The light-harvesting antenna of higher plants contains at least 10 different LHC polypeptides (Jansson, 1994). The four LHC polypeptides specifically associated with PSI have molecular masses between 20 and 25 kD and are encoded by the nuclear Lhca1, Lhca2, Lhca3, and Lhca4 genes. We have previously characterized Lhca1, Lhca2, and Lhca3 cDNAs from the gymnosperm Scots pine (Pinus sylvestris) (Jansson and Gustafsson, 1991), and report here on the isolation of two Lhca4 cDNA clones from the same organism, using the homologous tomato cDNA (Schwartz et al., 1991) as the probe. Lhca4 sequences have so far been characterized only in the three angiosperm species tomato (Lycopersicon esculentum) (Schwartz et al., 1991), Arabidopsis thaliana (Zhang et al., 1991), and barley (Hordeum vulgare) (Anandan et al., 1993).

The longest cDNA clone encodes a Lhca4 precursor protein of 251 amino acids. Hybridization to Southern blots of Scots pine genomic DNA using the cDNA as probe showed that there is only one Lhca4 gene in the Scots pine genome, named Lhca4.1 (Table I), according to Jansson et al. (1992). A near full-length cDNA clone corresponding to another allele of the Lhca4.1 gene was also isolated. This allele differed by three nucleotides within the coding region and three in the 3' end. One substitution resulted in a change of the protein sequence, an Ala-to-Gly substitution in the transit peptide, but the others were silent. We frequently isolate cDNA clones originating from different alleles from our libraries, which are derived from a commercial seed lot with large genetic diversity.

The N terminus of the mature protein (after removal of the transit peptide) has been identified as two Lys residues in spinach, (Spinacia oleracea), pea (Pisum sativum) (Ikeuchi et al., 1991), and tomato (Schwartz et al., 1992). The Scots pine preprotein has only one Lys residue at the corresponding position, which most likely is the N terminus in the mature Scots pine protein. Yet it is possible that the cleavage occurs before the preceding Ala residue, which would yield a mature Scots pine protein starting with Ala-Lys, as in barley (Knoetzel et al., 1992). Provided that the former assumption is correct, the Scots pine Lhca4 protein contains 200 amino acids.

**Acknowledgment**

We want to thank Gail Huchins for technical assistance.

Received May 19, 1994; accepted June 23, 1994.

Copyright Clearance Center: 0032-0889/94/106/1693/02.

The EMBL accession numbers for the sequences reported in this article are Z17226 and Z16409.

**Literature Cited**

Schwartz E, Shen D, Aebersold R, McGrath JM, Pichersky E, Green BR (1991) Nucleotide sequence and chromosomal location of Cab11 and Cab12, the genes for the fourth polypeptide of the photosystem I light-harvesting antenna (LHCI). FEBS Lett 280: 229–234