Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical Systematics and Ecology 34 (2006) 327–337
biochemical systematics and ecology
www.elsevier.com/locate/biochemsyseco

Evolution of 'front-end' desaturases in *Echium* (Boraginaceae)

F. García-Maroto [a,*], J.A. Garrido-Cárdenas [a], M.A. Vilches-Ferrón [b], A. Mañas-Fernández [a], D. López Alonso [b]

[a] *Area de Bioquímica, Facultad de Ciencias Experimentales, Universidad de Almería, 04120 Almería, Spain*
[b] *Departamento de Biología Aplicada, Facultad de Ciencias Experimentales, Universidad de Almería, 04120 Almería, Spain*

Received 17 October 2005; accepted 30 November 2005

Abstract

'Front-end' desaturases of higher plants catalyse the desaturation of either fatty acids attached to phospholipids (Δ^6-desaturases), or the long-chain base of sphingolipids (Δ^8-desaturases). In a few plant families like the Boraginaceae, a Δ^6-desaturase activity is responsible for the synthesis of unusual fatty acids like the γ-linolenic acid (18:3$n-6$) or the octadecatetraenoic acid (18:4$n-3$). Due to their sporadic appearance in nature, the biological role of these fatty acids has been questioned, and remains unknown. In this paper we describe a Δ^6-desaturase related gene (*D6DESΨ*) from *Echium* (Boraginaceae), which likely represents a pseudogene recently originated in this evolutionary lineage. This finding allowed us to check whether selective pressure is acting on the gene encoding the functional desaturase (*D6DES*) that is responsible for the synthesis of GLA/OTA. Determination of the substitution rates for the *D6DES*/*D6DESΨ* pairs from three *Echium* species shows that they are considerably higher for *D6DESΨ* genes than for their functional counterparts. This provides further support for a possible biological role of the products of the active paralogues.

Keywords: Molecular evolution; Phylogenetic analysis; Plant front-end desaturases; Nucleotide rates

1. Introduction

Acyl-desaturases of the cellular membranes make up a very heterogeneous group of enzymes catalysing diverse oxidative reactions on lipid substrates (Shanklin and Cahoon, 1998). Among them, the so-called front-end desaturases generally introduce a double bond between the carboxyl group and a pre-existing unsaturation of the acyl chain (Napier et al., 1997). These enzymes are encoded by an ancient multigene family that has evolved to desaturate, with different regioselectivities, a variety of substrates, both in plant and animals (Napier et al., 2002; Sperling et al., 2003). Their protein structure is characterised by the presence of an N-terminal cytochrome b_5-domain along with a typical fatty acid desaturase domain that includes three highly conserved histidine boxes. Within this group are the Δ^6-desaturases of plants, which are responsible for the synthesis of the unusual γ-linolenic (GLA; 18:3$n-6$) and octadecatetraenoic (OTA; 18:4$n-3$) fatty acids in a limited number of plant families like the Boraginaceae. Though

* Corresponding author. Tel.: +34 950015033; fax: +34 950015008.
E-mail address: fgmaroto@ual.es (F. García-Maroto).

0305-1978/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bse.2005.11.007

the function of these compounds in the plant remains unknown, they are interesting for human health mainly because they act as precursors for the synthesis of long-chain polyunsaturated fatty acids with diverse biological activities (Gunstone, 1998). Functional Δ^6-desaturase genes have been previously cloned and characterised from *Echium* species such as *Echium pitardii* and *Echium gentianoides*, which are endemic to the Canary Islands (García-Maroto et al., 2002). The analysis of their genomic sequences showed that they are intron-less genes, as it is true for other front-end desaturases of higher plants (Whitney et al., 2003). A possible retrotranscriptional origin of these genes was suggested based on phylogenetic analysis (López Alonso et al., 2003). In plants, other members of the same family, Δ^8-desaturases, are able to act on sphingolipid substrates, introducing a double bond on the long-chain base (LCB) moiety of the molecule (Sperling et al., 2003).

Few cases of non-functional genes have been reported for desaturases. Among them, are genes encoding enzymes involved in the biosynthesis of insect pheromones (Roelofs et al., 2002), and the human stearoyl-CoA desaturase (Zhang et al., 1999). In spite of the large number of genes described within the front-end desaturase family, the existence of pseudogenes has not been reported to date. The study of pseudogenes has focused a considerable interest by evolutionary biologists because their study may shed light on the evolutionary mechanisms of protein families. Pseudogenes are non-functional sequences that have been disabled, usually due to the lack of promoter activity. Less frequently, transcripts are generated but they cannot be translated, or the protein product is inactive. Two main mechanisms are likely to generate pseudogenes (Vanin, 1985; Mighell et al., 2000). Unprocessed or duplicated pseudogenes are thought to have arisen from a gene duplication event, with the subsequent functional loss of one of the copies. Processed pseudogenes, also called retropseudogenes, are the result of retrotranscription from processed mRNA, followed by further integration in the genome. As a result of non-functionality, pseudogenes are usually released of selective pressure and accumulate different mutations, such as frameshifts, in-frame stop codons or interspersed repeats that are often used as diagnostic characters (Zhang and Gerstein, 2004). When originated by retrotranscription, pseudogenes can be sometimes recognised by the lack of introns and recognisable promoter relative to its functional paralog, and the existence of short direct repeats flanking the coding sequence along with 3′-poliA stretches (Loguercio and Wilkins, 1998; Bennetzen, 2000). Processed pseudogenes appear to be less frequent in plants than in animals. Indeed, presently there are few reports (Drouin and Dover, 1987; Loguercio and Wilkins, 1998; Kvarnheden et al., 1998), though, with the progress in the analysis of the complete plant genomes, the number of characterised retropseudogenes will surely expand.

In this article we report a detailed characterisation of a genomic sequence from *Echium* that is highly related to the functional Δ^6-desaturase gene. A detailed analysis reveals that it represents a disabled gene version with a putative retrotranscriptional origin. This finding provided us a reference to check whether selection is acting on the gene encoding the functional desaturase (*D6DES*) that is responsible for the synthesis of GLA/OTA by comparison of the substitution rates.

