Curr Genet (1991) 19:417–422

Current Genetics

A small gene family in barley encodes ribosomal proteins homologous to yeast YL17 and L22 from archaebacteria, eubacteria, and chloroplasts

Lene H. Madsen, Jette D. Kreiberg *, and Kirsten Gausing

Department of Molecular Biology and Plant Physiology, University of Aarhus, DK-8000 C Aarhus C, Denmark

Received January 14, 1991

Summary. The amino acid sequences of two barley ribosomal proteins, termed HvL17-1 and HvL17-2, were decoded from green leaf cDNA clones. The N-terminal sequences of the derived barley proteins are 48% identical to the N-terminal amino acid sequence of protein YL17 from the large subunit of yeast cytoplasmic ribosomes. Via archaebacterial ribosomal proteins this homology extends to ribosomal protein L22 from eubacteria and chloroplast. Barley L17, and ribosomal proteins L22 and L23 from the archaebacteria *Halobacterium halobium* and *H. marismortui*, are 25–33% identical. Interestingly, the barley and archaebacterial proteins share a long, central stretch of amino acids, which is absent in the corresponding proteins from eubacteria and chloroplasts.

Barley L17 proteins are encoded by a small gene family with probably only two members, represented by the cDNA clones encoding HvL17-1 and HvL17-2. Both these genes are active in green leaf cells. The expression of the L17 genes in different parts of 7-day old barley seedlings was analyzed by semiquantitative hybridization. The level of L17 mRNA is high in meristematic and young cells found in the leaf base and root tip. In the leaf, the L17 mRNA level rapidly decreases with increasing cell age, and in older root cells this mRNA is undetectable.

Key words: *Hordeum vulgare* – Plant ribosomal protein – cDNA sequences – Developmental gene expression

Introduction

Protein synthesis is a fundamental process in cellular housekeeping metabolism, and the protein synthetic capacity of a cell obviously depends on the supply of ribosomes. In bacteria, ribosome content is proportional to growth rate. Similarly, the synthesis of rRNA and ribosomal proteins (r-proteins) in eukaryotes is extensively modulated during growth and differentiation. In this way the rate of ribosome formation is adjusted to meet the cellular need for protein synthesis. The synthesis of ribosomal components is coordinately regulated and a variety of regulatory mechanisms, operating at different levels of gene expression, is used to control r-protein accumulation. The precise regulatory strategy depends on the organism or cell type in question (for reviews see Gausing 1982; Mager 1988; Warner 1989).

In plants, protein synthesis takes place in three cellular compartments: the cytoplasm, the chloroplasts, and the mitochondria. Each compartment has a unique set of ribosomes. Cytoplasmic ribosomes are typically eukaryotic, whereas chloroplast ribosomes are similar to prokaryotic ribosomes (Boynton 1980). Moreover, the primary sequences of individual chloroplast r-proteins are clearly homologous to *E. coli* r-proteins (Wittmann-Liebold 1986). Mitochondrial ribosomes are structurally diverse and contribute only little to the total cellular ribosome population (Boynton 1980). Chloroplast ribosomes, on the other hand, can account for more than half the total ribosome content in sections of wheat and barley leaves (Dean and Leech 1982; Barkardottir et al. 1987).

Primary sequences of plant cytoplasmic r-proteins have been deduced from cDNA clones and so far include the sequences of S14 from maize (Larkin et al. 1989) and S11 from *Arabidopsis* and soybean (Gantt and Thompson 1990). The expression of cytoplasmic r-protein genes was studied in soybean and maize and elevated levels of r-protein mRNAs were related to cell division in both systems. In soybean the relative levels of individual r-protein messengers coordinately increased up to 8-fold in rapidly proliferating cells from hypocotyls treated with 1,2-dichlorophenoxyacetic acid, a synthetic auxin (Gantt and Key 1983, 1985). In developing maize endosperm, the r-protein S14 mRNA level is high and correlates with the mitotic index (Larkin et al. 1989).

