Vitis **50** (1), 47–48 (2011)

Research Note

Alternative transcription and EST/ cDNA assisted characterization of *Vitis vinifera FCA* gene

Ke Jiang and L. Goertzen

Department of Biological Sciences, Auburn University, Auburn, USA

K e y w o r d s : Alternative transcription, EST, *FCA*, *Vitis vinifera.*

Flowering time of plants is controlled through interactions of several pathways and many genes (Putterill *et al.* 2004), integrating both endogenous and environmental signals. *FCA*, a flowering promoter, is an RNA processing factor influencing flowering time in *Arabidopsis*. *FCA* undergoes alternative transcription in an autoregulated fashion to control the amount of fully functional protein products, and thus flowering time in *Arabidopsis* (Quesada *et al.* 2003). As many as four transcript forms have been identified although 90 % of all transcripts were of two kinds, transcript α and γ (MacKnight *et al.* 1997) in *Arabidopsis*. The functional FCA protein encoded by transcript γ contains two RNA recognition motifs (RRM) and one WW domain. The FCA protein interacts with FY, a polyadenylation factor, forming a protein complex with both RNA recognition and cleavage-polyadenylation abilities. The FCA-FY complex directs cleavage and polyadenylation within intron 3 of *FCA* pre-mature mRNA, producing transcript form β, which leads to a non-functional protein in flower development (MacKnight *et al.* 1997, MacKnight *et al.* 2002). As a result, FCA protein promotes flowering while limiting its own amount by negative-feedback autoregulation of mature mRNAs in *Arabidopsis*. Multiple transcript forms also have been identified in cabbage, pea and rice, suggesting that FCA autoregulation may be conserved among these plants (MacKnight *et al.* 2002, Lee *et al.* 2005). However, the *fca* mutant of *Arabidopsis* is late flowering due to improper processing of *Arabidopsis Flowering Locus C* (*FLC*). As no *FLC* gene has been identified outside Brassicaceae, it is not clear whether *FCA* is relevant to the timing of flowering process in other plant species.

Vitis vinifera 'Pinot Noir' is the fourth complete sequenced plant genome. As an important horticultural crop, its flowering time is of high value in both fundamental research and crop improvement. Even with detailed analyses, *FCA* gene was not correctly identified nor annotated in the genome (Jaillon *et al.* 2007). Here we combine EST data, transcript analysis, and sequence homology to improve the annotation of grapevine *FCA*.

Correspondence to: Dr. Ke Jiang, Department of Biological Sciences, Auburn University, 101 Life Sciences Building, Auburn, AL 36849, USA. Fax: 1-334-844-1645. E-mail: kjiang@cshl.edu

Genomic regions of *V. vinifera FCA* was identified by tBLASTn search against grapevine genome using *Arabidopsis FCA* protein as a query. *V. vinifera FCA* EST sequences were identified by BLASTn search against grapevine EST database using *FCA* genomic sequence as a query (Altschul *et al.* 1997, Jaillon *et al.* 2007). The *V. vinifera* 'Pinot Meunier' was kindly provided by Peter S. Cousins of USDA Grape Genetics Research. Total RNAs were extracted from fresh leaves. cDNA synthesis was conducted in two steps: 1. Denaturation of total RNAs at 65 °C for 5 min with FCA X13R primer; 2. cDNA synthesis with DTT, RNAase inhibitor and AMV Reverse-Transcirptase at 50 °C for 30 min followed by 85 °C for 5 min. The *FCA* cDNA was used as templates for following PCR with primer FCA X1F and FCA X13R (denaturation at 72 °C for 3 min, annealing at 54 °C for 0.5 min, elongation at 72 °C for 1.5 min, 35 cycles). The product was cloned using Topo-TA cloning kit and harvested by PCR using plasmids as templates. The final products were cleaned and submitted for sequencing by AU-GSL (GenBank accession number GU300763).

12 EST sequences of *FCA* were identified in *V. vinifera* by BLAST search. When mapped to genomic regions, 10 ESTs are composed of first three exons, with various portions of first exon and third intron. Most of these ESTs have poly-A tails within intron 3, suggesting that they come from *FCA* transcripts β. Two remaining ESTs (at 3' end) contain the last five exons and only the last exon, respectively. They apparently come from *FCA* transcript α or γ if alternative transcription is conserved between *Vitis* and *Arabidopsis* (Table). A new transcript was recovered from cDNA library of 'Pinot Meunier', a very close related variety of 'Pinot Noir'. The transcript spans from exon 1 to exon 9, with canonical splicing sites (GT/AG) in every intron. Intron 3 was correctly spliced out, indicating that this transcript does not belong to the α or β form. If alternative transcription of *FCA* is highly conserved between *Vitis* and *Arabidopsis*, this transcript can be γ or δ. The relative abundance of the two forms in *Arabidopsis* young seedlings are ~35 % and ~10 %, respectively. Therefore, the newly identified transcript in *Vitis* leaves is very likely transcript γ.

