Characterization and expression of U1snRNA genes from potato

Petra Vaux, François Guerineau, Robbie Waugh and John W.S. Brown*
Scottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, Scotland (*author for correspondence)

Received 7 January 1992; accepted in revised form 10 April 1992

Key words: gene expression, gene variants, pre-mRNA splicing, pseudogenes, U1 small nuclear RNA

Abstract

U1 small nuclear RNAs (U1snRNAs) occur in the nucleus of plants and animals where, complexed with several proteins in the form of U1 small nuclear ribonucleoprotein particles (U1snRNPs), they play an important role in precursor messenger RNA (pre-mRNA) splicing. Ten potato U1snRNA genes have been isolated on two genomic clones illustrating the clustering of this multigene family on the potato genome. Based on both the sequence of their coding regions and upstream regulatory elements, seven of the genes are potentially functional. The other three genes were pseudogenes with defective promoter or coding region sequences. Analysis of expression of individual cloned U1snRNA genes in transfected tobacco protoplasts was impossible due to the similarity of U1snRNA sequences in tobacco. However, by marking the coding regions with oligonucleotides or constructing chimaeric genes consisting of a potato U1snRNA promoter region and maize U5snRNA coding region, three of the U1 promoter regions were shown to be transcriptionally active.

Introduction

The majority of eukaryotic pre-mRNAs are composed of protein-coding sequences (exons), interrupted by one or more non-protein coding intervening sequences (introns). Intron removal by pre-mRNA splicing involves cleavage and ligation reactions that result in the precise excision of introns and joining of adjacent exons to generate the open reading frame of the mature mRNA. Splicing occurs in a large ribonucleoprotein complex called the spliceosome, which forms on the pre-mRNA. The major components of the spliceosome are UsnRNPs consisting of UsnRNA molecules complexed with a number of proteins.

The spliceosomal UsnRNPs (U1, U2, U4/U6 and U5) associate with other protein factors in an ordered pathway on the pre-mRNA to form the spliceosome [reviewed by 45, 26, 40]. Spliceosomal UsnRNAs have been characterized from animals, plants and smaller eukaryotes such as yeast, slime moulds and dinoflagellates [reviewed by 18, 17]. In animal nuclei, U1snRNAs are the most abundant of the spliceosomal UsnRNAs. In the majority of cases they are encoded by a multigene family, in which the individual members are often found clustered in the genome [8]. U1snRNAs associate with the core proteins common to all spliceosomal snRNPs and with three U1snRNP-specific

The nucleotide sequence data reported will appear in the EMBL, GenBank and DDBJ Nucleotide Sequence Databases under the accession numbers Z11881, Z11882 and Z11883.
proteins, U1-A, U1-C and U1-70K, to form U1snRNP [reviewed by 26]. The major role of U1snRNP lies in the recognition of intron 5' splice sites which involves base-pairing of the 5' single-stranded region of U1snRNA to the 5' splice site of the pre-mRNA [60]. This interaction appears to be augmented, either directly or indirectly, by the U1-C protein [22] and other splicing factors such as SF2/ASF [14, 25] and/or SPP1 and SPP2 [46].

Pre-mRNA splicing is essential for gene expression and regulation of expression at the level of splicing is well-documented in animal systems where it often involves alternative or differential splicing [reviewed by 44]. In plants relatively few cases of alternative splicing have been described [12, 30, 29, 56, 35, 2, 15]. Although in most cases the function of the resulting protein products is unknown, alternative splicing of the pre-mRNA of the phospholipid transfer protein from maize produces two proteins with different C-terminal regions which may have a role in protein targeting [2]. In the case of Digitaria streak virus (DSV), inefficient splicing of the DSV intron is required for viral replication and infectivity [35]. In order to understand how such gene systems are regulated at the level of splicing more information of plant splicing components is essential.

Analyses of UsnRNA genes from monocotyledonous and dicotyledonous species have indicated that plant UsnRNA populations are more heterogeneous in their coding regions than their yeast or animal counterparts [49, 10, 1, 54, 55, 21]. Plant UsnRNA gene promoters are also different and consist of two essential promoter elements, the upstream sequence element (USE) and a TATA element [50]. The expression of the relatively rare UsnRNA variants in animal systems is regulated in a developmental or tissue-specific manner [reviewed by 8] and it has been postulated that such variants may control developmental gene expression by regulating alternative splicing pathways [reviewed by 32]. Some evidence for differential expression of UsnRNA variants has been observed for wheat U1snRNAs [10], pea U1, U2, U4 and U5snRNAs [21] and maize U5snRNAs (D.J. Leader, personal communication). In order to provide a basis for studying differential and tissue-specific expression of potato UsnRNA genes we have isolated and characterized ten potato U1snRNA gene variants. In addition, we have utilized two different procedures for studying the expression of individual potato U1snRNA genes.

Materials and methods

Materials

Restriction enzymes were purchased from GIBCO-BRL, Pharmacia, Koch-Light, Boehringer-Mannheim and Northumbria Biological Laboratories. T4 DNA ligase, alkaline phosphatase, RNase A, RNase T1, SP6 RNA polymerase, and RNase inhibitor were obtained from Boehringer-Mannheim. RNase-free DNase was obtained from Promega. Nitrocellulose/nylon transfer membrane and radio-nucleotides were from Amersham. Sequenase was obtained from United States Biochemicals.

Isolation of potato U1snRNA genes

A Sau 3A partial potato genomic library in \( \lambda \)EMBL 4 was screened using standard procedures with the 1.6 kb Eco RV/Xba I fragment of phvU1, containing a Phaseolus vulgaris U1snRNA gene [51]. Two U1snRNA-containing clones, lambda Pot U1-6 and lambda Pot U1-21, were characterized (Fig. 1). The 4.5 kb Eco RI fragment from lambda Pot U1-6 and the 6.4 kb and 8.5 kb Eco RI fragments from lambda Pot U1-21 were cloned into pUC13 to produce the plasmids pPotU1-6E4.5, pPotU1-21E6.4 and pPotU1-21E8.5, respectively. The coding regions and extensive 5' and 3' flanking regions of the U1snRNA genes contained on these plasmids were obtained by the chain termination method with Sequenase using universal forward and reverse strands primers, as well as U1-specific primers based on the U1snRNA gene sequences of tomato, in positions 1 to 25, 112 to 90 and 141 to 122 [1].