Theor Appl Genet (1998) 96 : 177–186 

P. Barret · R. Delourme · M. Renard · F. Domergue
R. Lessire · M. Delseny · T. J. Roscoe

# A rapeseed *FAE1* gene is linked to the E1 locus associated with variation in the content of erucic acid

Received: 4 April 1997 / Accepted: 30 July 1997

**Abstract** The synthesis of very long chain fatty acids occurs in the cytoplasm via an elongase complex. A key component of this complex is the $\beta$-ketoacyl-CoA synthase, a condensing enzyme which in Arabidopsis is encoded by the *FAE1* gene. Two sequences homologous to the *FAE1* gene were isolated from a *Brassica napus* immature embryo cDNA library. The two clones, CE7 and CE8, contain inserts of 1647 bp and 1654 bp, respectively. The CE7 gene encodes a protein of 506 amino acids and the CE8 clone, a protein of 505 amino acids, each having an approximate molecular mass of 56 kDa. The sequences of the two cDNA clones are highly homologous yet distinct, sharing 97% nucleotide identity and 98% identity at the amino acid level. Southern hybridisation showed the rapeseed $\beta$-ketoacyl-CoA synthase to be encoded by a small multigene family. Northern hybridisation showed the expression of the rapeseed *FAE1* gene(s) to be restricted to the immature embryo. One of the *FAE1* genes is tightly linked to the E1 locus, one of two loci controlling erucic acid content in rapeseed. The identity of the second locus, E2, is discussed.

**Key words** $\beta$-ketoacyl-CoA synthase · *FAE1* · *Brassica napus* · Erucic acid · E1 locus

**Abbreviations** *FAE1 Fatty Acid Elongation 1* · KCS $\beta$-ketoacyl-Coenzyme A synthase · *C22:1* erucic acid · *VLCFA* very long chain fatty acid · *LEAR* low erucic acid rapeseed · *HEAR* high erucic acid rapeseed · *DAF/P* days after flowering/pollination · *TAG* triacylglycerol · *RAPD* random amplification of polymorphic DNA · *RFLP* restriction fragment length polymorphism

## Introduction

Very Long Chain Fatty Acids (VLCFA), contain at least 20 carbon atoms and occur ubiquitously in nature. In plants, VLCFAs are found in epidermal barrier waxes and in seed storage oils. Many Brassica species, including mustard (*B. juncea*, *B. carinata* and *B. nigra*) and rape (*B. napus*), contain the very long chain monounsaturated fatty acids eicosanoic (C20 : 1) and erucic (C22 : 1) as components of their seed oil. Certain varieties of oilseed rape (High Erucic Acid Rapeseed, HEAR) produce a triacylglycerol (TAG) that contains a high proportion of erucic acid (45–60% of the oil) which is used as an industrial feedstock for the production of carburants, lubricants and in a variety of other products, (Murphy and Sonntag 1991). An increase in the demand for rapeseed rich in erucic acid for industrial purposes has led to an interest in developing very high erucic acid rapeseed lines by genetic manipulation. Such an approach is required to overcome the theoretical limit of 66% erucic acid content of TAG imposed by the substrate specificity of the acyltransferase

*Note*: The nucleotide sequence corresponding to one of the cDNAs (CE8) reported in this paper has been entered in the Genbank data base under the accession number U50771

Communicated by J. S. Beckman

P. Barret · R. Delourme · M. Renard
INRA Station d'Amélioration des Plantes, B.P. 29,
35650 Le Rheu, France

F. Domergue · R. Lessire
Laboratoire de Biogenèse Membranaire, C.N.R.S. URA 1811,
Université Bordeaux II, 146 rue Léo Saignat, 33070 Bordeaux,
France

M. Delseny · T. J. Roscoe (✉)
Laboratoire Physiologie et Biologie Moléculaire des Plantes,
C.N.R.S. UMR 5545, Université de Perpignan,
52 Avenue de Villeneuve, 66860 Perpignan cedex, France
Tel.: (33) 04 68 66 21 31
E-mail: roscoe@univ-perp.fr

responsible for acylation at position *sn-2* of the glycerol moiety of the TAG, which discriminates against VLCFAs (Bernerth and Frentzen 1990; Cao et al. 1990).

VLCFAs are synthesised by the microsomal fatty acid elongation complex (elongase) using acyl-CoA substrates from a cytoplasmic pool maintained by de novo lipid biosynthesis in the plastid. Fatty acid elongation is achieved by the sequential condensation of two carbon units derived from a malonyl-CoA donor to a long chain acyl-CoA primer. The elongase complex introduces a $\beta$-ketoacyl group into the growing acyl chain via a $\beta$-ketoacyl-CoA synthase, and the keto group is removed by a series of three reactions: $\beta$-ketoreduction to the $\beta$-hydroxylacyl-CoA, $\beta$-hydroxydehydration to the enoyl-CoA and an enoyl reduction of the double bond, (reviewed by Cassagne et al. 1994; von Wettstein, 1994). Thus, elongase requires a minimum of four activities (excluding the transacylase providing the donor malonyl-CoA) for each C2 addition. These reactions are analogous to the plastidal FAS and KASIII reactions which utilise Acyl Carrier Protein (ACP) as the acyl carrier.

The structure and functioning of the acyl-CoA elongase complex is poorly understood because of the difficulty in purifying functional membrane proteins to homogeneity. Recent progress in this area has resulted from studies on the seed elongases of jojoba and rape. The acyl-CoA elongase complex has been partially purified from developing rape embryos and has resulted in the enrichment of four proteins between 54 and 67 kDa in size (Creach and Lessire 1996). The $\beta$-ketoacyl-CoA synthase (KCS) was purified from jojoba embryos by Lassner et al. (1996). The corresponding cDNA, homologous to the Arabidopsis *FATTY ACID ELONGATION1* (*FAE1*) gene (James et al. 1995), was used to transform rapeseed plants. Subsequent KCS activity in developing embryos of Low Erucic Acid Rapeseed (LEAR) plants resulted in an enrichment (up to 33.5% by weight) of the seed oil with VLCFAs, thereby demonstrating that KCS activity had been confered. The restoration of the elongation activity in LEAR plants transformed with the jojoba KCS gene suggests that the mutations that gave rise to the LEAR phenotype are associated with either the structural gene encoding KCS or with genes regulating the expression of KCS.

