OTC and *AUL1*, two convergent and overlapping genes in the nuclear genome of *Arabidopsis thaliana*

Víctor Quesada, María Rosa Ponce, José Luis Micol*

División de Genética, Universidad Miguel Hernández, Campus de San Juan, 03550 Alicante, Spain

Received 12 August 1999; received in revised form 12 October 1999

Abstract In contrast to bacterial, fungal and vertebrate ornithine transcarbamylases (OTCs; EC 2.1.3.3), very little is known about the enzyme in plants. We report here the isolation of a T-DNA-tagged mutant displaying sensitivity to ornithine, whose characterization has allowed for the identification of several complementary and genomic DNA clones encoding the OTC and auxilin-like 1 (AUL1) proteins of the crucifer *Arabidopsis thaliana*. Transcript mapping revealed that at least 22 bp within the *OTC-AUL1* intercoding region are transcribed from both strands, which makes this one of the rarely described cases of convergent and overlapping transcription units in the nuclear genome of a multicellular eukaryote. Transcription of the *OTC* gene was shown to be ubiquitous in aerial organs of adult plants, whereas that of *AUL1* was obscured by the existence of a putative second copy of the gene. The *OTC-AUL1* locus maps at the bottom of chromosome 1.

Key words: Ornithine transcarbamylase gene; Auxilin-like 1 gene; Ornithine sensitivity mutation

1. Introduction

L-Citrulline formation from carbamoyl phosphate and L-ornithine in the de novo biosynthesis of arginine is catalyzed by ornithine transcarbamylase (OTC; EC 2.1.3.3; carbamoyl phosphate: L-ornithine carbamoyltransferase) [1–3]. As is the case for most other enzymes of the arginine metabolic pathway, OTC is ubiquitous in prokaryotes and eukaryotes, with 38–39% similarity between the rat or human and *Escherichia coli* [4]. The OTC function and expression pattern has undergone considerable changes during evolution, mainly as a consequence of the appearance of the urea cycle in primitive vertebrates. It is assumed that the gene encoding OTC acquired information for mitochondrial import during evolution after separation of *Saccharomyces cerevisiae* and *Neurospora crassa*, since it is localized in the cytosol in the former and in the mitochondria in the latter of these two eukaryotic microorganisms [5]. Among the genes encoding OTC, the human one is the best known, since it has been subjected to extensive studies, given that OTC deficiency is an X-linked inborn error of metabolism of the urea cycle causing hyperammonemia. More than 90 mutations in the human *OTC* gene have been characterized [6].

In contrast to bacterial, fungal and vertebrate OTCs, little has been studied about the enzyme in plants, where it is known to participate not only in arginine metabolism but also in related pathways such as pyrimidine nucleotide and polyamine biosynthesis [7]. OTC enzymatic activity has been studied in several plant species, its subcellular localization being reported cytosolic and mitochondrial in sugar cane [8] and chloroplastic in pea [9]. Only one gene encoding an OTC has so far been isolated and characterized in a plant species, *Pisum sativum* [10]. We report here the isolation of an *Arabidopsis* T-DNA-tagged mutant displaying sensitivity to ornithine. Its characterization has allowed for the cloning of complementary and genomic DNA clones corresponding to two novel genes, whose transcription units are convergent and partially overlap, one of them encoding the OTC of *Arabidopsis*.

2. Materials and methods

Wild-type (Ws-2, Col and L*er* ecotypes) and transgenic (lines carrying T-DNA insertions, Feldmann and Marks [11]) *Arabidopsis thaliana* (L.) Heyhn seeds were supplied by the Nottingham *Arabidopsis* Stock Centre and grown as described in [12]. A λ-PRL2 library made from equal amounts of mRNA from several tissues and developmental stages of plants of the Col ecotype [13], containing *Sal*I-*Not*I inserts ranging from 400 to 2000 bp was provided, as well as the expressed sequence tag (EST) clones used as probes, by the *Arabidopsis* Biological Resource Center (Ohio State University, OH, USA). Synthetic oligonucleotides were bought from Perkin-Elmer Applied Biosystems UK. Their nucleotide sequences (5′→3′) and positions in Fig. 2 are as follows: OTC3E, ATGGCGGCTGCAATGGCTTC (14–33); OTC5C, GAATCTCCGAAGAGTAGTACC (118–98); OTC5E, GTAGTTTCTCAATCAATAGAGGC (365–343); OTC3F, GCCATGCGTTATATCTAGGTC (663–683); OTC5F, TCCTCCCGCTTACCCATCTG (714–695); OTC5D, GACCATTGACAACTGGAACAC (1053–1033); OTC3D, TCCTTGCCAAATCATGGCCG (1070–1089); OTCF, ATCTCGACATTCAAGATTGTCC (1261–1282); OTC5, ATGGACAATCTTGAATGTCGAG (1284–1263); OTC5B, CATGTTGTTCCCATCTCCAAC (1418–1398); OTCR, GCATGCATGCGATTCTCCGC (1906–1887); OTC3B, GCTATAATGCTTCACTTGCTC (1914–1934); OTC3, ATCTCTAGATCTGTCTTCATCC (2136–2157); AUXD, TGCGCCACGTTGTTGAAGCTTGT (2460–2482); AUXC, CTTTATGTTCATCCCGACAAG (2498–2478); AUXE, ACCTTTGGCAGTCTGGTTTG (2805–2824); AUXF, AATGCCCTTTCGCGGGCTTC (3175–3194); AUXG, TAGAAGCAATGTGGAGGCAGC (3560–3540).