An improved annotation of *FCA* was suggested based on ESTs and the new transcript (Figure). The 3'-neighboring gene of *FCA* of current annotation was included in the newly annotated *FCA* (Jaillon *et al.* 2007). This was supported by the presence of two 3' ESTs, and protein similarity between grapevine and *Arabidopsis* at this end. In addition, canonical intron splicing sites for the newly introduced intron are present after connecting the last exon of *FCA* and the first exon of the downstream gene.

An improved annotation of *Vitis vinifera* 'Pinor Noir' *FCA* gene was proposed based on EST data and sequence homology. The incorrect annotation split FCA exons into separate genes due to ambiguous intron position and splicing sites identification. This phenomenon is fairly common in annotation of *V. vinifera* genome as many genes contain extraordinary long introns. About 10 % of all grapevine genes contain introns larger than 3 kb, supported by ESTs or cDNAs and canonical splicing sites. In *V. vinifera*, there

T a b l e

V. vinifera FCA ESTs and cDNA

GenBank Accession	Type	Range Start	Range End	Transcript Form	*Vitis vinifera* variety	Tissue
CB009457	EST	exon 1	intron 3	β	Chardonnay	berry
EC928409	EST	exon 1	intron 3	β	Cabernet Sauvignon	seed
EC980349	EST	exon 1	intron 3	β	Muscat Hamburg	berry
EC978833	EST	exon 1	intron 3	β	Muscat Hamburg	pericarp
EC952213	EST	exon 1	intron 3	β	Cabernet Sauvignon	mixed
CD013791	EST	exon 1	intron 3	β	Chardonnay	berry
CD013721	EST	exon 1	intron 3	β	Chardonnay	berry
EC953975	EST	exon 1	intron 3	β	Cabernet Sauvignon	mixed
EC972841	EST	exon 1	intron 3	β	Muscat Hamburg	pericarp
EE091476	EST	exon 1	intron 3	β	Thompson-seedless	berry
CB975855	EST	exon 14	exon 18	α or γ	Cabernet Sauvignon	berry
CB976362	EST	exon 18	exon 18	α or γ	Cabernet Sauvignon	berry
GU300763	cDNA	exon 1	exon 9	γ or δ	Pinot Meunier	mixed

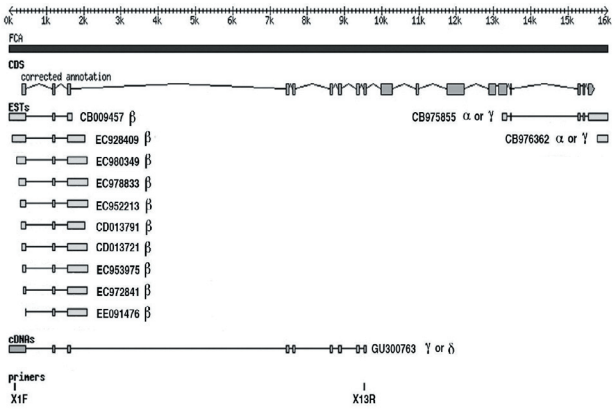

Figure: Diagram of improved annotation of *Vitis vinifera FCA* gene, dashed bar: genomic coordinates (kbp); dark bar: genomic sequences; shade boxes: exons; lines: introns; filled boxes: primers. ESTs and cDNAs annotated with GenBank accession numbers and putative transcript form (β, γ or δ).

are 2800 introns in the size range of 3 bkb to 100 kb, compared to only 10 and 50 such introns in *Arabidopsis* and *Populus*, respectively (JIANG and GOERTZEN, unpubl.). The presence of large introns presents a major difficulty to automated genome annotation (WANG *et al.* 2003). Here, we suggest an annotation combing evidence from genomic sequences, homology and expression data. As more EST and cDNA data are available in both model and non-model systems, large-scale verification of gene annotation by comparing with expression data is both practical and necessary to improve annotation quality (COYNE *et al.* 2008).