2. Materials and methods

2.1. Biological material

Leaf material from seedlings was used as a DNA source for *E. pitardii* A. Chev. ex D. Bramwell (=*Echium lancerottense* Lems et Holz), *E. gentianoides* Webb ex Coincy, and *Echium sabulicola* Pomel. Different tissues of *E. pitardii* used in Northern blot analysis were sampled from plants grown at 25 °C, under the controlled conditions of growth cabinets with a 16 h light/8 h dark photoperiod.

2.2. Cloning of Δ^6-desaturase related genes in Echium

Cloning of the *EPD6DESΨ* gene was achieved by polymerase chain reaction (PCR) amplification of a partial sequence, followed by bi-directional genomic walking through inverse PCR (IPCR). Initially, a 550 bp PCR fragment corresponding to amino acid positions 187–369 (Fig. 1) was obtained using the degenerate oligonucleotides BO-1 (5′-AT(A/C)AG(T/C)AT(T/C)GGTTGGTGGAA(A/G)TGG-3′) and BO-2 (5′-AATCCACC(A/G)TG(A/G)AACCA(A/G)TCCAT-3′) as primers, and *E. pitardii* genomic DNA as a template. The PCR program consisted of

a denaturation step of 5 min at 96 °C, followed by 38 cycles of 10 s at 94 °C, 1 min at 48 °C and 1 min at 72 °C, ending with a 5 min step at 72 °C. The product was cloned into the pGEM-T-Easy® vector (Promega), and a small plasmid library was generated and subsequently screened by plasmid restriction analysis and Southern blot under non-stringent conditions using Δ^6- and Δ^8-desaturase specific probes from *Echium* sp. (García-Maroto et al., 2002). Selected clones were sequenced on both strands using a Perkin–Elmer ABI-310 DNA automated sequencer and the Big Dye v3.1 chemistry. Three new sequences were obtained, and the corresponding genes designated as *EPD6DESΨ*, *EPD8DES1* and *EPD8DES2*. To complete the sequence of *EPD6DESΨ*, two nested upstream primers (5′-GTAAGAAGCTTGGAGGATACAACGATG-3′ and 5′GGCTTACGAAGAATTTTGATAAAGAATCG-3′) and two nested downstream primers (5′-TTCATGACTGGAATTCAACAACTTCAA-3′ and 5′-CTTCAATTCTCTT-TAAGTCACGTCTCTT-3′) were designed to perform the IPCR essentially as described by Ochman et al. (1990). From the several restriction enzymes used to cut the DNA, *Dra*I and *Vsp*I resulted most useful, allowing, after circularization and two nested rounds of PCR amplification, the obtaining of suitable 1.7 and 1.8 Kbp long fragments, respectively, that were analyzed and sequenced as described above. A genomic DNA fragment containing the whole coding sequence for *EPD6DESΨ* was further obtained by PCR amplification using suitable upstream (5′-AAGTCGA-CATATAATGGATTCTGCAAGAAG-3′) and downstream (5′-AAGGATCCATATTACAAACTCGACGTCCA-GAGC-3′) primers and the reading-proof polymerase Pfx (Invitrogen). The fragment was further cloned and sequenced as indicated before.

Cloning of partial fragments for *EPD6DESΨ* orthologues in other *Echium* species (*E. gentianoides* and *E. sabulicola*) was achieved by using the same degenerated primers BO-1 and BO-2. In this case PCR was performed under less stringent conditions using an annealing temperature of 45 °C. Clones were screened by Southern blot using a *EPD6DESΨ*-specific probe under low stringency conditions.

2.3. Phylogenetic analysis

Alignment of the amino acid and nucleotide sequences was performed with the program Clustal X v.1.7 (Thompson et al., 1997) using the default settings, and manual adjustments as required after visual inspection. Protein sequences were used to reconstruct the phylogenetic trees since they are less prone to the noise introduced by multiple substitutions with highly diverged sequences. The amino acid matrix used in Fig. 2a comprised a total of 419 residues that excluded few amino acids, of uncertain alignment, at the C- and N-terminus. The alignment used in Fig. 2b included 183 amino acid positions, derived from available partial sequences corresponding to the middle coding region of the desaturases. For those genes that were not previously cloned, these sequences were obtained by PCR using the BO-1/BO-2 primers (see previous section). Phylogenetic trees were generated using different approaches as indicated in the legends to the figures. Bayesian analysis (Huelsenbeck and Ronquist, 2001) was conducted with MrBayes v3.1 (Ronquist and Huelsenbeck, 2003), Maximum Parsimony (Eck and Dayhoff, 1966) was performed using PAUP 4.0b10 (Swofford, 2003), and Minimum Evolution (Rzhetsky and Nei, 1992) as implemented in the MEGA package v3.0 (Kumar et al., 2004). The ProtTest software v1.2.6 (Abascal et al., 2005) was used to infer the best-fit model of protein evolution for our data set. The Jones–Taylor–Thornton (JTT) metric (Jones et al., 1992) was favored both under the Akaike and Bayesian criteria, and it was therefore used in minimum evolution (ME) and Bayesian analysis (BA). Rooting of the trees was accomplished by using the Δ^6-desaturase sequence from the moss *Physcomitrella patens* as the outgroup.

BA analysis of the whole protein data set (Fig. 2a) was conducted with prior for the amino acid model set to JTT (aamodelpr = Jones), and estimation of Γ (shape parameter of gamma distribution) and I (proportion of invariable sites) parameters was performed by MrBayes (not previously fixed), as recommended by the software authors. Similar results were obtained when the model jumping option (aamodelpr = mixed) over the 10 different amino acid models implemented in MrBayes was used. A duplicate analysis of 500,000 generations was performed with four Markov chains starting from a random tree. Trees were sampled every 250 generations, and a burnin value of 100 samples was applied. The majority-rule consensus tree containing the posterior probabilities (the percentage of times a particular clade is recovered) of the phylogeny was determined, and branches supported by a probability lower than 80% were further collapsed.

Maximum parsimony (MP) analysis (Fig. 2a) was performed using tree bisection reconnection (TBR) branch swapping and 100 random stepwise additions. Gaps in the sequence data were treated as missing data. Bootstrap values were calculated using 1000 TBR replicates, with 10 random taxon additions per bootstrap replicate.