2.1. Sequence analysis

Sequencing reactions were carried out with ABI PRISM dye terminator cycle sequencing kits according to the instructions of the manufacturer. The Clustal X [14] and Treeview [15] programs were used for amino acid sequence alignments and construction and plotting of phylogenetic trees.

2.2. Synthesis of probes and screening of cDNA libraries

A λ-PRL2 library was screened to isolate *OTC* and auxilin-like 1 (*AUL1*) cDNA clones. A total of 10^5 plaque-forming units were screened, plating the phages on *E. coli* Y1090 ZL. Duplicate plaque lifts were made on nylon filters (Amersham Hybond N), which were

*Corresponding author. Fax: (34) 96 591 94 34.
E-mail: jlmicol@umh.es

Abbreviations: EST, expressed sequence tag; IPCR, inverse PCR; GAK, cyclin G-associated kinase

0014-5793/99/$20.00
PII: S0014-5793(99)01426-X

subjected to denaturation (5 min in 0.5 M NaOH, 1.5 M NaCl) and neutralization (15 min in 1 M Tris-HCl, 1.5 M NaCl), washed with 2×SSC (10 min), air- (15 min) and oven-dried (10 min at 80°C) and crosslinked (15 s under 70 000 μJ/cm^2). Pre-hybridization (3 h at 68°C) and hybridization (overnight at 68°C, in a total volume of 5 ml in sealed bags) were carried out in 0.25 M Na_2HPO_4 pH 7.2, 1 mM Na_2EDTA, 0.5% blocking reagent (Boehringer Mannheim) and 7% SDS [16] in a hybridization oven with shaking. Filters were hybridized with 15 ng/ml of already denatured (8 min at 100°C) probe.

The H76953 EST clone [17] was used as a template to make a probe against *OTC* cDNA. A miniprep of plasmid DNA of the EST clone was performed, 3 μg of which was digested overnight with 15 U of *Not*I and *Sal*I. The restriction fragments were electrophoresed in a 0.8% agarose gel and the band corresponding to the 0.7 kb insert purified with Geneclean II (BIO 101). Approximately 100 ng of such DNA was denatured (10 min at 95°C) to synthesize a probe by random priming with 400 ng of 10×p(dN)$_6$ (Boehringer Mannheim), in a 20 μl reaction mixture including 2 U Klenow enzyme (Promega) and 100 μM of each dNTP (dATP, dCTP, dGTP and a 3:1 mixture of dTTP and digoxigenin-11-dUTP). After 2 h at 37°C, the reaction was stopped by adding Na_2EDTA to a final concentration of 200 mM, ethanol-precipitated (using LiCl to a final concentration of 0.4 M) and the pellet was resuspended in H_2O.

For *AUL1* cDNA screening, 10 pmol of OTC3 and AUXC (Fig. 1) were used as primers and 100 ng of Col genomic DNA as a template in a 25 μl polymerase chain reaction (PCR) mixture, which included 1.5 U BioTaq and 200 μM of each dNTP. An amplification product of 363 bp was obtained, electrophoresed in a 0.8% agarose gel, purified with Geneclean II and used as a template in a second amplification performed under similar conditions, except that a 3:1 mixture of dTTP and digoxigenin-11-dUTP was used instead of only dTTP. Both thermocycling programs started with an initial denaturation of 2 min at 94°C, followed by 35 cycles of 30 s at 94°C, 15 s at 60°C and 1 min at 72°C and a final 10 min incubation at 72°C.

Immunological detection was performed with anti-DIG (Boehringer Mannheim) using the chemiluminescent alkaline phosphatase CDP-*Star* substrate (Boehringer Mannheim) according to the instructions of the manufacturer. Autoradiograms were obtained on Hyperfilm MP (Amersham) exposed for about 2 h. Several positively hybridizing phages were plaque-purified and rescreened and pZL-1 plasmids containing the cDNA inserts were in vivo-excised as described in [13].