In rapeseed, the two elongation steps from oleoyl-CoA to erucic acid are each controlled by alleles at two loci, E1 and E2, which exhibit additive gene action (Harvey and Downey 1963; Stefansson 1983). Using random amplified polymorphic DNA (RAPD) markers, Jourdren et al. (1996a) mapped the two loci and determined erucic content in a doubled-haploid population. Assignment of the two loci to independant linkage groups via a quantitative trait locus (QTL) approach confirmed the findings of Thormann et al. (1996). However, in contrast, Jourdren et al. (1996a) correlated individual segregating genotypes with erucic content to demonstrate that the two loci do not contribute equally to erucic acid content.

In this paper we describe the isolation of two rapeseed cDNAs corresponding to a $\beta$-ketoacyl-CoA synthase using an heterologous oligonucleotide probe made to the Arabidopsis *FAE1* gene. We show that at least two genes encode the rapeseed enzyme and that these are temporally expressed in immature embryos. Furthermore, we show that one of the structural genes is tightly linked to the E1 locus, which is associated with the greatest variation in erucic acid content in rapeseed.

## Materials and methods

### Plant material

The progeny of a cross between varieties 'B002' ($< 1\%$ erucic acid) and 'Hokkaido' ($> 50\%$ erucic acid) were used for the construction of cDNA libraries and the isolation of the *FAE1* cDNAs. High erucic acid segregant lines from a backcrossed B5$F_2$ population (recurrent parent 'B002') were obtained by allowing flowers of a B5 $F_1$ plant to self-pollinate by covering with paper bags. Seeds of these plants were vernalised for 2 months at 4°C and long daylength and transferred to a field nursery in spring 1994. For the mapping studies, the doubled-haploid (DH) population was used that was described in Foisset et al. (1996). The DH population was derived from the cross between varieties 'Darmor-*bzh*' ($< 1\%$ erucic acid) and 'Yudal' ($> 50\%$ erucic acid) and was used for the construction of a genetic map of rapeseed (Foisset et al. 1996). This doubled-haploid population was previously evaluated for erucic acid content by Jourdren et al. (1996a).

### Extraction of RNA

Total RNA was extracted from embryos dissected from the seeds of 20 high erucic segregant B5 $F_2$ plants isolated at 7-day intervals from day 15 to day 49 after pollination, together with various other tissues by a protocol based on the method of Kay et al. (1987).

### Construction of an immature embryo cDNA library of *Brassica napus*

Poly(A)$^+$RNA was isolated from total RNA extracted from immature embryos of 20 independant HEAR segregant lines of the B5 $F_3$ population 28 days after pollination, according to a protocol described in Sambrook et al. (1989). Approximately 5 µg Poly(A)$^+$ RNA was used to construct a cDNA library in the vector lambda ZAPII cDNA according to the manufacturer's protocols (Stratagene). The primary library was titred at $4.46 \times 10^6$ pfu ml$^{-1}$ with a non-recombinant background of 0.65%, and an aliquot was amplified to $1.3 \times 10^9$ pfu ml$^{-1}$.

### Isolation of cDNA clones encoding rapeseed $\beta$-ketoacyl-CoA synthase

A 34-mer oligonucleotide probe was designed from the sequence of the Arabidopsis *FAE1* gene (James et al. 1995). It was located within

a conserved region common to condensing enzymes and corresponded to the following sequence:

5′ GAT CAA CGT TAC ATA GAT TTG GGA ATA CTT CAT C 3′.

The oligonucleotide was labelled with fluorescein (Amersham) and used to screen 90 000 recombinant phage of the immature embryo cDNA library. Hybridising plaques were detected with fluorescein-HRP antisera and visualised by ECL (Amersham). Recombinant pBluescript plasmids were isolated from purified phage clones by in vivo excision as described in the lambda ZAPII cloning protocol (Stratagene).

### DNA sequencing

DNA sequences were determined using DyeDeoxy Terminator cycle sequencing (Applied Biosystems) on double-stranded DNA templates with an ABI 373A sequencer. A series of overlapping size deletions was obtained by exonuclease III digestions (Erase-a-base kit, Promega) performed on the largest clone isolated. The DNA sequences were analysed using the SEQUENCHER, DNA STRIDER, FASTA and BLAST (Altshul et al. 1990) programmes.

### Genomic Southern analysis

Genomic DNA was extracted from young green leaf tissue according to the protocol of Doyle and Doyle (1990). A series of reactions containing 10 μg DNA were digested with the appropriate restriction enzyme, and the resulting fragments were separated on a 0.8% agarose gel, transfered to Hybond $N^+$ nylon membrane (Amersham) via an alkali blotting protocol. A 1.026-kbp probe contained within the coding sequence was generated by polymerase chain reaction (PCR) using the primers:

CE8IN: 5′ ACA CGC CAT GGG TGC GCT AGA AAA TCT ATT CAA G 3′
CE8B: 5′ CGG GAT CCT TAG GAC CGA CCG TTT GG 3′.

The amplimer was radiolabelled by random priming. Following hybridisation in 5 × SSC, 5 × Denhardt's, 0.5% SDS and 20 μg $ml^{-1}$ calf thymus DNA, the membrane was washed initially at 5 × SSC, 0.1% SDS at 65°C for 15 min, autoradiographed and then washed subsequently at 0.125 × SSC, 0.1% SDS, and re-exposed.

### Northern hybridisation analysis

Total RNA (20 μg) or poly $(A)^+$ RNA (1 μg) was fractionated on 1.2% agarose gels containing formaldehyde and transferred to Hybond N (Amersham) nylon membrane. Hybridisation was at 65°C in 5 × SSC, 5 × Denhardt's solution, 0.5% SDS and 20 μg $ml^{-1}$ calf thymus DNA. The filter was washed at 65°C in 0.125 × SSC, 0.1% SDS and then autoradiographed with intensifying screens at − 80°C. The filters were subsequently stripped and reprobed with a 18S RNA probe in order to compare the relative amounts of RNA in each lane.

### Amplification and radioactive labelling of the *FAE1* probe for mapping

The DNA insert contained within the cDNA clone CE8 was amplified by PCR using the following primers: pD-5′ GTT TTC CCA GTC ACG A 3′ and pR-5′ GGA AAC AGC TAT GAC C 3′. The PCR reaction mixture (50 μl) contained 1 × *Taq* polymerase buffer (Eurobio, Les Ullis, France), 3 m*M* $MgCl_2$, 75 μ*M* each of dATP, dCTP, dGTP, dTTP (Boehringer, Meylan, France), 20 pmole of each primer, 0.5 unit of *Taq* polymerase (Eurobio) and 4 μl of an overnight culture of the bacterial clone as template. PCR was performed in a Perkin Elmer 480 or 9600 apparatus under the conditions described by Szewc-McFadden et al. (1996) for amplification of rapeseed microsatellites. After amplification, the DNA was desalted by microcentrifugation on a Sephadex G50 spin column (Pharmacia, Uppsala, Sweden). DNA amount was estimated on agarose minigels, 50–100 ng of DNA was used for the radioactive labelling reaction (Quickprime labelling kit, Pharmacia).