For ME analysis (Fig. 2a and b), the JTT model was used along with the gamma shape parameter obtained from the BA analysis, and the search was performed by the close-neighbor-interchange (CNI) method set to level 1, using a neighbor joining (NJ) initial tree. Node confidence was checked by bootstrap on one thousand replicates.

Synonymous (d*S*) and non-synonymous (d*N*) substitution rates per site were calculated from the codon-aligned partial sequences available for the central coding region. The Kumar method was used, a variant of the Pamilo–Bianchi–Li method (Pamilo and Bianchi, 1993; Li, 1993) that is included in the MEGA v3.0 software.

For the amino acid sequences selected in Fig. 1, the alignment was visualised using the ''Boxshade'' v. 3.21 program.

2.4. Southern and Northern blot analysis

Genomic DNA was isolated from *Echium* seedlings by a CTAB-based extraction procedure (Taylor and Powel, 1982). DNA (about 5 μg) was restricted with the appropriate restriction enzymes, separated on a 0.8% agarose gel, and transferred by capillarity onto Hybond®-N^+ nylon membranes (Amersham). Filters were UV-crosslinked, prehybridised at 42 °C during 5 h in the 50% formamide/high SDS buffer recommended by the DIG manufacturer (Boehringer–Mannheim), and hybridised at the same temperature and same buffer solution (stringent conditions), containing the digoxigenin-labelled *EPD6DESΨ*-specific probe. High-stringency washes were performed twice at 65 °C during 15 min in buffer containing 0.5× SSC, 0.1% SDS, and the luminogenic substrate CSPD® was used for the detection following the instructions provided with the DIG detection kit. Images were obtained by exposure of Biomax® ML films (Kodak) for 10–25 min and final developing by standard procedures. The *EPD6DESΨ* probe was obtained by random primed labeling from a PCR fragment spanning 375 bp of the 3′-coding sequence and about 135 bp of the 3′-untranslated region. Total RNA was extracted from different tissues of either young or mature (flowering) *Echium* plants, following the method of Chang et al. (1993). About 10 μg per lane of total RNA were loaded onto an agarose/formaldehyde gel, electrophoretically separated, and transferred to Hybond®-N membranes. Filters were hybridised at 50 °C (stringent conditions) as described for Southern analysis, and using the same *EPD6DESΨ*-specific probe. Stringent washes were accomplished at 68 °C, and the detection of DIG-labelled probe was as indicated before. As a control, the filters were re-hybridised with either a *EPD6DES* specific probe from *E. pitardii*, or a 900 bp cDNA probe from tobacco, which encodes part of the cytosolic glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) gene (Shih et al., 1986). In this case, hybridisation was done in the same conditions, but the final washes were performed at 65 °C. For radioactive detection, radiolabelling with ^{32}P of DNA fragments used as hybridisation probes was done as described by Feimberg and Vogelstein (1984). Hybridisation was performed in Church buffer (Church and Gilbert, 1984) under stringent conditions (65 °C) Washes were accomplished at 65 °C, once in buffer 40 mM sodium phosphate pH 8.0, 5% SDS, for 15 min, followed by a 10 min stringent wash in the same buffer containing 1% SDS. Film exposure was two days at −70 °C with intensifying screen.

3. Results

3.1. Cloning and genomic organisation of a Δ⁶-desaturase-related gene from E. pitardii

We have performed an extensive search for Δ^6-desaturase-related sequences in *E. pitardii*, an endemic from the Canary Islands that accumulates high amounts of GLA (Guil-Guerrero et al., 2001). Degenerated primers, corresponding to highly conserved motives (ISIGWWKW and MDWFHGG, Fig. 1) found in front-end desaturases of plants, were used to amplify the central part of the coding sequence by PCR on genomic DNA. The PCR product (about 550 bp) was cloned in a T-vector, and a plasmid library was generated for further analysis. A careful screening, based on fragment restriction analysis and Southern hybridisation using different desaturase probes, allowed us the identification of two slightly different sequences related to Δ^6-desaturases. The whole genomic sequences for the two Δ^6-desaturase related clones were completed by reconstruction from overlapping clones obtained by IPCR (see Section 2.2). This allowed the design of primers flanking the coding sequence and the obtaining of the complete genomic clones by PCR using a proof-reading polymerase. As we will show later, these sequences likely represent two different alleles, designated *EPD6DESΨ1* and *EPD6DESΨ2*, of a single-copy gene. The coding sequences of the two *EPD6DESΨ* alleles differ by only six nucleotide changes resulting in two amino acid changes (one of them conservative),