2.3. Reverse transcriptase (RT-) PCR

To isolate total RNA, tissues were homogenized in 500 μl Trizol (Gibco BRL) and incubated for 5 min at room temperature. RNA was chloroform-extracted, isopropyl alcohol-precipitated and resuspended in H_2O. Genomic DNA was removed by adding 5 U DNase I (Gibco BRL) and incubated for 30 min at 37°C and then 10 min at 70°C, to inactivate the enzyme. First-strand cDNA synthesis was performed in a 20 μl reaction mixture containing 0.5 mM of each dNTP, 10 mM DTT, 200 U Superscript II enzyme (Gibco BRL), 40 U RNaseOUT, 400 ng 10×p(dN)$_6$ primers (Boehringer Mannheim) and approximately 1 μg RNA as a template. Samples were incubated for 10 min at 25°C and then 1 h at 42°C.

In order to distinguish between genomic DNA and cDNA amplification products, primers were designed to span at least one intron. Gene-specific primers (OTC3D and OTCR, OTC3 and AUXC; Fig. 1) were used to PCR amplify 1/20 of each first-strand cDNA synthesis reaction mixture, 0.5 U BioTaq, 10 pmol of each oligonucleotide and 200 μM of each dNTP, in a total volume of 25 μl. The thermocycling programs started with an initial denaturation at 94°C for 2 min, followed by 35 cycles of 30 s at 94°C, 15 s at 55°C and 1 min at 72°C, with a final extension step of 10 min at 72°C.

3. Results

3.1. Isolation of the OTC gene

In a large-scale screening for mutations causing salt-tolerant germination [18], we have isolated and studied several mutants derived from T-DNA mutagenesis. One such mutant strain, T78e, was found to carry three independent T-DNA insertions, none of which co-segregated with the salt-tolerant mutant phenotype. Following the procedure described in [12], an inverse PCR (IPCR) amplification product of about 1 kb was obtained using oligonucleotides (LB2 and ORI1, see [12]) targeting sequences within one of the T-DNA inserts in the T78e strain. It was shown to contain *Arabidopsis* genomic DNA flanking the left border of the DNA insert, including part of a gene that was named *OTC*, since its deduced amino acid sequence was found homologous to the OTCs of different organisms, especially to that of *P. sativum* [10]. Two oligonucleotides (OTC3B and OTC5B, see Fig. 1) were designed from the obtained partial sequence of the *OTC* gene, which allowed us to IPCR amplify a molecule of 900 bp from the DNA of Ws-2, the wild-type ancestor of T78e. Its characterization allowed us to identify a deletion of 47 bp, 27 bp away from the left border of the T-DNA insert (see Fig. 2), in the genomic DNA of T78e and provided 400 bp additional to the already determined sequence of the *OTC* gene.

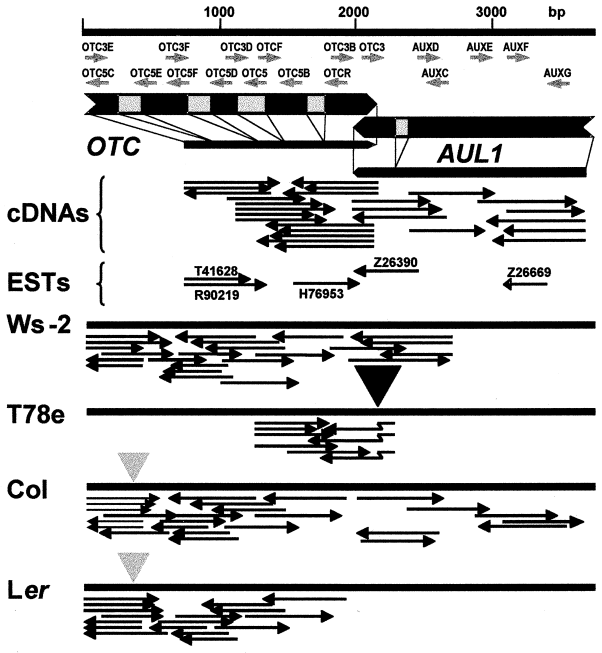

Fig. 1. Sequencing strategy of the *OTC-AUL1* locus. Indications are provided for introns (gray squares), exons (black squares), sequencing runs on genomic DNA or cDNA (black arrows), oligonucleotides used (gray arrows) and EST clones that were found identical to the genes under study. The black and gray triangle symbols indicate, respectively, a T-DNA insertion in the T78e strain and the polymorphic variant of the first *OTC* gene intron in Col and L*er*, which is 23 bp longer than that of Ws-2.