### Restriction fragment length polymorphism (RFLP) analysis

RFLP analysis was performed as described in Sharpe et al. (1995). DNA was extracted from young leaves according to the protocol of Doyle and Doyle (1990). Labelled Southern blots were exposed to BioMax autoradiography films (Kodak, Rochester, USA).

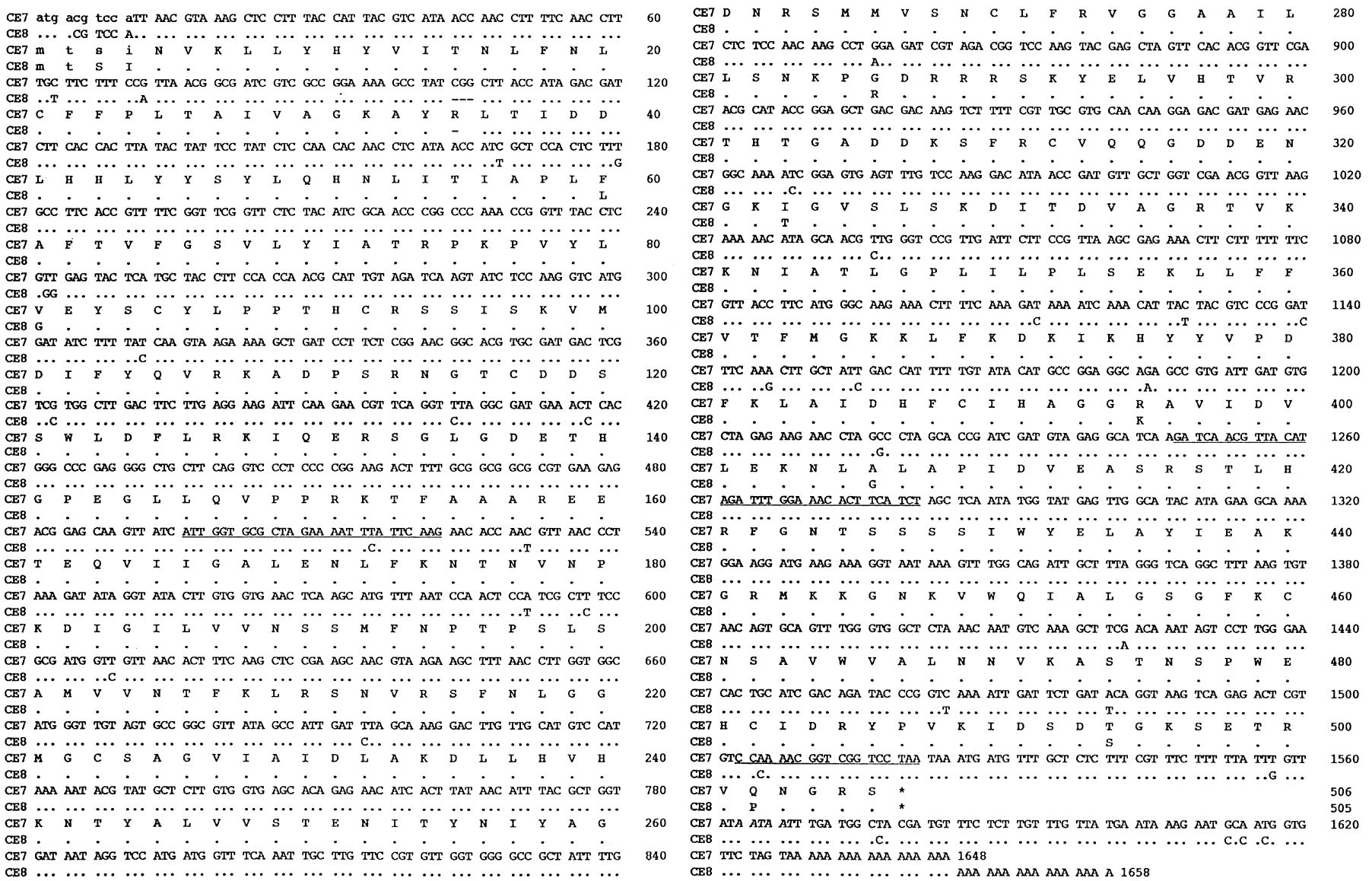
### Genetic analysis

The CE8/*Hin*dIII marker was scored for each genotype on the DH population of 126 lines used by Foisset et al. (1996) and Jourdren et al. (1996a). Linkage analysis was performed using MAPMAKER software (Lander et al. 1987) and MAPMAKER QTL software as described in Jourdren et al. (1996a).

## Results

### Isolation of cDNAs encoding a β-ketoacyl-CoA synthase

An oligonucleotide probe corresponding to a sequence of the *FAE1* gene (James et al. 1995) of Arabidopsis and contained within a consensus sequence common to the condensing enzymes β-ketoacyl-CoA synthase, chalcone synthase and stilbene synthase was used to screen an immature embryo cDNA library of *Brassica napus*. Twelve hybridising clones were isolated after three rounds of screening. The 2 largest clones, CE7 and CE8, were sequenced to completion and found to contain 1647- and 1654-bp cDNA inserts, respectively. These sequences, when aligned with the nucleotide database, were found to be homologous to the Arabidopsis *FAE1* gene (James et al. 1995). Alignment with the *FAE1* gene indicated that the rapeseed CE8 sequence was truncated 4 nucleotides from the 5′ end of the corresponding Arabidopsis open reading frame (ORF) and thus lacked the presumed initiation codon. That the identity of the missing nucleotides was ATGA was confirmed after isolation of this region of the rapeseed genomic DNA via a PCR walking protocol (Devic et al. 1997). This ATG is likely to be the initiating codon since a stop codon is present in the same reading frame 3 nucleotides upstream. A second clone, CE7, was subsequently isolated that was highly homologous to clone CE8; CE7 was presumed to be truncated 10 nucleotides from the 5′ end of the ORF. The complete nucleotide sequence and deduced amino acid sequence of clone CE7 is presented in Fig. 1, as an