We have previously described two gene families with members preferentially expressed in young basal cells of

* *Present address:* Danish Distillers, Danisco A/S, The Science Park, Gustav Wieds Vej 10, DK-8000 Aarhus C, Denmark

Offprint requests to: K. Gausing

barley leaves (Barkardottir et al. 1987; Gausing and Barkardottir 1986). One is the subclass of ubiquitin genes encoding ubiquitin r-protein fusion peptides (Gausing and Barkardottir 1986; Gausing and Jensen 1990). We now find that the second gene family also encodes r-proteins; these are homologous to r-protein YL17 from yeast (Otaka et al. 1984). Via archaebacterial r-proteins the homology extends to r-protein L22 of eubacteria and chloroplasts, suggesting equivalent roles for these proteins in ribosome assembly and function.

Materials and methods

Plant material. Barley seedlings (*Hordeum vulgare* cv. Bomi) were grown and harvested as previously described (Barkardottir et al. 1987).

Isolation and analysis of cDNA clones. The construction (in the *Pst*I site of pBR327, Covarrubias et al. 1981) and analysis of a barley green leaf cDNA library containing 4800 clones has been described (Gausing and Barkardottir 1986; Gausing 1987). A second cDNA library with 6.5×10^6 recombinants was constructed in lambda ZAP1 from poly $(A)^+$ RNA purified from the basal 2.5 cm of 7-day old barley seedlings (Stratagene Custom Services). The 680 bp *Pst*I insert of the HvL17-1 cDNA clone pKG1973 (Barkardottir et al. 1987) was used to screen the lambda ZAP1 cDNA library. Purification of recombinant phages and rescue of pBluescript SK(–) plasmids was performed according to the Stratagene lambda ZAP1 manual. cDNA inserts and overlapping restriction fragments covering the inserts were transferred into pBS- (Short et al. 1988). Sequencing was performed with double-stranded DNA using a Sequenase® kit according to suppliers instructions (United States Biochemicals, Cleveland, Ohio).

Isolation and analysis of RNA and DNA. Nucleic acids were isolated and analyzed by Northern and Southern hybridization as described by Gausing and Barkardottir (1986).

Amino acid sequence comparisons. The homology search was performed using the FASTA program (Pearson and Lipman 1988) to search the mipsx database.

Results and discussion

Isolation and characterization of L17 cDNA clones

Because of the similar expression pattern of ubiquitin-ribosomal protein fusion peptide genes and the L17 genes represented by the pKG1973 group of cDNA clones (Barkardottir et al. 1987) we were encouraged to investigate the possible relationship between the products of these genes.

As none of the original four L17 cDNA clones isolated from our cDNA library of 4800 clones were full length, the *Pst*I insert of clone pKG1973 (680 bp) was used to screen an independent cDNA library in which 44 positive clones were identified among 100000 recombinants. Selected clones which were purified and further analyzed fell in two classes, representing the transcripts of two related genes. The structure of the two classes of cDNA clones, as well as the originally isolated clones pKG1973 and pKG989, is shown in Fig. 1. The nucleotide and derived amino acid sequences are shown in Fig. 2. In

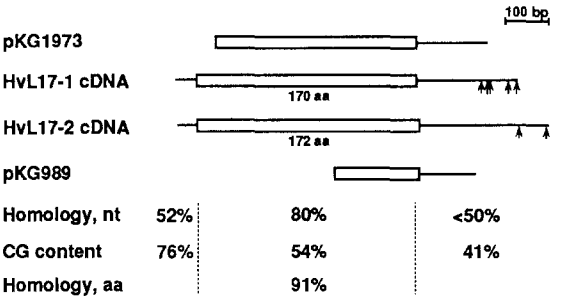

Fig. 1. Structural map of barley HvL17-1 and HvL17-2 cDNA clones. *Thin line*, non-coding regions; *open box*, protein coding regions; *arrowheads*, polyadenylation sites. The position, relative to the full length clones, of the HvL17-1 cDNA clone pKG1973 and the HvL17-2 cDNA clone pKG989 is indicated

agreement with earlier estimates of messenger size (850 nt, Barkardottir et al. 1987), the cDNA inserts of the longest clones are between 750 and 850 bp in length without the poly (A) tails. The different length of the 3′-noncoding region in the two types of cDNAs, together with the occurrence of multiple, widely spaced, polyadenylation sites, account for this variation in size (Fig. 1). The maps and nt sequences showing the polyadenylation sites (Figs. 1 and 2) represent data from 12 HvL17-1 and seven HvL17-2 clones.