3.2. Isolation of OTC and AUL1 cDNA clones

One EST (EMBL accession number H76953) was found to be identical to a segment of the partial *OTC* gene sequence obtained. It was requested from the *Arabidopsis* Biological Resource Center and used as a template to synthesize a probe that was used to screen a cDNA library made from RNA of several tissues and different developmental stages of *Arabidopsis* [13]. Nine positive clones were obtained, the largest of which, of 1.5 kb (its 5′ and 3′ ends corresponding, respectively, to positions 1 and 2092 in Fig. 2), contained an open reading frame encoding a protein of 375 amino acids. Some segments of the nucleotide sequence of the *OTC* cDNA were found identical to two *Arabidopsis* ESTs (T41628 and

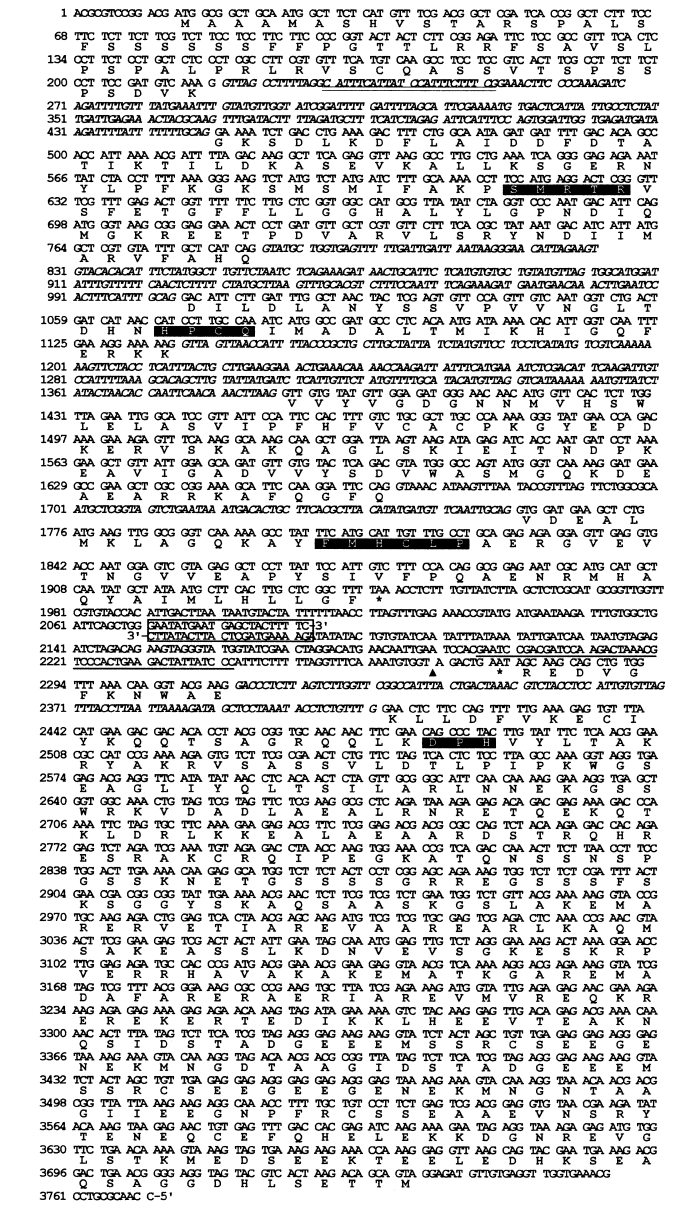

Fig. 2. Nucleotide and predicted amino acid sequences of the *OTC* and *AUL1 Arabidopsis* genes (EMBL accession numbers AJ000476, AJ002524 and AJ007450). Nucleotides are numbered on the left. Introns are indicated in italics, stop codons by asterisks. Nucleotides in the region of gene overlap are boxed. Double underlined nucleotides are present in the Col and L*er OTC* alleles, but absent in that of Ws-2. Underlined nucleotides are deleted in the T78e strain, which carries a T-DNA insertion in position 2269, indicated by the ▲ symbol. Amino acids considered to be important for OTC and AUL1 functions are black boxed (see text).

R90219), as well as similar to others (Z25591, H76031, N38372, Z34794, T04805 and AA395155).

Analysis of the *OTC* gene sequence revealed the existence of an open reading frame in the strand complementary to that encoding the OTC protein, whose deduced product was found similar to two proteins already described in vertebrate species: the *Bos taurus* auxilin, a clathrin binding protein found in some neurons [19], and the rat and human cyclin G-associated kinases (GAKs) [20,21]. We decided to name this gene *AUL1*, which had two related *Arabidopsis* ESTs, one of them identical (Z26390) and another similar (H76055).

Aiming to obtain *AUL1* cDNA clones, a screening was accomplished in the above-mentioned cDNA library, using a probe whose sequence was assumed to be outside the *OTC* transcription unit (2136–2498 in Fig. 2). A total of six positive clones were detected and sequenced, two of which contained a putative *AUL1* full-length open reading frame, encoding a protein of 452 amino acids (Fig. 2), their 5′ and 3′ ends corresponding, respectively, to positions 3771 and 2071 in Fig. 2. Neither the *AUL1* cDNAs nor those of *OTC* contained obvious polyadenylation signals.

3.3. Organization and map position of the OTC-AUL1 locus

From the *OTC* and *AUL1* cDNA sequences, we designed oligonucleotides in order to PCR amplify the corresponding genomic sequences (Fig. 1). The *OTC* transcription unit was shown to include four introns, whose sizes range from 96 to 254 bp. A single intron of 100 bp was found 1322 bp downstream of the start codon of the *AUL1* gene. The most striking feature was the fact that the two transcription units were convergent and overlapped 22 bp in their 3′ non-coding regions.