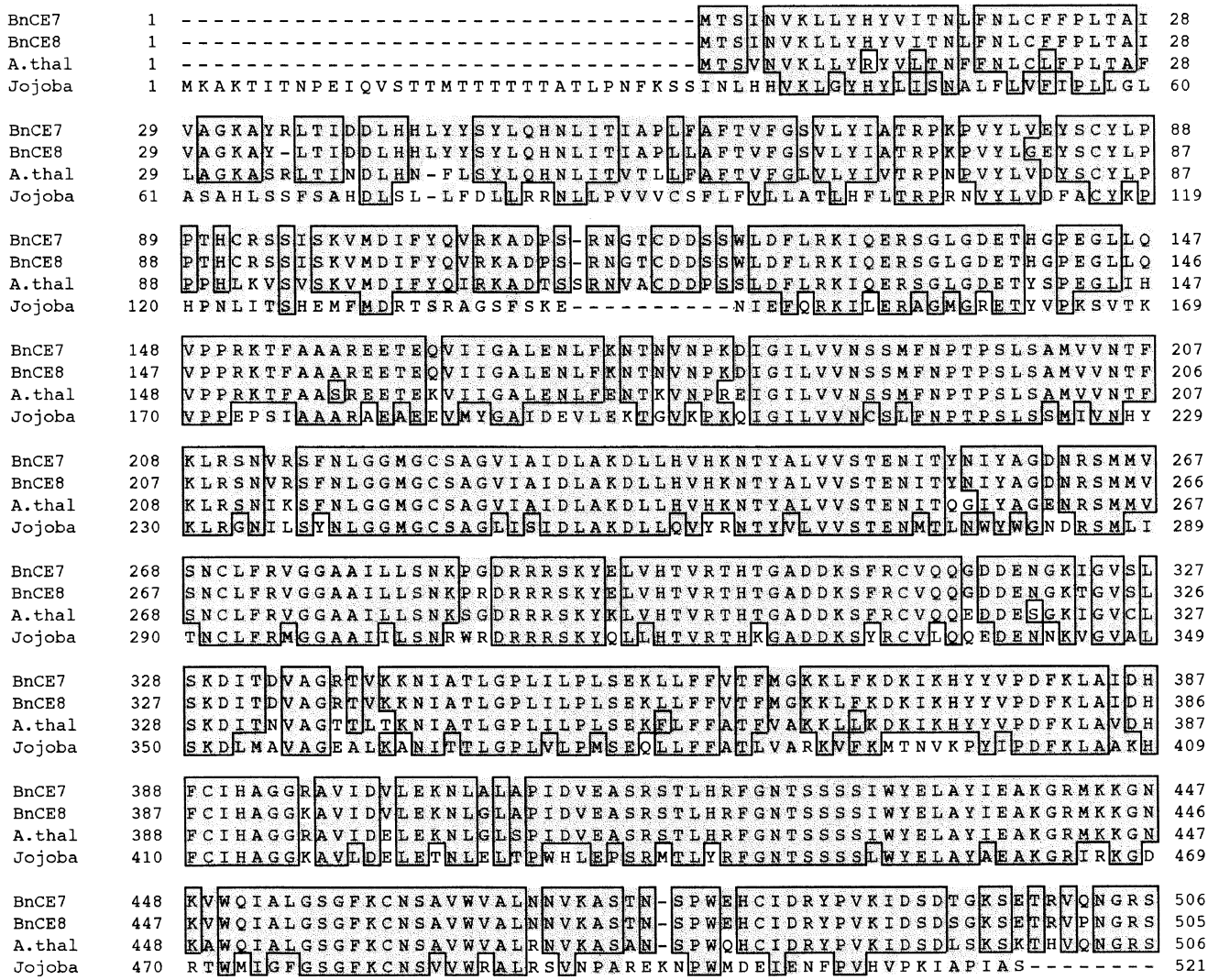



**Fig. 2** Comparison of the deduced amino acid sequences of rapeseed β-ketoacyl-CoA synthases with Arabidopsis and jojoba. *Outlined boxes* indicate identity; *shading* indicates conservative differences while the *absence of shading* indicates nonconservative differences

alignment with clone CE8 (Genbank accession U50771) to illustrate the differences between the cDNAs.

The principle difference between clones CE7 and CE8 is the absence of an arginine codon after nucleotide 102 in the CE8 sequence. Thus, CE8 encodes a protein of 505 amino acids and CE7 a protein of 506 amino acids of molecular masses 56 191 and 56 423 Da, respectively, and a common pI of 9.26. The 2 rapeseed sequences share 97.6% identity at the nucleotide level. This variation results in differences between the cDNAs in the deduced amino acid sequence at nine residues, five of which are nonconservative substitutions; thus, the two proteins share 98.2% identity over a 501 amino acid overlap. A low level of variation in the nucleotide sequence between the CE7 and CE8 sequences and the additional cDNA clones isolated was evident; for example, the CE3 cDNA differs from CE7 at 8 nucleotide positions over a sequence of 1155 bp, (data not shown). Several other hybridising clones were partially sequenced at their extremities. Each of the CE7 and CE8 cDNAs contain the sequence of the oligonucleotide used for screening (Fig. 1), differing in 2 nucleotides from the Arabidopsis sequence.

The deduced amino acid sequences for each of the rapeseed cDNAs were compared with those of Arabidopsis and jojoba and the identity of the rapeseed proteins confirmed as a *β*-ketoacyl-CoA synthase (Fig. 2). In general, a high degree of identity is evident

**Fig. 1** The nucleotide and deduced amino acid (*single letter code*) sequences of rapeseed *β*-ketoacyl-CoA synthase cDNA clones CE8 (Genbank accession U50771) and CE7. The complete sequence for CE7 is shown. For clone CE8 the sequence is shown only at positions of difference from clone CE7. Identical nucleotides or amino acids are indicated by *dots*. Gaps introduced for alignment are indicated by *dashes*. An *asterisk* indicates a stop codon. A *lower case letter* indicates nucleotide and deduced amino acid sequences obtained from the genomic sequence. The sequences of oligonucleotides used for screening and PCR are *underlined*

throughout the protein sequence with the least conservation at the amino and carboxy terminal extremities. The jojoba sequence is the most divergent, possessing a threonine-rich extension of 32 amino acids at the N-terminus and lacking two blocks of residues towards the N-terminus after E143 and at the C-terminal extremity. At the amino acid level, the rapeseed sequences share 85% identity with the Arabidopsis sequence and 54% identity with the jojoba sequence.