The similarities between alternative transcription of *FCA* in *Vitis* and *Arabidopsis* suggests that conserved *FCA* autoregulation mechanism in grapevine is possible. This is perhaps not surprising since the functional conservation can also be found between *Arabidopsis* and rice, indicting that *FCA* functional components could be present in the common ancestor of monocots and dicots (LEE *et al.* 2005). The homology of RRMs can also be traced back to a *Plasmodium* protein, indicating an extremely long history of *FCA* and its RNA-recognizing abilities.

This work was supported by Alabama Experimental Program to Stimulate Competitive Research [RII (05-08): NSF EPS-04476752 to K.J.

ALTSCHUL, S. F.; MADDEN, T. L.; SCHAFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J.; 1997: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl. Acids Res. **25**, 3389-3402.

COYNE, R. S.; THIAGARAJAN, M.; JONES, K. M.; WORTMAN, J. R.; TALLON, L. J.; HAAS, B. J.; CASSIDY-HANLEY, D. M.; WILEY, E. A.; SMITH, J. J.; COLLINS, K.; LEE, S. R.; COUVILLION, M. T.; LIU, Y.; GARG, J.; PEARLMAN, R. E.; HAMILTON, E. P.; ORIAS, E.; EISEN, J. A.; METHÉ, B. A.; 2008: Refined annotation and assembly of the *Tetrahymena thermophila* genome sequence through EST analysis, comparative genomic hybridization, and targeted gap closure. BMC Genomics **9**, 562.

JAILLON, O.; AURY, J. M.; NOEL, B.; POLICRITI, A.; CLEPET, C.; CASAGRANDE, A.; CHOISNE, N.; AUBOURG, S.; VITULO, N.; JUBIN, C.; VEZZI, A.; LEGEAI, F.; HUGUENEY, P.; DASILVA, C.; HORNER, D.; MICA, E.; JUBLOT, D.; POULAIN, J.; BRUYÈRE, C.; BILLAULT, A.; SEGURENS, B.; GOUYVENOUX, M.; UGARTE, E.; CATTONARO, F.; ANTHOUARD, V.; VICO, V.; DEL FABBRO, C.; ALAUX, M.; DI GASPERO, G.; DUMAS, V.; FELICE, N.; PAILLARD, S.; JUMAN, I.; MOROLDO, M.; SCALABRIN, S.; CANAGUIER, A.; LE CLAINCHE, I.; MALACRIDA, G.; DURAND, E.; PESOLE, G.; LAUCOU, V.; CHATELET, P.; MERDINOGLU, D.; DELLEDONNE, M.; PEZZOTTI, M.; LECHARNY, A.; SCARPELLI, C.; ARTIGUENAVE, F.; PÈ, M. E.; VALLE, G.; MORGANTE, M.; CABOCHE, M.; ADAM-BLONDON, A. F.; WEISSENBACH, J.; QUÉTIER, F.; WINCKER, P.; 2007: The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature **449**, 463-467.

LEE, J. H.; CHO, Y. S.; YOON, H. S.; SUH, M. C.; MOON, J.; LEE, I.; WEIGEL, D.; YUN, C. H.; KIM, J. K.; 2005: Conservation and divergence of *FCA* function between Arabidopsis and rice. Plant Mol Biol. **58**, 823-838.

MACKNIGHT, R.; BANCROFT, I.; PAGE, T.; LISTER, C.; SCHMIDT, R.; LOVE, K.; WESTPHAL, L.; MURPHY, G.; SHERSON, S.; COBBETT, C.; DEAN, C.; 1997: *FCA*, a gene controlling flowering time in Arabidopsis, encodes a protein containing RNA-binding domains. Cell **89**, 737-745.

MACKNIGHT, R.; DUROUX, M.; LAURIE, R.; DIJKWEL, P.; SIMPSON, G.; DEAN, C.; 2002: Functional significance of the alternative transcript processing of the Arabidopsis floral promoter *FCA*. Plant Cell **14**, 877-888.

PUTTERILL, J.; LAURIE, R.; MACKNIGHT, R.; 2004: It's time to flower: the genetic control of flowering time. Bioessays **26**, 363-373.

QUESADA, V.; MACKNIGHT, R.; DEAN, C.; SIMPSON, G. G.; 2003: Autoregulation of FCA pre-mRNA processing controls Arabidopsis flowering time. EMBO J. **22**, 3142-3152.

WANG, J.; LI, S.; ZHANG, Y.; ZHENG, H.; XU, Z.; YE, J.; YU, J.; WONG, G. K.; 2003: Vertebrate gene predictions and the problem of large genes. Nat. Rev. Genet. **4**, 741-749.

Received May 10, 2010