After obtaining *OTC* genomic sequences from the ecotypes Col, L*er* and Ws-2, we found an allelic variation between them: a sequence of 23 bp (5′-CATTTCATTATC-CATTTCTTTCG-3′) that is present in the most 5′ intron of the Col and L*er* alleles but absent from the Ws-2 allele (see Figs. 1 and 2). We used this polymorphism as a molecular

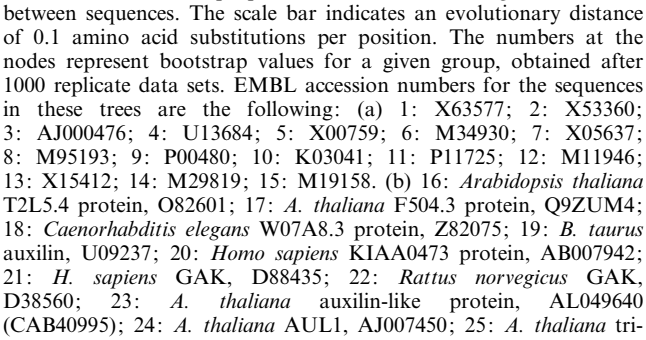

Fig. 3. Hypotheses of the molecular phylogenetic relationship of OTC proteins (a) and of the C-terminal region (the last 88 amino acids) of AUL1 and those of related proteins (b). The length of the horizontal branches is proportional to the estimated genetic distance between sequences. The scale bar indicates an evolutionary distance of 0.1 amino acid substitutions per position. The numbers at the nodes represent bootstrap values for a given group, obtained after 1000 replicate data sets. EMBL accession numbers for the sequences in these trees are the following: (a) 1: X63577; 2: X53360; 3: AJ000476; 4: U13684; 5: X00759; 6: M34930; 7: X05637; 8: M95193; 9: P00480; 10: K03041; 11: P11725; 12: M11946; 13: X15412; 14: M29819; 15: M19158. (b) 16: *Arabidopsis thaliana* T2L5.4 protein, O82601; 17: *A. thaliana* F504.3 protein, Q9ZUM4; 18: *Caenorhabditis elegans* W07A8.3 protein, Z82075; 19: *B. taurus* auxilin, U09237; 20: *Homo sapiens* KIAA0473 protein, AB007942; 21: *H. sapiens* GAK, D88435; 22: *Rattus norvegicus* GAK, D38560; 23: *A. thaliana* auxilin-like protein, AL049640 (CAB40995); 24: *A. thaliana* AUL1, AJ007450; 25: *A. thaliana* trichohyalin-like protein, Z99708 (CAB16841); 26: *S. pombe* AC17A5.12 protein, Z98849; 27: *S. cerevisiae* YDR320cp protein, U32517.

marker for linkage analysis, following the method described in [22]. The *OTC-AUL1* locus was found to map at the bottom of chromosome 1, 3.6 ± 2.0 cM away from the nga111 SSLP marker.

3.4. Analysis of the expression of the OTC and AUL1 genes

RT-PCR amplifications were made on total RNA from either vegetative leaves, flower buds or mature flowers of 4 week old Ws-2 and T78e plants, using the OTC3D and OTCR primers (Fig. 1), a single product of 487 bp being obtained in all cases. PCR amplifications performed with the OTC3D and OTCR oligonucleotides and Ws-2 or T78e genomic DNA as a template yielded a 868 bp product, as expected.

In order to assay the expression of *AUL1*, we used the OTC3 and AUXC primers (Fig. 1). Genomic DNA of Ws-2 yielded a major PCR product of the expected size, 363 bp, plus an unexpected one of 265 bp, the latter being identical in size to the single RT-PCR product obtained from total RNA. On the contrary, amplifications of both genomic DNA and total RNA of the mutant strain T78e yielded a single product of 265 bp, which was unexpected since the T-DNA insert in the *OTC-AUL1* locus maps within the region flanked by the primers used. When those PCR products of 265 bp were sequenced, they were found identical, lacking the *AUL1* 3′ intron as well as the 47 bp deletion present in the genomic DNA of T78e. Taken together, these results suggest that the *AUL1* gene is duplicated, that both copies are transcribed and that the copy that does not overlap with the *OTC* gene lacks the 3′ intron.

3.5. Phylogenetic analysis of the OTC and AUL1 protein sequences

The amino acid composition of the N-terminal region of the OTC protein is typical of the signal peptide of plant mitochondrial or chloroplastic proteins encoded by nuclear genes. It is rich in basic hydroxylated amino acids, mainly serine, arginine and alanine [23,24]. Multiple alignment of the OTC of *Arabidopsis* with those of other species points to the existence of several highly conserved residues, probably essential for their function. In fact, amino acids 123–127 (SMRTR) and 201–204 (HPCQ) seem to configure the binding site for carbamoyl phosphate, as it is known for the OTC of *E. coli* [25], while the residues 330–335 (FMHCLP) include the cysteine required to bind ornithine for its conversion into citrulline [26,27] (Fig. 2).