## The β-ketoacyl-CoA synthase of rapeseed is encoded by a minimum of two genes

Variation in the nucleotide sequence between the 2 cDNA clones CE7 and CE8, the low level but consistantly observed sequence variation among the other partially sequenced cDNA clones together with the variation in the 5′ nontranslated sequence found using PCR walking on genomic DNA (Devic et al. 1997) suggested the existence of several genes encoding β-keto-acyl-CoA synthase. In order to investigate this, we performed Southern hybridisation (Fig. 3). Genomic DNA isolated from plants of the HEAR segregating lines homozygous for the high erucic acid trait was digested with enzymes *Dra*I, EcoRI, *Eco*RV and *Xba*I. A 1.026-kbp CE8 truncated probe was hybridised and washed at low stringency. The probe did not discriminate between CE7 and CE8. There were no cleavage sites within the probe sequence for the enzymes used for digestion. Sequencing of the genomic clones obtained by PCR walking revealed an absence of intron sequences, (data not shown). Autoradiography revealed the presence of two strongly hybridising fragments from each restriction digest, except *Xba*I for which a single strong signal was evident. In addition, a minimum of six weakly hybridising fragments were detected with the DNA digested with *Dra*I and *Eco*RI. After washing under high stringency conditions, the number of hybridising fragments was unchanged. Similarly, when hybridised with a 117-bp probe made to the 3′ nontranslated sequence common to each cDNA, the pattern of hybridising fragments was similar to that of the coding sequence probe (data not shown). Thus, rapeseed β-ketoacyl-CoA synthase is encoded by a minimum of two highly homologous genes and may be present in a small multigenic family consisting of six to eight members.

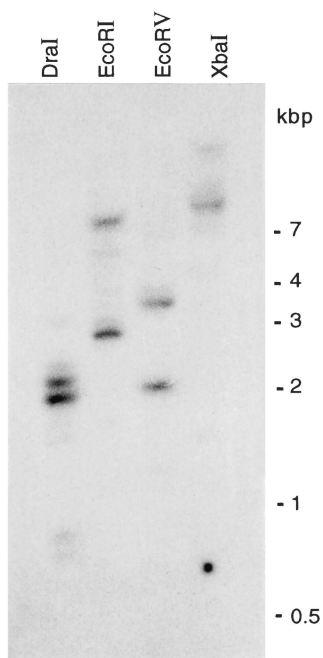



**Fig. 3** Southern hybridisation analysis of rapeseed *FAE1* genes. Pooled, genomic DNA isolated from 20 HEAR segregant plants was digested with the enzymes *Dra*I, *Eco*RI, *Eco*RV and *Xba*I, resolved on an agarose gel and transferred to a nylon membrane. A truncated coding sequence probe (1.026 kbp) was hybridised, the filter washed at a stringency of 0.125 × SSC, 0.1% SDS at 60°C and autoradiography was performed

## Expression of the β-ketoacyl-CoA synthase gene(s) during seed development

In order to determine the pattern of rapeseed β-ketoacyl-CoA synthase expression, we performed Northern hybridisation with total RNA isolated from embryos dissected at weekly intervals through seed development to maturation and from flower, root and leaf tissues (Fig. 4). Expression of the *FAE1* gene(s) was found to be restricted to the embryo and temporally regulated during seed development. A probe corresponding to the entire CE8 sequence was found to hybridise to a mRNA of embryos at 28–42 days after pollination (DAP). This transcript was detectable after 21 DAP and was substantially reduced at 49 DAP. A poly $(A)^+$ RNA preparation isolated from immature embryos confirmed the presence of the *FAE1* transcript(s) at 26 DAP.

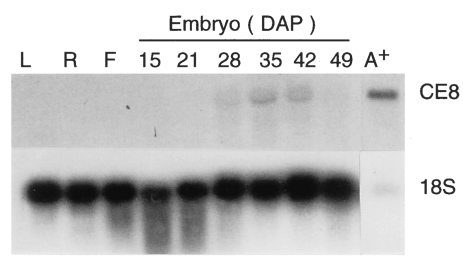



**Fig. 4** Northern hybridisation analysis of rapeseed *FAE1* expression. Total RNA was extracted from leaf (*L*), root (*R*), flower (*F*) and immature embryos dissected during seed development. The *numbers above* the lanes indicate days after pollination (*DAP*). Poly(A) + RNA was isolated at 26 DAP. The RNAs were resolved on formamide agarose gels, blotted to a nylon membrane and hybridised with a 1.5-kbp probe corresponding to the CE8 cDNA insert and subsequently with a ribosomal RNA probe. The filters were washed at a stringency of 0.125 × SSC, 0.1%SDS at 60°C, and autoradiography was performed

The *FAE1* gene is linked to the E1 locus controlling erucic acid content in rapeseed

The RAPD markers linked to the E1 and E2 loci identified in a previous study (Jourdren et al. 1996a) were subsequently assigned to two linkage groups (DY7 and DY9) on the rapeseed genetic map (Foisset et al. 1996). Hybridisation of the CE8 probe to Southern blots of DNA of the parental varieties 'Darmor-*bzh*' and 'Yudal' digested with various restriction endonucleases revealed at least two loci, (data not shown). In the case of digestion by *Hin*dIII, five hybridising fragments were detected (Fig. 5), since CE8 contains three internal restriction sites for this enzyme. Linkage analysis showed that one of these loci (CE8/*Hin*dIII) mapped to the DY 9 linkage group and cosegregated with the E1 locus. From a test population of 100 plants, no recombinants were identified. The position of the E1 locus was determined in Jourdren et al. (1996a). The highest LOD score value for the E1 QTL was found at CE8/*Hin*dIII locus (Fig. 6). The same locus was mapped with a CE8 *Pst*I RFLP. To date, the second locus has not exhibited polymorphism despite the use of eight different restriction enzymes with the 'Darmor-*bhz*' and 'Yudal' genomic DNAs.

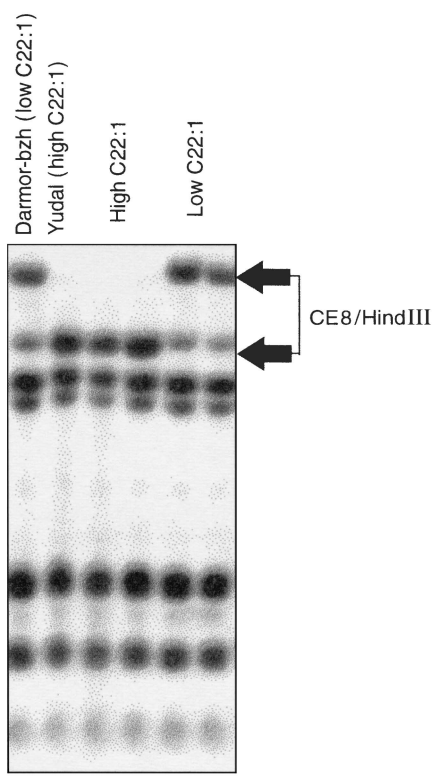



**Fig. 5** RFLP analysis of *FAE1* gene. Genomic DNA isolated from 'Yudal' (HEAR), 'Darmor-*bzh*' (LEAR) and doubled-haploid progeny of this cross (labelled high or low) was digested with *Hin*dIII, transferred to a nylon membrane and hybridised with a FAE1 probe (CE8) and autoradiography was performed