The nucleotide sequence homology between the two types of cDNA clones is 80% in the coding region, whereas the 5′- and 3′-noncoding regions differ considerably, with identities of 52 and less than 50%, respectively (Fig. 1).

The distribution of C+G nucleotides is identical in HvL17-1 and HvL17-2 cDNA clones; a changed C+G content marks the boundaries of the coding region (Fig. 1). In cereals, the C+G content in coding regions is normally higher than the average genomic C+G content (Salinas et al. 1988), which is 42% in barley (Shapiro 1970). The 54% and 41% C+G content of the coding and 3′-noncoding regions of the cDNA clones is, therefore, not unexpected. The 76% C+G nucleotides in the 5′-leaders is exceptional, however, since 5′-leader sequences in plants are normally rich in A+T (Joshi 1987). It may be significant that the leader sequences of two barley ubiquitin-ribosomal protein mRNAs have a comparably high C+G content (Gausing and Jensen 1990).

The single, open reading frames following the first ATG of the HvL17-1 and HvL17-2 cDNA clones are 510 and 516 bp long, respectively. The encoded proteins, HvL17-1 and HvL17-2, are 170 and 172 amino acids long (Figs. 1 and 2). The amino acid sequences are 91% identical. Calculated molecular masses are 19.5 and 19.7 kDa, respectively, and the proteins contain 21% basic (K+R) and 9% acidic (D+E) amino acids.

Identification of homologous proteins from other organisms

By computer-assisted sequence comparisons, the barley L17 proteins were found to be homologous to well-characterized r-proteins from other organisms. Figure 3