The percentage of identical amino acids between the OTC of *Arabidopsis* and that of *P. sativum* is 71% and with that of *Bacillus subtilis* 44.3%. For the remaining OTC proteins, the percentages of identity ranged from 42.7% (*Mus musculus*) to 31.8% (*S. cerevisiae*). The *Arabidopsis* OTC also shows similarity with the amino acid sequence deduced from ESTs of plant species such as *Ricinus communis*, *Medicago truncatula*, *Oryza sativa* and *Lycopersicon esculentum* (86.2, 76.0, 63.3 and 57.0% identity, respectively).

From the comparison of the amino acid sequences of *Arabidopsis* AUL1, *B. taurus* auxilin and the human and rat GAKs, a novel protein family can be proposed. Additional candidates to be included in this family are the *Arabidopsis* CAB16841 (which has been named trichohyalin-like) and T2L5.4 (a putative GAK) proteins and the *Schizosaccharomyces pombe* C17A5.12 and *S. cerevisiae* YDR320C proteins, all of them hypothetical, and the KIAA0473 human neuronal protein (Fig. 3). A recently sequenced *Arabidopsis* BAC (T20K18), which represents a segment of chromosome 4, contains a predicted gene encoding a protein that has been named auxilin-like (CAB40995), due to its similarity with the bovine auxilin. Multiple alignment of these proteins revealed that similarity was restricted to the C-terminal region, ranging from 28% (F504.3) to 70% (trichohyalin-like) of identity with regard to AUL1 (data not shown). AUL1 (Fig. 2) and all the above-mentioned proteins present the HPD residues in their C-terminal regions, a motif that characterizes the DnaJ domain of *B. taurus* auxilin [28] and other prokaryotic and eukaryotic proteins. Only two of the four *Arabidopsis* proteins studied cluster together with AUL1 in the tree shown in Fig. 3.

We have identified four *Arabidopsis* EST sequences similar to that of *AUL1*, two of which (Z26390 and Z26669, Fig. 1) seem to correspond to the transcripts of this gene, the remaining ones (H76055 and AA585826) representing putative paralogs. Three ESTs similar to *AUL1* were found in the *O. sativa* database (AU029942, AA752243 and AA750887), two in that of *Zea mays* (AI621869 and AI664895), one in those of *Citrus sinensis* (C21910) and *Gossypium hirsutum* (AI727268) and another in that of *Brassica napus* (AI352708). The amino acid sequence deduced from the latter showed 80% identity with that of AUL1.

3.6. Mutant phenotype caused by the T-DNA insertion in the OTC-AUL1 locus

We attempted to ascertain whether or not the T-DNA insertion in the *OTC-AUL1* locus caused some phenotypic effect. Since plant OTCs participate in three different reactions, one of which is the conversion of ornithine into citrulline, a key step in the biosynthesis of arginine [3], we studied the germination and growth of the mutant strain T78e and its ancestor ecotype, Ws-2, in media supplemented with different concentrations (0.1–10 mM) of the amino acids ornithine, citrulline and arginine. Plants were collected 28 days after sowing, no differences being found between T78e and Ws-2 individuals in media supplemented with citrulline or arginine. On the contrary, 1 mM ornithine was toxic for T78e, causing a reduction of its fresh weight (83.4%) substantially greater than that suffered by Ws-2 (7.8%). Co-segregation of the phenotype of sensitivity to ornithine with the T-DNA insert in the *OTC-AUL1* locus was confirmed in the F_3 progeny of a T78e × Col cross.

4. Discussion

The isolation of a genomic DNA flanking the left border of a T-DNA insert in the mutant strain T78e allowed us to characterize two novel *Arabidopsis* genes, whose transcription units converge and overlap in 22 nucleotides. This phenomenon has been described in prokaryotic [29,30] and mitochondrial [31] genomes but only a few cases are known in the nuclear genomes of fungi [32,33] and animals [34–36]. Although gene overlapping could be expected to occur in *Arabidopsis*, due to the small size of its nuclear genome, only two cases have been described previously [37,38]. Shintani et al. [35] have postulated that gene overlap arose during evolution as a result of chromosomal rearrangements that brought two previously separated and unrelated genes togeth-

er. Such a kind of rearrangement would be maintained only when each of the genes fortuitously found a polyadenylation signal on the non-coding strand of the other gene. In accordance with this, the *OTC* and *AUL1* genes do not seem to be structurally or functionally related, their 3′ untranslated regions displaying a high AT content (63 and 67%, respectively) as would be expected if the likelihood of finding a potential polyadenylation signal was to be increased. The isolation of cDNA clones from the *OTC* and *AUL1* genes suggests that their overlap does not impair transcription. Expression of both genes at the same time in the same cell will have to be demonstrated to validate such a hypothesis.