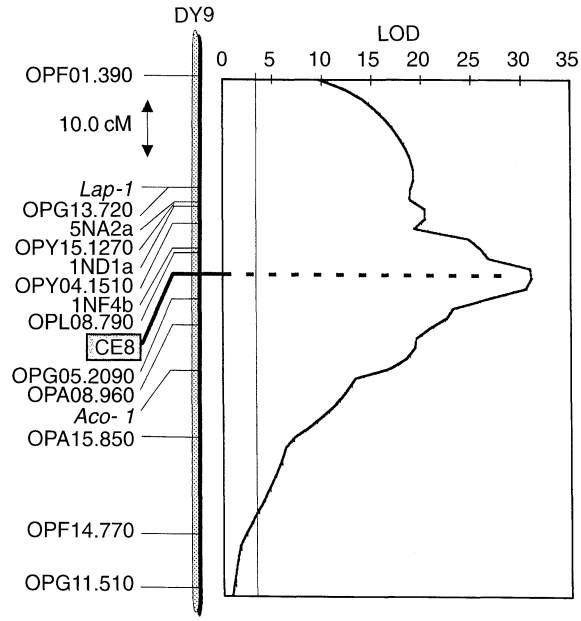



**Fig. 6** Genetic mapping of the CE8 gene. Localisation of the CE8/*Hin*dIII locus on the DY9 linkage group. The highest LOD score for the E1 QTL was found at the CE8/*Hin*dIII locus

## Discussion

As part of our efforts to understand the regulation of erucic acid biosynthesis in *Brassica napus* seeds and in order to gain insight into the molecular basis of the difference between the LEAR and HEAR phenotypes, we report here the characterisation of two distinct cDNAs encoding a $\beta$-ketoacyl-CoA synthase. We show that a gene encoding at least one of these cDNAs cosegregates with a QTL (E1) controlling erucic acid content in rapeseed.

### Heterogeneity among rapeseed $\beta$-keto-acyl-CoA synthase cDNAs

The identity of the rapeseed clones was established by alignment with the Genbank and Swissprot databases. The rapeseed sequences were highly homologous at the nucleotide (Fig. 1) and protein (Fig. 2) levels to the Arabidopsis sequence (James et al. 1995) and to the jojoba sequence (Lassner et al. 1995) which was shown to encode $\beta$-ketoacyl-CoA synthase activity.

Two highly homologous, yet distinct, cDNAs were isolated, discriminated principally by the absence of an arginine codon in clone CE8 together with a low level of nucleotide variation (Fig. 1). Previous reports have demonstrated the existence of highly homologous *Brassica napus* cDNAs encoding other proteins involved in fatty acid biosynthesis; for example, Loader et al. (1993) describe two ACP-thioesterase cDNAs sharing 95.7% nucleotide identity and de Silva et al.

(1990) describe two acyl carrier protein genes sharing 94% nucleotide identity.

Additional nucleotide variation was evident among the other clones isolated for which partial sequence data was obtained. Variation in the nucleotide sequence between clones CE3 and CE7 inferred variation in three amino acid residues, two of which are nonconservative substitutions, $F_{208}$ to S and $Q_{543}$ to P. Other clones sequenced at their extremities, for example CE4 and CE9, resembled clone CE7, and in the case of clone CE12, resembled clone CE8. In addition, clone CE3 possessed a 3′ nontranslated sequence 80 bp longer than the sequence common to that of CE7. The size difference in length of the 3′ nontranslated sequence between the CE3 and other clones may be due to differential processing of the 3′ end of the pre-mRNA or may be a consequence of the cloning. Clones CE7 and CE8 shared near identical 3′ nontranslated sequences that differed in only four nucleotides.

The sequence heterogeneity between CE7 and CE8 may be a consequence of the allotetraploid nature of the *Brassica napus* genome. *Brassica napus* is hypothesised to have arisen from a spontaneous cross between *Brassica campestris* and *B. oleracea* (U 1935). The variation between CE7 and CE8 cDNAs may reflect sequence polymorphism between the parental *FAE1* gene copy present in each donor genome, as was argued by Kater et al. (1991) for sequence variants of enoyl-ACP reductase. Alternatively, CE7 and CE8 may be linked members of a gene family. The two strong hybridisation signals detected on the Southern blot may correspond to the 2 cDNAs specifically expressed in embryo tissue. The minor sequence variation between the CE7 and CE8 cDNAs and the other clones, for example CE3, may be attributable to allelic variation and to residual heterozygosity despite intensive selection. The Southern hybridisation analysis of genomic DNA (Fig. 3) confirmed rapeseed β-ketoacyl-CoA synthase to be encoded by a minimum of two genes and possibly by a small multigene family having up to eight members. The detection of multiple, weakly hybridising fragments may thus be a consequence of the existence of structurally related genes encoding enzymes which function in the synthesis of VLCFA components of the plasma membrane and as precursors of waxes (Post-Beittenmiller 1996).

Some of the minor variation in sequence may be associated with the nature of the cDNA library. This was constructed from RNA isolated from a population of 20 individual segregating lines characterised by high erucic acid content.

## Expression of the rapeseed *FAE1* gene(s) is temporally regulated during embryo development

Expression of the rapeseed *FAE1* gene(s) was restricted to immature embryo tissue and thus confirms the identity of the encoded enzyme as a seed specific β-ketoacyl-CoA synthase. Analysis by Northern hybridisation using a coding sequence probe revealed the presence of transcript(s) approximately 1.7 kb in length in immature embryos between 21 and 42 DAP, with a maximum between 28 and 42 DAP, (Fig. 2). Under our growth conditions this corresponded to the early-to-late cotyledonary stage of embryo development according to the developmental key of Piffanelli (1996). Expression was not detected in pre-cotyledonary stage embryos nor in mature, pre-dessiccated embryos. No transcript was detected in leaf, root or flowers (Fig. 3). Also, *FAE1* transcripts were not detected in dry seed, (data not shown). The appearance and disappearance of the *FAE1* transcript together with our failure to detect the transcript in other tissues suggest that the β-ketoacyl-CoA synthase activity of seeds is regulated at the level of transcription of embryo-specific genes. These results are consistant with the observation that Arabidopsis *FAE1* is expressed in the seed only and that mutation of this gene affects VLCFA synthesis in the seed only (James et al. 1995; Kunst et al. 1992). The results are also consistant with the appearance of elongase activity at 35 days post-anthesis in developing seed of *Brassica napus* grown under similar conditions Creach et al. (1992).