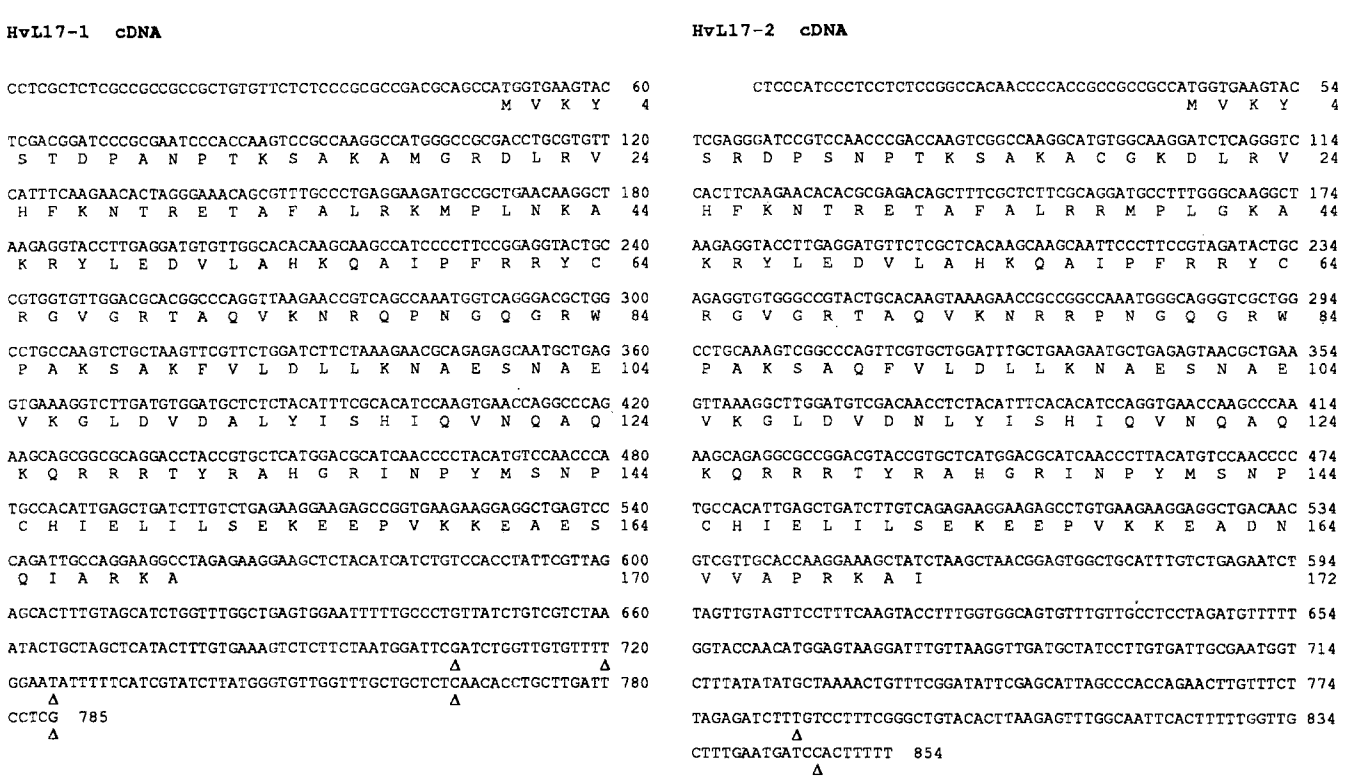

Fig. 2. Nucleotide and derived amino acid sequences of HvL17-1 and HvL17-2 cDNA clones. *Arrowheads*, polyadenylation sites. The longest HvL17-2 clone has no poly(A)-tail

Table 1. Percentages of amino acid identity between L17/L22 proteins

	H.h. L22	H.m. L23	YL17[a]	HvL17-1	HvL17-2	Lw, cpL22	Eug, cpL22	Tob, cpL22	Rice cpL22	Myc L22	E.c. L22
H.h. L22	·	67	22	27	25	17	15	13	15	19	17
H.m. L23	67	·	24	33	30	17	15	13	15	17	17
YL17	23	26	·	48	48	15	20	15	3	20	18
HvL17-1	29	36	48	·	90	19	14	13	12	18	17
HvL17-2	27	34	48	91	·	20	15	14	13	17	16
Lw, cpL22	23	23	17	28	30	·	42	41	40	37	36
Eug, cpL22	22	22	24	21	22	44	·	31	28	34	36
Tob, cpL22	21	21	18	19	20	55	42	·	34	26	26
Rice, cpL22	21	21	5	18	19	50	38	52	·	24	23
Myc L22	28	25	29	28	27	40	37	36	32	·	53
E.c. L22	24	24	25	26	26	39	39	37	31	53	·

The calculations were based on the alignment shown in Fig. 3. Upper right: percentage of identity including terminal residues and non-overlapping regions. Lower left: percentage of identity omitting terminal residues and non-overlapping regions. Matches of initiator methionine were not scored. For abbreviations of species names and sequences references see legend to Fig. 3

[a] The percentage of homology to yeast YL17 is calculated on the basis of the 40 known N-terminal amino acids

shows an alignment of barley L17 with the equivalent r-proteins from yeast, two archaebacteria, two eubacteria and chloroplasts from four distantly related taxa: a bryophyte, an alga, a dicotyledonous and a monocotyledonous angiosperm. Table 1 shows the percentage of identity between the proteins.

The N-terminal sequences of the barley proteins are 48% identical to the 40 amino acid N-terminal sequence known from the 60 S r-protein YL17 from yeast (Otaka and Kobata 1978; nomenclature according to Otaka and Osawa 1981; Otaka et al. 1984). This strongly suggests that the barley proteins are 60 S r-proteins equivalent to YL17. The molecular weight of YL17, 20 kDa, also compares well with the values calculated for the barley L17 proteins. The database search did not identify additional eukaryotic r-proteins with significant homology to barley L17 or to the N-terminal of yeast YL17. Barley L17 and r-proteins L22 and L23 from the archaebacteria *Halobacterium halobium* and *H. marismortui* are 25–33% identical (Table 1; Mankin 1989; Hatakeyama et al. 1988). These archaebacterial equivalents of eubacterial and chloroplast r-protein L22 contain a central 30 amino acid region not found in the eubacterial protein (Mankin 1989; Hatakeyama et al. 1988). Interestingly, an identically located insertion is found in barley L17, and the similarity between barley L17 and archaebacterial L22/