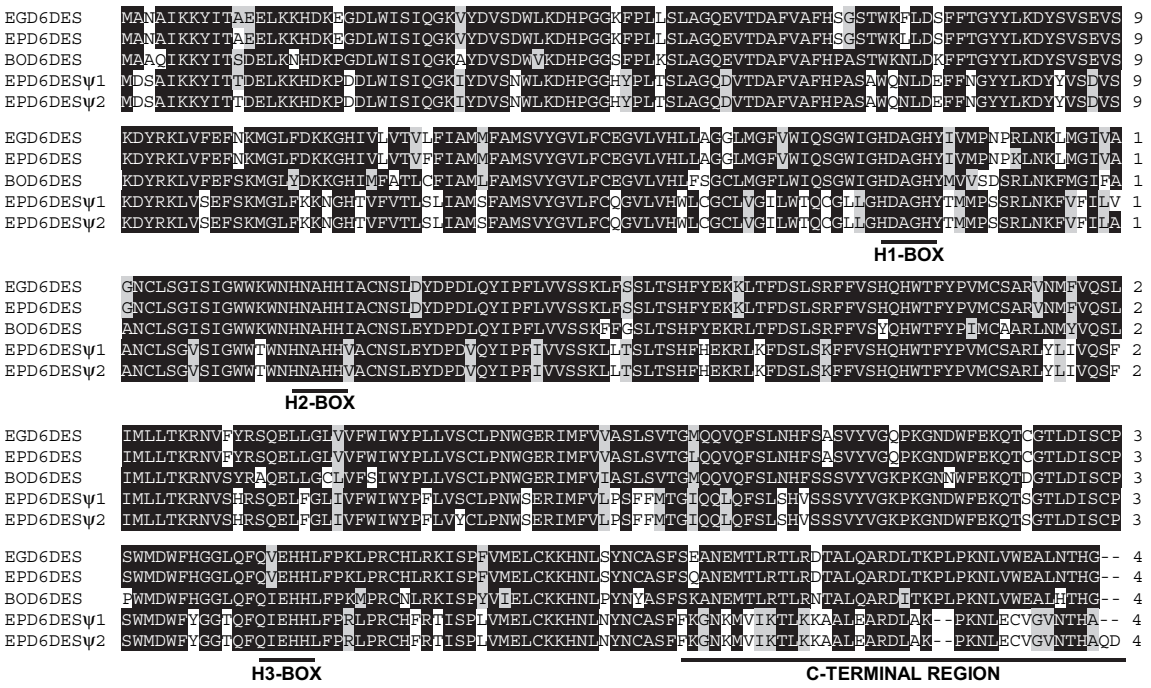

Fig. 1. Sequence comparison of the ORFs encoded by the *EPD6DESΨ* genes and related Δ^6-desaturases of Boraginaceae species. The amino acid sequences of Δ^6-desaturases of *Echium gentianoides* (EGD6DES), *E. pitardii* (EPD6DES), *Borago officinalis* (BOD6DES), and the ORFs encoded by the putative pseudogenes *EPD6DESΨ1* and *EPD6DESΨ2* (accession numbers in legend to Fig. 2) were aligned with the ClustalX (v1.7) software. The Boxshade program was used to highlight the homology between protein sequences. Shading is applied when there is agreement for a fraction of sequences above 0.5. Amino acids identical to EGD6DES are enclosed in black boxes, and similar residues in grey. Conserved histidine boxes, and the more diverged region, at the C-terminus, for EPD6DESΨ are underlined.

and a remarkable frame-shift leading to an extension of two amino acids at the C-terminus of *EPD6DESΨ2* (Fig. 1). *EPD6DESΨ* encodes a 448 amino acids protein showing an 82% of similarity (74% identity) to the functional Δ^6-desaturase of *E. pitardii* (EPD6DES, Fig. 1). In spite of the high similarity among those proteins, there are important changes, mainly those affecting the carboxyl-end of the protein, a region that is pretty well conserved in front-end desaturases (Fig. 1). Phylogenetic analysis, using the deduced amino acid sequences (Fig. 2a), indicated that *EPD6DESΨ* was closely related (node support $>95\%$) to functional Δ^6-desaturase genes of the Boraginaceae species. *EPD6DESΨ* genes appear as sister to the highly supported clade ($>85\%$ support) of functional Δ^6-desaturases of Boraginaceae, including *Echium* and *Borago* species. Moreover, *EPD6DESΨ* clustered apart from the main phylogenetic group of Δ^8-desaturases including that of the Boraginaceae *Borago officinalis* (Fig. 2a). As a result of the same initial screening, two partial sequences of *E. pitardii* (*EPD8DES1/2*) were also identified that clustered (95% bootstrap support) within the main Δ^8-desaturase group, including the *Borago* sphingolipid desaturase (Fig. 2b). These clones likely correspond to Δ^8-desaturase genes of *E. pitardii*, and represent a distinct evolutive lineage in the Boraginaceae apart from that of Δ^6-desaturases.

To investigate the genomic organisation of *EPD6DESΨ*, and the origin of the two similar versions of this gene, a Southern blot analysis was performed using genomic DNA of *E. pitardii*. Single hybridisation signals were obtained corresponding to fragments of the expected size (Fig. 3), when the DNA was digested with five different restriction enzymes cutting at least once outside of the coding region. This indicates that *EPD6DESΨ* is a single-copy gene, and that the two sequences, designated as *EPD6DESΨ1* and *2*, correspond to equivalent chromosomal locations in the diploid genome of *E. pitardii*.

Sequences flanking the coding region of both *EPD6DES* and *EPD6DESΨ* have been obtained in our lab (unpublished results). Comparison of the proximal sequences upstream the ATG for *EPD6DESΨ* and the functional *EPD6DES* does not reveal any homology (not shown). A short direct repeat (TNATCTCCATNCT) was found flanking the coding sequence, as well as putative polyadenylation signals, though the presence of a remnant polyA tail was not detected downstream the stop codon (not shown).