## *FAE1* is a candidate gene for the E1 locus

Erucic acid content in rapeseed oil is determined by the synthesis of the substrate, erucoyl-CoA, by the acyl-CoA elongase and by its incorporation into TAG by the action of acyltransferases specific for positions 1 and 3 of the glycerol moiety. Studies in rapeseed have shown that the synthesis of erucic acid occurs via two sequential condensation reactions controlled by elongase activities present in HEAR but absent in LEAR varieties (Stumpf and Pollard 1983). Thus, at least one of the two loci involved in erucic acid accumulation controls elongation activity. Arabidopsis *fae1* mutants have drastically reduced levels of seed VLCFAs (James and Dooner 1990; Lemieux et al. 1990) and a deficiency of both elongation activities (Kunst et al. 1992). Cloning of the Arabidopsis *FAE1* gene revealed homologies of the encoded protein to condensing enzymes (James et al. 1995).

E1 is one of the two loci controlling erucic acid content in rapeseed (Harvey and Downey 1963). RFLP analysis did not reveal any recombinant plants between CE8/*Hin*dIII and E1. Thus, one of the rapeseed *FAE1* genes is a good candidate for the E1 locus. The estimated confidence interval of the QTL position covers 2.7 cM (Fig. 6). Hence, several genes could be clustered in this area and may correspond to the E1 locus. Thus, similarly to the case of the *FAD3* candidate gene and linolenic acid content in rapeseed (Jourdren et al. 1996b), it is not possible to determine whether the

*FAE1* locus is localised adjacent to the E1 locus or corresponds to the E1 locus. Furthermore, because of the near identity in sequence of CE7 and CE8 it was not possible to obtain a probe that discriminated between the genes encoding the 2 cDNAs and hence which of the genes maps to the E1 locus. Such a gene-specific probe would be of great value when selecting for high C22 : 1 content, especially for use in haplodiploidisation schemes.

The linkage of the *FAE1* structural gene with the E1 locus is consistant with the inferrence of Ohlrogge et al. (1991) that the lack of enzymic activity in LEAR plants (Stumpf and Pollard 1983) suggests that at least one of the two loci involved in erucic acid accumulation controls elongation activity. The complementation of a LEAR variety by a transformation with a jojoba homologue of the *FAE1* gene (Lassner et al. 1996), confirmed that the LEAR phenotype is a consequence of a mutation in the structural gene encoding seed $\beta$-ketoacyl-CoA synthase or in a gene controlling the activity of the synthase.

Experiments designed to obtain a functional complementation of a LEAR plant with the *FAE1* cDNA originating from a HEAR segregant are in progress in order to test the hypothesis that the *FAE1* gene corresponds to the E1 locus.

## The identity of the E2 locus

If one of the rapeseed *FAE1* structural genes is confirmed as a gene corresponding to the E1 locus, what is the identity of the E2 locus? Since rapeseed possesses an amphidiploid genome, the E2 locus may be the homoeologous gene of *FAE1*. Other markers have been described close to E1 and E2 genomic regions on the rapeseed genetic map (Foisset et al. 1996). The three RFLP probes localised on the DY9 group (E1 group) revealed only one polymorphic locus, and thus it is unclear as to whether these two regions are homoeologous. Indirect evidence associating a *FAE1* gene with the E2 locus is suggested by the preliminary identification of the DY7 linkage group as being homeologous to that of DY9 (data not shown). This hypothesis will be tested using other combinations of probe and enzymes to map the second locus revealed by the CE8 probe. In addition, a polymorphic locus detected with the CE8 probe in the cross between varieties 'Stellar' and 'Drakkar' was found to be linked to a RAPD marker. This RAPD marker is common to both the 'Stellar/Drakkar' and 'Darmor-*bzh* /Yudal' maps and is linked to the E2 locus on the 'Darmor-*bzh*/Yudal' map, (Barret and Delourme unpublished).

Since elongase comprises three additional activities, the E2 locus could correspond to one of these structural genes or to a regulatory gene controlling the expression or activity of one or more of the three subunits. However, as pointed out by Ohlrogge et al. (1991), in the case of the LEAR phenotype a mutation in one of the steps subsequent to the condensation reaction could be deleterious because of sequestration of coenzyme A. Furthermore, the condensing enzyme has been shown to be the key regulatory subunit of the elongase complex (Cassagne et al. 1994). It is unlikely that the E2 locus corresponds to a seed-specific form of the acyltransferases since in the LEAR phenotype a mutation would have to alter substrate specificities of two enzymes, glycerol-3-phosphate acyltransferase and diacylglycerol acyltransferase. Thus, it is probable that the E2 locus corresponds to the second *FAE1* gene.

E1 and E2 loci do not contribute equally to erucic acid content in rapeseed (Jourdren et al. 1996a). The percentages of the total variation of the C22 : 1 level determined by the E1 and E2 loci were 56.4 and 28.6%, respectively, and 90.6% of the trait variation was explained when the effects of these two genes were combined. Several explanations to account for the difference in the contribution of the two genes are possible. Should each of the 2 cDNAs identified in this study correspond to the E1 and E2 loci, then the level of expression of each gene may not be equal in HEAR possibly as a consequence of parental dominance or transinactivation of duplicated sequences at unlinked loci (reviewed in Matzke and Matzke 1993). Thus, the variation in amino acid sequence could influence the kinetic properties of one or other of the two enzymes. Alternatively, transcriptional and/or post-transcriptional regulation or modification may differentially affect the activity of the enzymes. The possibility that different high erucic acid alleles may exist in different genotypes may reconcile the lack of variation observed by Thormann et al. (1996) between the effects of two loci controlling erucic acid content in rapeseed.

**Acknowledgements** The authors thank Martine Devic for critical appraisal of this manuscript and Michele Laudie and Cristel Berger for help with the automatic sequencer. The authors thank Drs. J.-C. Kader and P. Guerche for helpful comments during the preparation of this manuscript. This work was supported by a grant from the French Ministère de la Recherche et de la Technologie and the Organisation Nationale Interprofessionelle des Oléagineux, Paris.