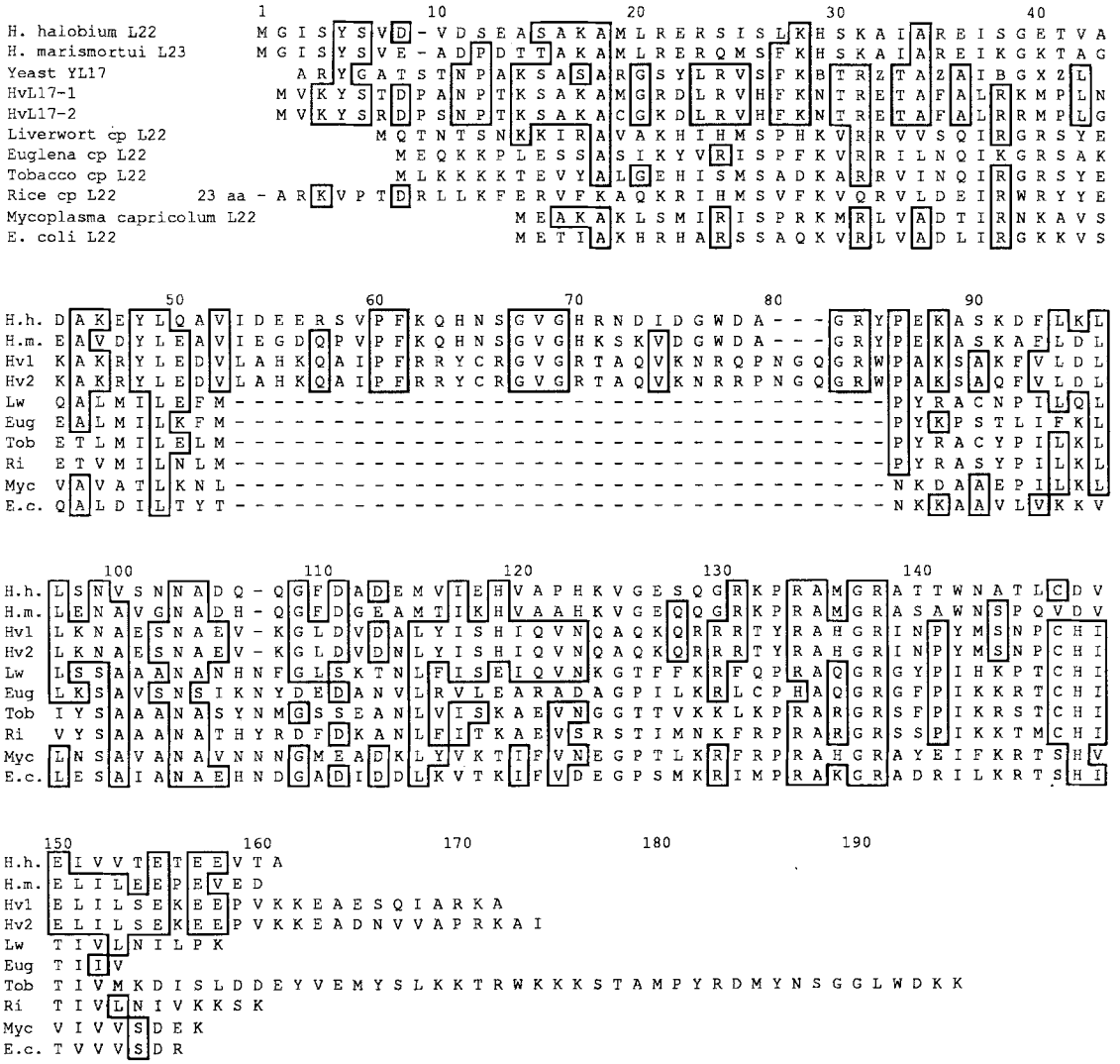

Fig. 3. Alignment of barley L17 proteins with homologous ribosomal proteins from other organisms. Amino acids identical to both barley sequences are *boxed*. Only gaps necessary to maintain alignment are introduced. Numbering is arbitrary, starting with the first methionine of the archaebacterial sequences. An N-terminal extension of 23 amino acids of rice chloroplast L22 (Ri) (Hiratsuka et al. 1989) is omitted for clarity. The alignment of *H. Marismortui* (H.m.) L22 with *E. coli* L22 is taken from Hatakeyama et al. 1988. Additional sequence information: *Euglena gracilis* (Eug) (Christopher et al. 1988), yeast (Otaka et al. 1984), *E. coli* (E.c.) (Wittmann-Liebold and Greuer 1980), *H. halobium* (H.h.) (Mankin 1989), *Mycoplasma capricolum* (*Myc*) (Ohkubo et al. 1987), liverwort *Marchantia polymorpha* chloroplast (Lw) (Ohyama et al. 1986), Tobacco chloroplast (Tob) (Tanaka et al.) *Hordeum vulgare* (Hv) (this work). cp = chloroplast. In the yeast sequence B = Asx, Z = Glx, X = unknown amino acid