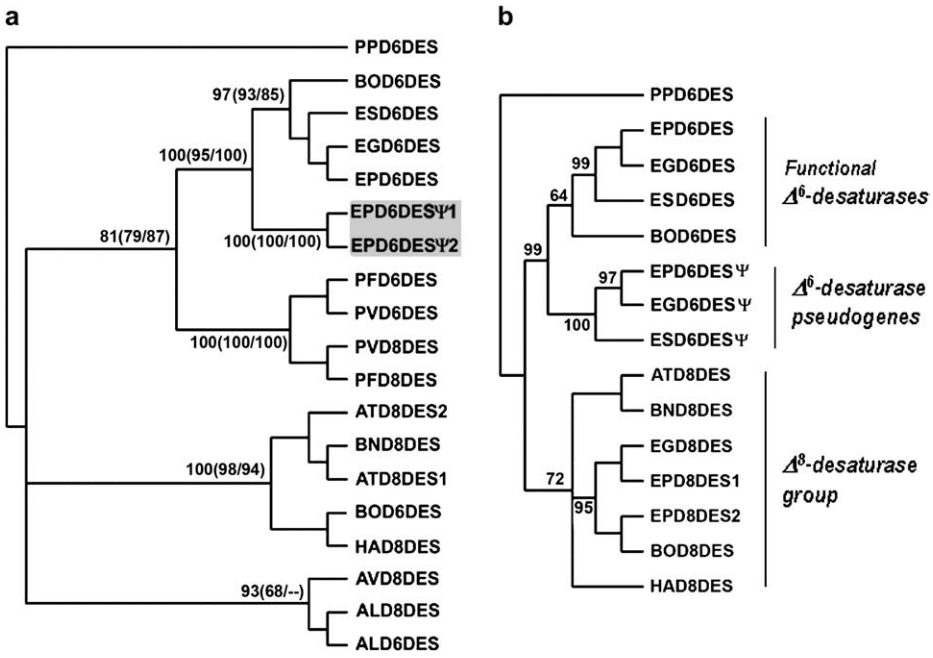

Fig. 2. Phylogenetic analysis of front-end desaturases from plants. Sequences are designated by using the first letter of the genus and the species name, followed by the protein name, D6DES or D8DES, for Δ^6- or Δ^8-desaturases, respectively. PP: *Physcomitrella patens* (acc. no. CAA11032); PF: *Primula farinosa* (acc. no. AAP23034 and AAP23033, for PFD6DES and PFD8DES, respectively); PV: *Primula vialli* (AAP23036 and AAP23035, for PVD6DES and PVD8DES); EP: *Echium pitardii* (AAL23581, DQ069309 and DQ069310 for EPD6DES, EPD8DES1 and EPD8DES2); BO: *Borago officinalis* (AAC49700 and AAG43277, for BOD6DES and BOD8DES); ES: *E. sabulicola* (DQ067612); EG: *E. gentianoides* (AAL23580 and DQ069308 for EGD6DES and EGD8DES); AT: *Arabidopsis thaliana* (CAA11858); BN: *Brassica napus* (CAA11857); HA: *Helianthus annuus* (CAA60621); AV: *Aquilegia vulgaris* (AAN03619); AL: *Anemone leveillei* (AAQ10731 and AAQ10732, for ALD6DES and ALD8DES). Pseudogenes are designated ending with the Ψ symbol (acc. no. EPD6DES Ψ1: DQ067610; EPD6DES Ψ2: DQ067611; ESD6DES Ψ: DQ067613; EGD6DES Ψ: DQ067614). (a) Tree reconstructed by Bayesian analysis on the whole protein data set, as implemented in the MrBayes 3.0b4 program. Numbers along relevant branches indicate posterior probabilities (as percentage) obtained after 500.000 generations. Branches under an 80% support were collapsed. Additional figures within the brackets are node bootstrap supports after 1000 replicates obtained from either maximum parsimony (before dash) or minimum evolution (after dash) methodologies on the same data set, for which a similar tree topology was also obtained. (b) Minimum evolution tree obtained from the data set of partial amino acid sequences of Boraginaceae species including putative *D6DESΨ* orthologues. Bootstrap values from 1000 replicates are indicated on relevant nodes.

3.2. Expression analysis of EPD6DESΨ

We have checked the expression pattern of *EPD6DESΨ* by Northern blot, using total RNA from different tissues of *E. pitardii* (Fig. 4). As it is shown, the *EPD6DESΨ* transcript was not detected in any of the main plant tissues when hybridisation was performed under our non-isotopic standard conditions (Fig. 4b). As a positive control, hybridisation of the same filter was performed with a probe for the functional Δ^6-desaturase gene of *E. pitardii* (*EPD6DES*). Clear hybridisation bands, corresponding to the known pattern of this gene (García-Maroto et al., 2002), were obtained in this case (Fig. 4c). *EPD6DESΨ* signals are only detected when hybridisation was performed under more sensitive conditions, using a radioactively labelled probe and a film exposure of two days (Fig. 4e). The pattern obtained was similar to that of *EPD6DES*, and it is therefore possible that the observed signals are simply the result of a weak cross-hybridisation. These results indicate that transcription of the *EPD6DESΨ* gene is taking place at a much lower rate, if any, compared to that of its functional counterpart, *EPD6DES*.

Since premature stop codons are not present in the coding sequence of *EPD6DESΨ*, we have investigated if the protein encoded by this gene (*EPD6DESΨ1* and *2* alleles) was able to perform some of the desaturation activities that could be expected for this family of proteins. The Δ^6-desaturase activity on fatty acids was checked by heterologous expression of the *EPD6DESΨ* coding sequence both in yeast and in transgenic tobacco plants following the methods previously used in the characterisation of the functional gene (García-Maroto et al., 2002). No desaturation

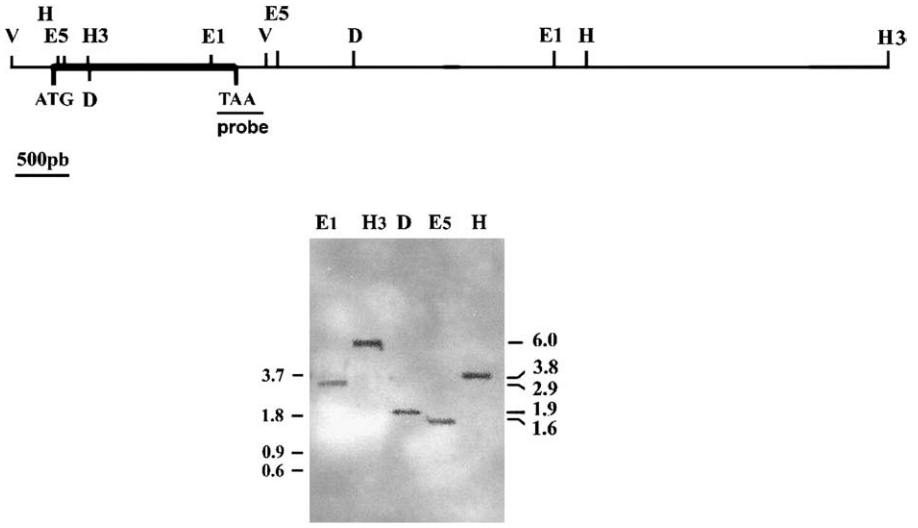

Fig. 3. Southern blot analysis of *EPD6DESΨ* in *Echium pitardii*. Genomic DNA was digested with *Eco*RI (E) *Hind*III (H3), *Dra*I (D), *Eco*RV (E5), or *Hae*III (H) restriction endonucleases, and hybridisation was performed using a *EPD6DESΨ*-specific probe under stringent conditions, as described in Section 2. Size markers positions (Kbp) are included. The restriction map for the *EPD6DESΨ* genomic sequence, deduced from the IPCR clones (see Section 2), is shown for reference. Position of the *EPD6DESΨ*-specific probe and map scale, are indicated by underlying bars.