## References

Altshul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. J Mol Biol 215 : 403–410

Bernerth R, Frentzen M (1990) Utilisation of erucoyl-CoA by acyltransferases from developing seeds of *Brassica napus* L. involved in triacylglycerol biosynthesis. Plant Sci 67 : 21–28

Cao Y, Oo KC, Huang AHC (1990) Lysophosphatidate acyltransferase in the microsomal fraction from maturing seeds of meadow foam (*Limnanthes alba*). Plant Physiol 1199–1206

Cassagne C, Lessire R, Bessoule J-J, Moreau P, Creach A, Schneider F, Sturbois B (1994) Biosynthesis of very long chain fatty acids in higher plants. Prog Lipid Res 33 : 55–69

Creach A, Lessire R (1993) Solubilisation of acyl-CoA elongases from developing rapeseed (*Brassica napus* L.) J Am Oil Chem Soc 70 : 1129–1133

Creach A, Lessire R, Cassagne C (1992) Biosynthesis of very long chain fatty acids by subcellular fractions of developing rapeseeds. In: Cherif A, Miled-Daoud DB, Marzouk B, Smaoui A, Zarrouk M (eds) Metabolism, structure and utilisation of plant lipids. Tunis, C.N.P., pp 428–431

de Silva J, Loader NM, Jarman C, Winhurst JHC, Hughes SG, Safford R (1990) The isolation and sequence analysis of two seed-expressed acyl carrier protein genes from *Brassica napus*. Plant Mol Biol 14 : 537–548

Devic M, Albert S, Delseny M, Roscoe TJ (1997) Efficient PCR walking on plant genomic DNA. Plant Physiol Biochem 35 : 331–339

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissues. Focus 12 : 13–15

Foisset N, Delourme R, Barret P, Hubert N, Landry BS, Renard M (1996) Molecular-mapping analysis in *Brassica napus* using isozyme, RAPD and RFLP markers on a double haploid progeny. Theor Appl Genet 93 : 1017–1025

Harvey BL, Downey RK (1963)The inheritance of erucic acid content in rapeseed (*Brassica napus* L.). Can J Plant Sci 44 : 104–111

James DW, Dooner HK (1990) Isolation of EMS-induced mutants in Arabidopsis altered in seed fatty acid composition. Theor Appl Genet 80 : 241–245

James DW, Lim E, Keller J, Plooy I, Dooner HK (1995) Directed tagging of the Arabidopsis *FATTY ACID ELONGATION (FAE1)* gene with the maize transposon *Activator*. Plant Cell 7 : 309–319

Jourdren C, Barret P, Horvais R, Foisset N, Delourme R, Renard M (1996a) Identification of RAPD markers linked to the loci controlling erucic acid level in rapeseed. Mol Breed 2 : 61–71

Jourdren C, Barret P, Brunel D, Delourme R, Renard M (1996b) Specific molecular markers of the genes controlling the linolenic acid level in rapeseed. Theor Appl Genet 93 : 512–518

Kater MM, Koningstein GM, Nijkamp HJJ, Stuitje AR (1991) cDNA cloning and expression of *Brassica napus* enoyl-acyl carrier protein reductase in *Escherichia coli*. Plant Mol Biol 17 : 895–909

Kay R, Chan A, Daly M, McPherson J (1987) Duplication of CaMV 35S promoter sequences creates a strong enhancer for plant genes. Science 236 : 1299–1302

Kunst L, Taylor DC, Underhill EW (1992) Fatty acid elongation in developing seeds of *Arabidopsis thaliana* Plant Physiol Biochem 30 : 425–434

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1 : 174–181

Lassner MW, Levering CK, Davies HM, Knutzon DS (1995) Lysophosphatidic acid acyltransferase from meadowfoam mediates insertion of erucic acid at the sn-2 position of triacylglycerol in transgenic rapeseed oil. Plant Physiol 109 : 1389–1394

Lassner MW, Lardizabal K, Metz JG (1996) A Jojoba $\beta$-ketoacyl-CoA synthase cDNA complements the canola Fatty Acid Elongation mutation in transgenic plants. The Plant Cell 8 : 281–292

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

Loader, NM, Woolner EM, Hellyer A, Slabas AR, Safford R (1993) Isolation and characterisation of two *Brassica napus* embryo acyl-ACP thioesterase cDNA clones. Plant Mol Biol 23 : 769–770

Matzke M, Matzke AJM (1993) Genomic imprinting in plants: parental effects and trans-inactivation phenomema. Annu Rev Plant Physiol Plant Mol Biol 44 : 53–76

Murphy IS, Sonntag JG (1991) Erucic, behenic: feedstocks of the 21st century. Inform 2 : 449–463

Ohlrogge JB, Browse J, Somerville C (1991) The genetics of plant lipids. Biochim Biophys Acta 1082 : 1–26

Piffanelli P (1996) Molecular and biochemical studies of stearoyl-ACP desaturase in gametophytic and sporophytic generations of plants. PhD thesis, University of East Anglia, Norwich, UK

Post-Beittenmiller D (1996) Biochemistry and molecular biology of wax production in plants. Annu Rev Plant Phys Plant Mol Biol 47 : 405–430

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Press, New York

Sharpe AG, Parkin AP, Keith DJ, Lydiate DJ (1995) Frequent non-reciprocal translocations in the amphidiploid genome of oilseed rape (*Brassica napus*). Genome 38 : 1112–1121

Stefansson RB (1983) The development of improved rapeseed cultivars. In: Kramer JKG, Sauer FD, Pigden WJ (eds) High and low erucic acid rapeseed oils. Academic Press, New York, pp 143–159

Stumpf PK, Pollard M (1983) Pathways of fatty acid biosynthesis in higher plants with particular reference to developing rapeseed. In: Kramer JKG, Sauer FD, Pigden WJ (eds) High and low erucic acid rapeseed oils. Academic Press, New York, pp 131–143

Szewc-McFadden AK, Kresowitch S, Bliek SM, Mitchell SE, McFerson JR (1996) Identification of polymorphic, conserved simple sequence repeats (SSRs) in cultivated *Brassica species*. Theor Appl Genet 93 : 534–538

Thormann CE, Romero J, Mantet J, Osborn TC (1996) Mapping loci controlling the concentrations of erucic and linolenic acids in seed oil of *Brassica napus* L. Theor Appl Genet 93 : 282–286

U N (1935) Genome analysis in Brassica with special reference to the experimental formation of *Brassica napus* and peculiar mode of fertilisation. Jpn J Bot 7 : 389–452

Von Wettstein P (1994) Waxes, cutin and suberin. In: Moore TS (ed) Lipid metabolism in plants. CRC Press, Bocca Raton

**Note added in Proof** Since the acceptance of this paper, a sequence corresponding to that of the CE7 cDNA of this report has been deposited in the database under the accession number: AF009563