Structural analysis of the *OTC* gene in the ecotypes Col, L*er* and Ws-2 evidenced a polymorphism among them, consisting in a sequence of 23 bp in the most 5′ intron of the Col and L*er* alleles, which is absent from Ws-2. Such a polymorphism is easily visualized in agarose gels and allows for the use of the *OTC* gene as a molecular marker for linkage analysis. In fact, it allowed us to map the *OTC-AUL1* locus at the bottom of chromosome 1.

As has already been shown in *P. sativum* [10], analysis of *Arabidopsis OTC* gene expression indicates that it is active in different plant organs, as is to be expected from its participation in different metabolic pathways. The sensitivity to ornithine shown by the T78e mutant strain could be a consequence of the T-DNA insert in the *OTC-AUL1* locus, which is easily explained by a putative accumulation of ornithine, due to a partial loss of function of the mutant allele of the *OTC* gene. Given that transcription of the *OTC* gene, as measured by RT-PCR, is apparently normal in the T78e strain, the above-mentioned partial loss of function caused by the T-DNA insert might be due to a reduced efficiency of mRNA 3′ end formation. In fact, both the T-DNA insertion and the associated deletion found in the T78e strain are downstream the polyadenylated 3′ ends of the studied *OTC* cDNAs. Further characterization of this mutant strain will be useful in the dissection of OTC function in plants.

The *OTC* gene of *Arabidopsis* (this work) and that of *P. sativum* (named *argF* [10]) are the two only members of the OTC superfamily so far cloned in the plant kingdom. An observation of evolutionary interest is that these two plant proteins are closer to their prokaryotic orthologs than to those of fungi and vertebrates. In addition, we have shown that the *Arabidopsis OTC* gene is similar but not identical to several *Arabidopsis* ESTs, which might represent some of its paralogs.

The predicted product of the *AUL1* gene presents similarity with eukaryotic proteins such as the auxilin of *B. taurus* and the human and rat GAKs, as well as with others of unknown function. The sequence of *AUL1* also shows similarity with EST clones of several plant species. These results lead us to propose the existence of a novel gene family, not yet described in plants, that would count on further representatives in vertebrates, nematodes and fungi. It has been shown in bovine neurons that the auxilin DnaJ domain binds to heat shock protein 70, whose ATPase activity is then stimulated in order to remove the clathrin coat of vesicles that transport selective membrane components from the plasma membrane and the *trans*-Golgi network to the endosomal system [28,39]. Hence, interaction between AUL1 and other polypeptides can be considered likely, given that it seems to contain a DnaJ domain, which is known to be involved in protein-protein interactions. The induction of the expression of a probable *AUL1* ortholog in *B. napus* in response to the infection by the fungus *Leptosphaeria maculans* [40] suggests its participation in defense functions.

Our structural analysis of the *OTC-AUL1* locus in the T78e mutant strain indicates that *AUL1* transcription must be disturbed by the T-DNA insertion that interrupts the gene at 6 bp from its stop codon. However, amplifications of either genomic DNA or total RNA from the T78e mutant strain yielded a single product of 265 bp, whose sequencing revealed that they were identical, both lacking the intron contained in the *AUL1* gene. This observation can be explained assuming that there are two copies of the *AUL1* gene in the genome of *Arabidopsis*, both of them being transcribed. Only one of such *AUL1* copies, the one in the *OTC-AUL1* locus, would contain an intron. Further studies will be required to ascertain whether or not any of these two transcribed sequences is a pseudogene.

Acknowledgements: This research was supported by Grants from the Comisión Interministerial de Ciencia y Tecnología of Spain (BIO94-0253 and BIO97-1050). We thank H. Candela, A. Martínez-Laborda, J.M. Pérez-Pérez, P. Piqueras, P. Robles and A. Vera for comments on the manuscript, S. Gerber and J.M. Serrano for their expert technical help, the Nottingham *Arabidopsis* Stock Centre for providing seeds and the *Arabidopsis* Biological Resource Center for clones and K. Feldmann for making publicly available his-tagged lines. V. Quesada was a fellow of the Conselleria de Cultura, Educació i Ciència of the Generalitat Valenciana.