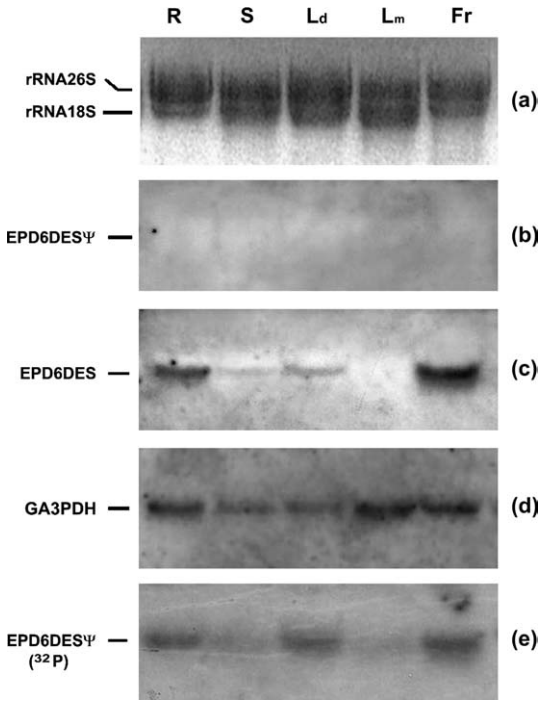

Fig. 4. Expression analysis of the *EPD6DESΨ* gene by Northern blot. Equivalent amounts of total RNA from young developing (Ld) or mature (Lm) leaves, stems (S), roots (R), or developing fruits (Fr) were hybridised either with a *EPD6DESΨ*-specific probe (panels b and e), or with a probe for the functional Δ^6-desaturase gene *EPD6DES* (panel c) under high-stringency conditions, as indicated in Section 2. The expression pattern for the constitutive gene encoding the cytosolic glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is also shown (panel d), besides the ethidium bromide staining of the RNA (panel a). Non-radioactive digoxigenin-based detection was used except for the experiment shown in panel e, where a more sensitive detection was achieved using a ^{32}P-labelled probe.

Table 1
Nucleotide substitution rates of functional Δ^6-desaturase genes and putative pseudogene counterparts from *Echium* species

Genes	d*S*	d*N*	d*S*/d*N*
EPD6DESΨ/EGD6DESΨ	0.066 (0.026)	0.026 (0.008)	2.5
EPD6DESΨ/ESD6DESΨ	0.117 (0.032)	0.039 (0.010)	3.0
EGD6DESΨ/ESD6DESΨ	0.111 (0.031)	0.025 (0.008)	4.4
Group average	0.098 (0.022)	0.030 (0.008)	3.3
EPD6DES/EGD6DES	0.055 (0.027)	0.006 (0.004)	9.2
EPD6DES/ESD6DES	0.209 (0.046)	0.008 (0.004)	26.1
EGD6DES/ESD6DES	0.191 (0.045)	0.004 (0.003)	47.8
Group average	0.152 (0.032)	0.006 (0.003)	25.3

Synonymous (d*S*) and non-synonymous (d*N*) rates per site are calculated according to Kumar method, as explained in Section 2. Values obtained are expressed besides standard errors inside brackets. Sequences are designated by using the first letter of the genus and the species name, EP: *Echium pitardii*; ES: *E. sabulicola*; EG: *E. gentianoides*.

products were detected in any case, contrary to what is found for the functional Δ^6-desaturase gene (*EPD6DES*) used as a positive control. On the other hand, it is known that proteins closely related to Δ^6-desaturases are able to desaturate sphingolipids by introducing a double bond at the Δ^8-position of the LCB. Therefore, Δ^8-desaturase activity on sphingolipids was also investigated, using a yeast expression system as previously described (Whitney et al., 2003). A negative result was also obtained (data not shown), indicating that EPD6DESΨ lacks both activities.

3.3. Analysis of the evolutionary rates of D6DESΨ in Echium

To this aim, we have cloned partial sequences of about 550 nts (see Section 2.2) for the putative orthologues of *EPD6DESΨ* in other *Echium* species, namely *E. gentianoides*, which is also a Macaronesian endemic, and the continental relative *E. sabulicola*. They have been designated *EGD6DESΨ* and *ESD6DESΨ*, respectively. For *E. sabulicola*, the complete sequence of the functional Δ^6-desaturase counterpart (*ESD6DES*) has also been obtained in this work. The *E. gentianoides* sequence was previously reported by us (García-Maroto et al., 2002). A cladistic analysis (Fig. 2b) for the deduced amino acid sequences shows that D6DESΨ sequences for the three *Echium* species cluster together (100% bootstrap) and are sister to the group made up by their functional counterparts, including the *Borago* Δ^6-desaturase. This indicates that *D6DESΨ* sequences of the three *Echium* species likely represent orthologous genes. Moreover, the D6DESΨ sequences appear as sister to the *Borago* functional gene (Fig. 2a), as further discussed.

We have determined the synonymous (d*S*) and non-synonymous (d*N*) substitution rates, for all species combinations within both functional and putative pseudogene groups (Table 1). In our calculations the Kumar method based on Kimura's two-parameter model was used, though other methods tested for comparison gave similar results. The d*S*/d*N* ratio measures how often nucleotide substitutions in a DNA sequence change the amino acid a measure that is often used to test whether a sequence is under selective constraints. This ratio ranged from 9.2 to 47.8 for functional Δ^6-desaturases of *Echium*. In contrast, d*S*/d*N* ratio for *D6DESΨ* sequences was clearly lower, ranging from 2.5 to 4.4. This means around eight times lower on average than those obtained for their functional counterparts. This difference increases up to 10 times if we consider substitution rates of *E. sabulicola* against *E. pitardii* or *E. gentianoides*, for which more reliable values are expected due to earlier divergence between Macaronesian and continental *Echium* species. This difference is probably a reflection of the different selective pressures acting over both kinds of genes.

