotype do not produce a functional elongase, and can be termed null alleles (although they may, as explained above, produce an elongase which catalyses the conversion of methionine to homomethionine). At the *Gsl-elong-C* locus in natural forms of *B. napus*, only butyl glucosinolates are produced. Hence, the alleles at this locus produce a functional elongase which adds a carbon to homomethionine to produce the precursors of butyl glucosinolates, but cannot use these precursors as a further substrate for the addition of further carbons.

The alleles at the *Gsl-elong-A* locus in the A genome donor of the synthetic *B. napus* line are similar, in that predominantly butyl glucosinolates are produced. In contrast, the alleles at the *Gsl-elong-A* locus in natural forms of *B. napus* produce an elongase which can add carbon to both homomethionine and to 2-amino-6-methylthiohexanoic acid, resulting in the production of butyl and pentyl glucosinolates (Fig. 9). *Landsberg erecta* has null alleles at the *Gsl-elong-Ar* locus so that no chain-elongated glucosinolates are produced, whereas the alleles at this locus in Columbia must produce an elongase which predominantly produces butyl glucosinolates. However, if this elongase also has slight affinities for 2-amino-6-methylthiohexanoic acid and homologues of longer length, this would explain the trace levels of long chain aliphatic glucosinolates found within this ecotype (Hogge *et al.*, 1988; Haughn *et al.*, 1991; Fig. 9).

Gene products of the *Gsl-elong* loci

This study has demonstrated that the side chain elongation of aliphatic glucosinolates is under simple genetic control. The biochemical nature of the side chain elongation remains to be elucidated. It has been suggested that elongation is a multienzyme reaction system which involves the transamination of methionine homologues to produce the corresponding α -keto acids, followed by carbon chain elongation by condensation with acetyl-CoA, and then conversion back to the amino acid homologue, analogous to the biosynthesis of leucine from valine (Haughn *et al.*, 1991). Isolating the *Gsl-elong* gene from *A. thaliana* through a map-based cloning strategy and subsequent studies on its structure may provide insights to the nature of its gene product(s) and the biochemistry of side chain elongation. Furthermore, the use of the *Arabidopsis* gene as a probe to isolate *Brassica* homologues from genotypes with different side-chain-elongated glucosinolates may help to elucidate the basis of specificity of products of different alleles at the *Gsl-elong* locus for methionine homologues of different side chain lengths.

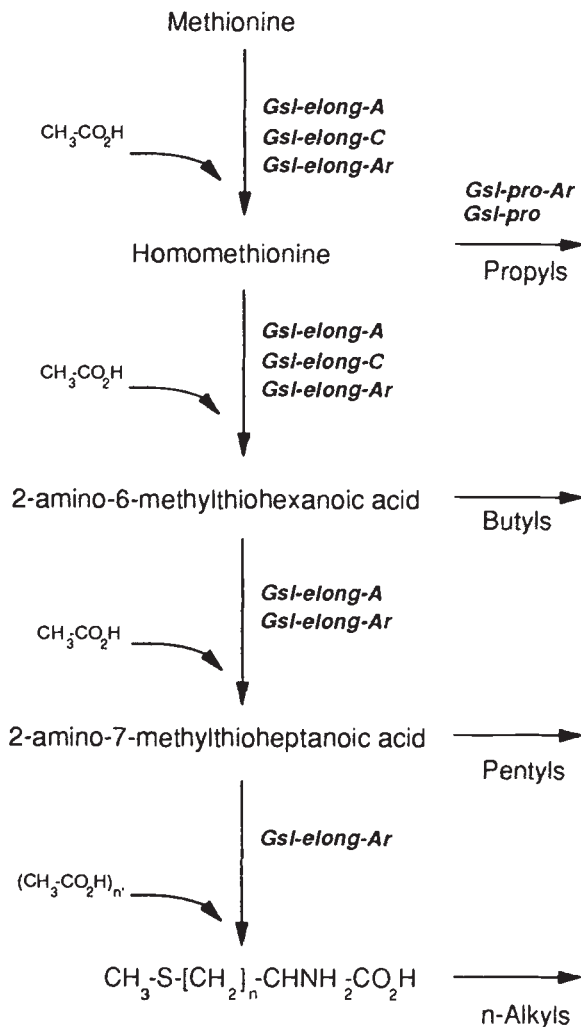


Fig. 9 Hypothetical pathway of side chain elongation showing the possible steps regulated by *Gsl-elong* genes; see Underhill *et al.*, (1973) for further details.