References

[1] Thompson, J.F. (1980) Amino acids and derivatives (Miflin B.J., Ed.), in: The Biochemistry of Plants, Vol. 5, pp. 375–402, Academic Press, New York.
[2] Huygen, R., Crabeel, M. and Glandsdorff, N. (1987) Eur. J. Biochem. 166, 371–377.
[3] Acaster, M.A., Scott-White, S. and Weitzman, P.D.J. (1989) J. Exp. Bot. 40, 1121–1125.
[4] Takiguchi, M., Murakami, T., Miura, S. and Mori, M. (1987) Proc. Natl. Acad. Sci. USA 84, 6136–6140.
[5] Takiguchi, M., Matsubasa, T., Amaya, Y. and Mori, M. (1989) BioEssays 10, 163–166.
[6] Tuchman, M., Plante, R.J., García-Pérez, M.A. and Rubio, V. (1996) Hum. Genet. 97, 274–276.
[7] Srivenugopal, K.S. and Adiaga, P.R. (1981) J. Biol. Chem. 256, 9532–9541.
[8] Glenn, E. and Maretzki, A. (1977) Plant Physiol. 60, 122–126.
[9] deRuiter, H. and Kolloffel, C. (1985) Plant Physiol. 77, 695–699.
[10] Williamson, C.L., Lake, M.R. and Slocum, R.D. (1996) Plant Mol. Biol. 31, 1087–1092.
[11] Feldmann, K.A. and Marks, M.D. (1987) Mol. Gen. Genet. 208, 1–9.
[12] Ponce, M.R., Quesada, V. and Micol, J.L. (1998) Plant J. 14, 497–501.
[13] D'Alessio, J.M., Bebee, R., Hartley, J.L., Noon, M.C. and Polayes, D. (1992) Focus 14, 76–79.
[14] Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997) Nucleic Acids Res. 25, 4876–4882.
[15] Page, R.D.M. (1996) Comput. Appl. Biosci. 12, 357–358.
[16] Engler-Blum, G., Meier, M., Frank, J. and Miller, G.A. (1993) Anal. Biochem. 210, 235–244.
[17] Newman, T., Bruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N. and Thomashow, M. (1994) Plant Physiol. 106, 1241–1255.
[18] Quesada, V., Ponce, M.R. and Micol, J.L. (2000) Genetics (in press).
[19] Schroder, S., Morris, S.A., Knorr, R., Plessmann, U., Weber, K., Nguyen, G.V. and Ungewickell, E. (1995) Eur. J. Biochem. 228, 297–304.

[20] Kanaoka, Y., Kimura, S.H., Okazaki, I., Ikeda, M. and Nojima, H. (1997) FEBS Lett. 402, 73–80.
[21] Kimura, S.H., Tsuruga, H., Yabuta, N., Endo, Y. and Nojima, H. (1997) Genomics 44, 179–187.
[22] Ponce, M.R., Robles, P. and Micol, J.L. (1999) Mol. Gen. Genet. 261, 408–415.
[23] Smeekens, S., Weisbeek, P. and Robinson, C. (1990) Trends Biochem. Sci. 15, 73–76.
[24] Glaser, E., Sjöling, S., Tanudji, M. and Whelan, J. (1998) Plant Mol. Biol. 38, 311–338.
[25] Wild, J.R. and Wales, M.E. (1990) Annu. Rev. Microbiol. 44, 193–218.
[26] Kraus, J.P., Hodges, P.E., Williamson, C.L., Horwich, A.L., Kalousek, F., Williams, K.R. and Rosenberg, L.E. (1985) Nucleic Acids Res. 13, 943–952.
[27] Kuo, L.C., Caron, C., Lee, S. and Herzberg, W. (1990) J. Mol. Biol. 211, 271–280.
[28] Kelley, W.L. (1998) Trends Biochem. Sci. 23, 222–227.
[29] Stragier, P. and Patte, J.C. (1983) J. Mol. Biol. 168, 333–350.
[30] Nomura, T., Aiba, H. and Ishihama, A. (1985) J. Biol. Chem. 260, 7122–7125.
[31] Normark, S., Bergström, S., Edlund, T., Grundström, T., Jaurin, B., Lindberg, F.P. and Olsson, O. (1983) Annu. Rev. Genet. 17, 499–525.
[32] Peterson, J.A. and Myers, A.M. (1993) Nucleic Acids Res. 21, 5500–5508.
[33] Gerads, M. and Ernst, J.F. (1998) Nucleic Acids Res. 26, 5061–5066.
[34] Spencer, C.A., Gietz, R.D. and Hodgetts, R.B. (1986) Nature 322, 279–281.
[35] Shintani, S., O'hUigin, C., Toyosawa, S., Michalová, V. and Klein, J. (1999) Genetics 152, 743–754.
[36] Williams, T. and Fried, M. (1986) Nature 322, 275–279.
[37] Ito, T., Takahashi, N., Shimura, Y. and Okada, K. (1997) Plant Cell Physiol. 38, 248–258.
[38] Glover, J., Grelon, M., Craig, S., Chaudhury, A. and Dennis, E. (1998) Plant J. 15, 345–356.
[39] Ungewickell, E., Ungewickell, H., Holstein, S.E.H., Lindner, R., Prasad, K., Barouch, W., Martin, B., Greene, L.E. and Eisenberg, E. (1995) Nature 378, 632–635.
[40] Fristensky, B., Balcerzak, M., Daifen, H. and Peijun, Z. (1999) Molecular Plant Pathology Online, http://www.bspp.org.uk/mppl/1999/0301fristensky.