4. Discussion

A great deal of evidence indicates that *EPD6DESΨ* is a pseudogene. First, as we have shown, it is transcribed at a much lower rate in those tissues where its functional counterpart is active (Fig. 4). Second, the encoded protein does not catalyse any of the desaturation reactions that are usually achieved by front-end desaturases of plants, neither Δ^6-desaturase nor Δ^8-desaturase activities. One cannot discard the possibility of a desaturase achieving another reaction in some minoritary tissues, or under some particular induction or developmental conditions. However, we have shown an accelerated substitution rate for *D6DESΨ* as compared to their functional paralogs (Table 1), indicating a relaxation

in the selective pressure acting on the non-functional gene. Those results, together, indicate that the gene has been disabled along evolution.

Additional data further suggest that *EPD6DESΨ* might be a processed pseudogene. Analysis of the genomic sequences in *Echium* (García-Maroto et al., 2002) and *Primula* (Sayanova et al., 2003) indicate that front-end desaturase genes of higher plants are intron-less. Therefore, in this particular case, the absence of introns in *EPD6DESΨ* cannot be used as a diagnostic character. However, the lack of similarity to the functional gene in the 5′-region upstream the ATG, as well as the presence of short direct repeats flanking the coding region, points in favour of a retrotranscriptional origin (Mighell et al., 2000). Though in this case 3′-polyA stretches were not found, it is known that this characteristic is not always a hallmark of retropseudogenes (Schmitz et al., 2004).

Phylogenetic analysis shows an early divergence between pseudogenes of *Echium* and functional desaturases of *Borago* and *Echium* (Fig. 2). Therefore, it is likely that *D6DESΨ* orthologues could also be found in other Boraginaceae species. In that sense we have recently undertaken a similar search in *Borago* that yielded new Δ^6-desaturase related sequences that are currently under investigation (results not shown). In this case it will be interesting to check if a loss of functionality has also occurred, as it happened in the *Echium* lineage.

A high conservation between pesudogene/functional gene pairs has been reported in cases where pseudogenes have resulted from recently duplicated paralogues. This is the case of the three human aquaporin pseudogenes (95–99% similarity to the functional gene), or the human stearoyl-CoA desaturase pseudogene (95% similarity) (Kondo et al., 2002; Zhang et al., 1999). In our case, we found an ancient origin for *D6DESΨ*, as displayed by a much lower similarity level to the functional product (87% for *E. pitardii*). Typical pseudogenes often show numerous insertions/deletions which usually result in frame-shift mutations. In the case of EPD6DESΨ, a high number of amino acid changes are observed, primarily at the C-terminus, even though an almost intact ORF has been conserved. Moreover, histidine boxes that are critical for desaturase activity have also been maintained. This conservation suggests that the pseudogenisation of the *D6DESΨ* is probably recent. It that sense, a d*S*/d*N* ratio > 1 for *D6DESΨ* genes can also be interpreted as an indication of a recent relaxation of the selective constrains. Recent loss of function of paralogous genes has also been described in *Ostrinia* (Lepidoptera) for desaturase genes involved in pheromone production (Roelofs et al., 2002).

Biological functions of the products for both kinds of front-end plant desaturases, Δ^6- and Δ^8-desaturases, remain largely unknown. In particular, Δ^6-desaturase activity is relatively rare among plant families (Gunstone, 1992), and the gene is even lacking in some organisms like *Arabidopsis*. This could lead to the notion that Δ^6-desaturation products such as GLA or OTA, are either redundant to other common polyunsaturated fatty acids, or they are simply by-products, lacking any function. Finding of pseudogenes allows us to establish a reference point that is devoid of selective constrains, allowing comparison of the substitution rates. Both the high d*S*/d*N* ratios of active Δ^6-desaturases, and the fact that they are about eight times higher than for the pseudogenes, argue in favour of the long-term existence of some unknown biological role for those unusual fatty acids in plants.

Pseudogenes have been proposed to serve as a sequence pool for generating genetic diversity (Balakirev and Ayala, 2003). Genes and pseudogenes can recombine and produce new genes, and it is paradigmatic the example of the human immune system (Vargas-Madrazo et al., 1995). Moth pheromone desaturase genes compose a family subjected to a birth-and-death evolutionary process (Roelofs et al., 2002) that shows interesting parallelisms to the front-end desaturases. Resurrection of a functional Δ^{14}-desaturase gene in a non-functional lineage of moths and flies has been described in *Ostrinia* leading to a major shift in the pheromone blend, thus illustrating the potential role of pseudogenes. Within the front-end desaturase family, it has been recognised the relative easiness to which genes with Δ^6- or Δ^8-desaturase activities are generated in relatively short evolutionary periods (López Alonso et al., 2003). This is well illustrated in the case of the Δ^6/Δ^8-desaturase genes of *Primula* and *Anemone* (see Fig. 2a), where both activities have evolved independently to those of the Boraginaceae lineage for example. In this context, it is tempting to speculate that the presence of pseudogenes could contribute to generate variability in the front-end desaturase family.

Acknowledgements

José A. Garrido-Cárdenas and Aurora Mañas-Fernández were recipients of two fellowships from the Spanish Ministerio de Ciencia y Tenología (M.C.Y.T.) and from the Junta de Andalucía, respectively. The work was financially

supported by a grant from the M.C.Y.T. (AGL2001-3453). We also thank the Estación Experimental Las Palmerillas (CAJAMAR) for providing facilities and help in plant cultivation. We are indebted to Drs. Johnathan Napier and Louise Michaelson, for the assay of Δ^8-desaturase activity on the EPD6DESΨ product, and to Dr. Pablo Vargas for helpful suggestions and revision of the manuscript.

References

Abascal, F., Zardoya, R., Posada, D., 2005. A ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.

Balakirev, E.S., Ayala, F.J., 2003. Pseudogenes: are they "junk" or functional DNA? Annu. Rev. Genet. 37, 123–151.

Bennetzen, J.L., 2000. Transposable element contributions to plant gene and genome evolution. Plant Mol. Biol. 42, 251–269.

Chang, S., Puryear, J., Cairney, J., 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Rep. 11, 13–116.

Church, G.M., Gilbert, W., 1984. Genomic sequencing. Proc. Natl Acad. Sci. USA 81, 1991–1995.

Drouin, G., Dover, G.A., 1987. A plant processed pseudogene. Nature 328, 557–558.

Eck, R.V., Dayhoff, M.O., 1966. Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Springs, Maryland.