Manipulating aliphatic glucosinolates in *B. napus*

Isolation of an *Gsl-elong-Ar* gene may also provide a novel method to manipulate the aliphatic glucosinolate profile of *B. napus*. Transformation of a *B. napus* line with the *Gsl-elong-Ar* gene (or a *Brassica* homologue) in an antisense configuration may result in the down-regulation of the endogenous *Gsl-elong* genes, which would reduce or eliminate side-chain-elongated derivatives of methionine. This would result in the reduction in the total level of aliphatic glucosinolates. As leaf aliphatic glucosinolates have an important role in deterring non-specific herbivores (Mithen, 1993), it is desirable to reduce aliphatic glucosinolates only in seeds. This could be achieved by driving the antisense gene with a pod-specific promoter, the pod being the site of biosynthesis of seed aliphatic glucosinolates (Magrath & Mithen, 1993). However, reducing aliphatic glucosinolates within leaf tissue would be of value in testing hypotheses concerning the role of these metabolites in pest and pathogen interactions (Chew, 1988). Genes which regulate side chain modifications of aliphatic glucosinolates in *A. thaliana* may also be of value in manipulating the aliphatic glucosinolate profile of *B. napus*, and will be discussed in subsequent publications.

Acknowledgements

The authors thank Cambridge Plant Breeders Ltd and Zeneca Ltd for making RFLP probes available and for providing financial support to I. Parkin and A. Sharpe.

References

- CHEW, F. S. 1988. Biological effects of glucosinolates. In: Cutler, H. G. (ed.) *Biologically Active Natural Products - Potential Use in Agriculture*, pp. 155-181. American Chemical Society Symposium, Washington DC.
- CHISHOLM, M. D. AND MATSUO, M. 1972. Biosynthesis of allyl-glucosinolate and 3-methylthiopropylglucosinolates in horseradish, *Armoracia lapathifolia*. *Phytochemistry*, **11**, 203-207.
- FENWICK, G. R., HEANEY, R. K. AND MULLIN, W. J. 1983. Glucosinolates and their breakdown products in food and food plants. *Crit. Rev. Food Sci. Nutr.*, **18**, 123-301.
- FOISSET, N., DELOURME, R., LUCAS, O. M. AND RENARD, M. 1993. Segregation analysis of isozyme markers on isolated microspore-derived embryos in *Brassica napus* L.. *Plant Breed.*, **110**, 315-322.
- HAUGHN, G. W., DAVIN, L., GIBLIN, M. AND UNDERHILL, E. W. 1991. Biochemical genetics of plant secondary metabolites in *Arabidopsis thaliana*. The glucosinolates. *Plant Physiol.*, **97**, 217-226.
- HOGGE, L. R., REED, D. W., UNDERHILL, E. W. AND HAUGHN, G. W. 1988. HPLC separation of glucosinolates from leaves and seeds of *Arabidopsis thaliana* and their identification using thermospray liquid chromatography-mass spectrometry. *J. Chrom. Sci.*, **26**, 551-556.
- KOSAMBI, D. D. 1944. The estimation of map distance from recombinant values. *Ann. Eugen.*, **12**, 172-175.
- LANDER, E. S. AND BOTSTEIN, D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, **121**, 185-199.
- LANDER, E. S., GREEN, P., ABRAHAMSON, A., BARLOW, A., DALY, M. J., LINCOLN, S. E. AND NEWBURG, I. 1987. Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, **1**, 174-181.
- LISTER, C. AND DEAN, C. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.*, **4**, 745-750.
- MAGRATH, R. AND MITHEN, R. 1993. Maternal effects on the expression of individual aliphatic glucosinolates in seeds and seedlings of *Brassica napus*. *Plant Breed.*, **111**, 249-252.
- MAGRATH, R., HERRON, C., GIAMOUSTARIS, A. AND MITHEN, R. 1993. The inheritance of aliphatic glucosinolates in *Brassica napus*. *Plant Breed.*, **111**, 55-72.
- MITHEN, R. 1993. Leaf glucosinolate profiles and their relationship to pest and disease resistance in oilseed rape. *Euphytica*, **63**, 71-83.
- POULTON, J. E. AND MOLLER, B. L. 1993. Glucosinolates. In: Lea, P. J. (ed.) *Methods in Biochemistry, vol. 9: Enzymes of Secondary Metabolism*, pp. 209-237. Academic Press, London.
- STUBER, C. W., LINCOLN, S. E., WOLFF, D. W., HELENTJARIS, T. AND LANDER, E. S. 1992. Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers. *Genetics*, **132**, 823-829.
- UNDERHILL, E. W., WETTER, L. R. AND CHISHOLM, M. D. 1973. Biosynthesis of glucosinolates. *Biochem. Soc. Symp.*, **38**, 303-326.