L23 clearly persists in this region (Fig. 3). Furthermore, the barley and archaebacterial proteins share a single one amino acid gap (position 107, Fig. 3) relative to the eubacterial and chloroplast r-proteins.

With the exception of proteins of similar origin (e.g., archaebacterium/archaebacterium), the homologies between the proteins in region of overlap are all in the range of 20–35% (lower left, Table 1). However, the occurrence of insertions and deletions clearly discriminates between eukaryotic and archaebacterial proteins on the one hand and eubacterial/chloroplast proteins on the other. A similar conservation of insertions/deletions between eukaryotes and archaebacteria is seen by examination of other r-protein sequences (Hatakeyama et al. 1989; Köpke and Wittmann-Liebold 1988). As expected, barley chloroplast *rpl22* encodes a protein which lacks the central insertion and, therefore, clearly belongs to the eubacterial/chloroplast group of L17/L22 proteins (unpublished results).

In conclusion, barley L17 proteins are highly related to YL17 and significantly more related to archaebacterial than to eubacterial counterparts. This is in agreement with the general observation that eukaryotic and archaebacterial r-proteins are often more related to each other than to eubacterial r-proteins (Kimura et al. 1989; Köpke and Wittmann-Liebold 1989; Auer et al. 1989; Strobel et al. 1989). The conservation through evolution of the amino acid sequences of the L17/L22 proteins, which is evident from the comparison presented in Fig. 3 and Table 1, suggests a conserved role for these proteins in ribosome assembly and function. In *E. coli*, r-protein L22 is involved in the early steps of ribosome assembly, and is closely associated with the 5′ region of 23S rRNA (Pichon et al. 1975; Homann and Nierhaus 1971; Chen-

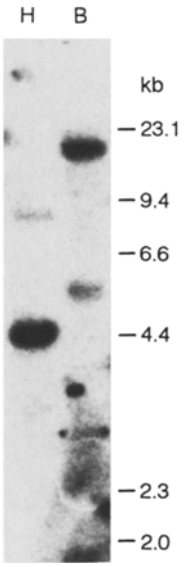

Fig. 4. Southern blot analysis of restriction enzyme digests of barley DNA (10 μg per lane). The probe used was the *Pst*I insert of pKG1973 (Fig. 1) *H*, *Hin*dIII and *B*, *Bam*HI digests. MW marker was lambda DNA digested with *Hin*dIII

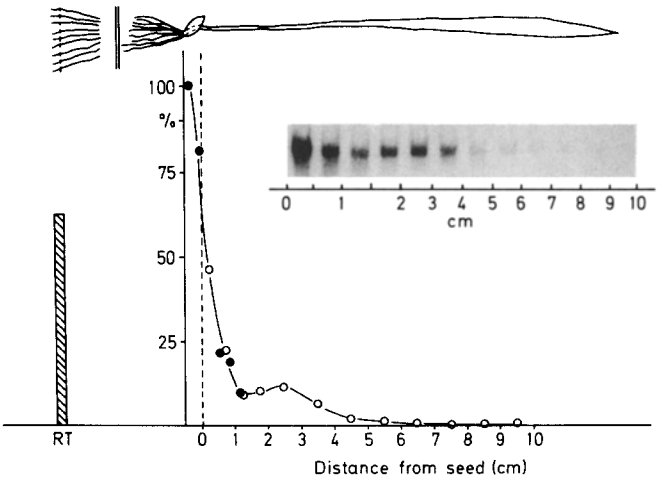

Fig. 5. Northern blot analysis of total RNA from sections of leaves and roots from 7-day old barley seedlings. The probe was the *Pst*I insert of pKG1973 (Fig. 1). The L17 mRNA levels was determined in 3 mm sections of leaves cut at their bases within the seeds (—●—), 0.5 and 1.0 cm section of leaves cut at seed level (—○—), 3 mm root tips (*RT bar*) and in the rest of the root (undetectable). The third 3 mm leaf section was lost during preparation. The highest L17 mRNA level detected (first basal 3 mm leaf section) was arbitrarily set at 100%. Differently exposed autoradiographs were scanned and the results were normalized to fit the same curve. Inset: an autoradiograph from Northern hybridization to total RNA from leaf sections (10 μg per lane)