Feimberg, A.P., Vogelstein, B., 1984. Addendum. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137, 266–267.

García-Maroto, F., Garrido-Cárdenas, J.A., Rodríguez-Ruiz, J., Vilches-Ferrón, M., Adam, A.C., Polaina, J., López Alonso, D., 2002. Cloning and molecular characterization of the Δ^6-desaturase from two *Echium* plant species: production of γ-linolenic acid by heterologous expression in yeast and tobacco. Lipids 37, 417–426.

Guil-Guerrero, J.L., Gómez-Mercado, F., Rodríguez-García, I., Campra-Madrid, P., García-Maroto, F., 2001. Occurrence and characterization of oils rich in γ-linolenic acid (III): the taxonomical value of the fatty acids in *Echium* (Boraginaceae). Phytochemistry 58, 117–120.

Gunstone, F.D., 1992. γ-Linolenic acid occurrence and physical and chemical properties. Prog. Lipid Res. 31, 145–161.

Gunstone, F.D., 1998. Movements towards tailor-made fats. Prog. Lipid Res. 37, 277–305.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17, 754–755.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. CABIOS 8, 275–282.

Kondo, H., Shimomura, I., Kishida, K., Kuriyama, H., Makino, Y., Nishizawa, H., Matsuda, M., Maeda, N., Nagaretani, H., Kihara, S., Kurachi, Y., Nakamura, T., Funahashi, T., Matsuzawa, Y., 2002. Human aquaporin adipose (*AQPap*) gene. Genomic structure, promoter analysis and functional mutation. Eur. J. Biochem. 269, 1814–1826.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Kvarnheden, A., Albert, V.A., Engström, P., 1998. Molecular evolution of *cdc2* pseudogenes in spruce (*Picea*). Plant Mol. Biol. 36, 767–774.

Li, W.-H., 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J. Mol. Evol. 36, 96–99.

Loguercio, L.L., Wilkins, T.A., 1998. Structural analysis of a *hmg-CoA* reductase pseudogene: insights into evolutionary processes affecting the *hmgr* gene family in allotetraploid cotton (*Gossypium hirsutum* L.). Curr. Genet. 34, 241–249.

López Alonso, D., García-Maroto, F., Rodriguez-Ruiz, J., Garrido, J.A., Vilches, M.A., 2003. Evolution of the membrane-bound fatty acid desaturases. Biochem. Syst. Ecol. 31, 1111–1124.

Mighell, A.J., Smith, N.R., Robinson, P.A., Markham, A.F., 2000. Vertebrate pseudogenes. FEBS Lett. 468, 109–114.

Napier, J.A., Sayanova, O., Stobart, A.K., Shewry, P.R., 1997. A new class of cytochrome b_5 fusion proteins. Biochem. J. 328, 717–720.

Napier, J.A., Michaelson, L.V., Sayanova, O., 2002. The role of cytochrome b_5 fusion desaturases in the synthesis of polyunsaturated fatty acids. Prost. Leuk. Essent. Fatty Acids 68, 135–143.

Ochman, H., Medhora, M., Garza, D., Hartl, D.L., 1990. Amplification of flanking sequences by inverse PCR. In: Innis, M.A., Gelfan, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: A Guide to Methods and Applications. Academic Press, London, pp. 219–227.

Pamilo, P., Bianchi, N.O., 1993. Evolution of the *Zfx* and *Zfy* genes: rates and interdependence between the genes. Mol. Biol. Evol. 10, 271–281.

Roelofs, W.L., Liu, W., Hao, G., Jiao, H., Rooney, A.P., Linn, C.E., 2002. Evolution of moth sex pheromones via ancestral genes. Proc. Natl Acad. Sci. USA 99, 13621–13626.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Rzhetsky, A., Nei, M., 1992. A simple method for estimating and testing minimum evolution trees. Mol. Biol. Evol. 9, 945–967.

Sayanova, O., Beaudoin, F., Michaelson, L.V., Shewry, P.R., Napier, J.A., 2003. Identification of *Primula* fatty acid Δ^6-desaturases with $n-3$ substrate preferences. FEBS Lett. 542, 100–104.

Schmitz, J., Churakov, G., Zischler, H., Brosius, J., 2004. A novel class of mammalian-specific tailless retropseudogenes. Genome Res. 14, 1911–1915.

Shanklin, J., Cahoon, E.B., 1998. Desaturation and related modifications of fatty acids. Annu. Rev. Plant Physiol. Plant Mol. Biol. 49, 611–641.

Shih, M.C., Lazan, G., Goodman, H.M., 1986. Evidence in favor of the symbiotic origin of chloroplasts: primary structure and evolution of tobacco glyceraldehyde-3-phosphate dehydrogenases. Cell 47, 73–80.

Sperling, P., Ternes, P., Zank, T.K., Heinz, E., 2003. The evolution of desaturases. Prost. Leuk. Essent. Fatty Acids 68, 73–95.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts.

Taylor, B., Powel, A., 1982. Isolation of plant DNA and RNA. Focus 4, 4–6.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24, 4876–4882.

Vanin, E.F., 1985. Processed pseudogenes: characteristics and evolution. Annu. Rev. Genet. 19, 253–272.

Vargas-Madrazo, E., Almagro, J.C., Lara-Ochoa, F., 1995. Structural repertoire in VH pseudogenes of immunoglobulins: comparison with human germline genes and human amino acid sequences. J. Mol. Biol. 246, 74–81.

Whitney, H.M., Michaelson, L.V., Sayanova, O., Pickett, J.A., Napier, J.A., 2003. Functional characterisation of two cytochrome b_5-fusion desaturases from *Anemone leveillei*: the unexpected identification of a fatty acid Δ^6-desaturase. Planta 217, 983–992.

Zhang, L., Ge, L., Parimoo, S., Stenn, K., Prouty, S.M., 1999. Human stearoyl-CoA desaturase: alternative transcripts generated from a single gene by usage of tandem polyadenylation sites. Biochem. J. 340, 255–264.

Zhang, Z., Gerstein, M., 2004. Large-scale analysis of pseudogenes in the human genome. Curr. Opin. Genet. Dev. 14, 328–335.