Schmeisser and Garrett 1976). Spinach chloroplast L22, which carries a long C-terminal extension compared to other L17/L22 proteins, may be exceptional. This protein binds to 5S rRNA (Zhou et al. 1989).

Genomic organization of L17 genes

The genomic organization of r-protein genes in eukaryotes varies considerably, ranging from large gene families with only one active member in mouse (Dudov and Perry 1984) to single copy genes in *Drosophila* (Burns et al. 1984; Rafti et al. 1988). The highest number of active genes reported for a single r-protein is three (Larkin et al. 1989; Gatermann et al. 1989). In the yeast *Saccharomyces cerevisiae* many, but not all, r-protein genes are duplicate and sometimes the two copies have promoters of different strengths (Mager 1988; Warner 1989). In plants, copy numbers from one to six have been suggested (Gantt and Key 1985; Larkin et al. 1989).

To estimate the number of L17 genes in the barley genome, isolated DNA was digested with restriction enzymes and probed with the *Pst*I insert of the HvL17-1 cDNA clone pKG1973 (Fig. 1). Two hybridizing bands of different intensity were detected with each of the restriction enzymes *Hin*dIII and *Bam*HI (Fig. 4). With a HvL17-2-specific cDNA probe (3′-noncoding region), only bands corresponding to the weakly hybridizing *Hin*dIII and *Bam*HI bands were detected (data not shown). From these results it is concluded that the weakly hybridizing bands (Fig. 4) represent the HvL17-2 gene, while more strongly hybridizing bands represent the HvL17-1 gene. It is, therefore, likely that r-protein L17 is encoded by only these two genes in barley.

Developmental regulation of L17 mRNA accumulation

The expression of the barley L17 genes was analyzed in different parts of 7-day old barley plants grown under normal, diurnal conditions (Barkardottir et al. 1987). In grass leaves the cells above the basal, meristematic zones are progressively older with increasing distance from the leaf base. In 7-day old barley plants cell age is nearly proportional to distance from the leaf base up to about 10 cm where cells are 3.5 days old (Barkardottir et al. 1987). The results of the expression studies (Fig. 5) show that the L17 genes are predominantly expressed in regions of meristematic and young, rapidly growing, cells in both leaf and root. In leaf cells, the L17 mRNA level quickly drops to low, but detectable, levels in leaf sections above the seed, whereas L17 mRNA could not be detected in roots with the tip meristem removed. The developmental modulation of L17 mRNA accumulation in leaf cells is highly similar to the pattern observed for r-protein-encoding ubiquitin-tail mRNAs in barley (Gausing and Barkardottir 1986; Gausing and Jensen 1990), which is in agreement with the concerted regulation of r-protein gene expression observed in other organisms (Falkis and Meyuhas 1982; Gantt and Key 1985; Kim and Warner 1983; Pierandrei-Amaldi et al. 1982).

Acknowledgements. We thank N. Pallisgaard for help with database searches, P. S. Nielsen for many valuable discussions and L. Kjeldbjerg and I. Bjørndal for technical assistance. This work was supported by the Commission of the European Communities (contract GBI-4-117-DK) and The Danish Biotechnology Programme. L.H.M. was supported by a fellowship from the Carlsberg Foundation.

References

Auer J, Spicker G, Böck A (1989) J Mol Biol 209:21–26

Barkardottir RB, Jensen BF, Kreiberg JD, Nielsen PS, Gausing K (1987) Develop Genet 8:495–511

Boynton JE, Gillham NW, Lambowitz AM (1980) Biogenesis of chloroplast and mitochondrial ribosomes. In: Chambliss G, Craven GR, Davies J, Davis K, Kahan L, Nomura M (eds) Ribosomes. Structure, function and genetics. University Park Press, Baltimore, pp 903–950
Burns DK, Stark BC, Macklin MD, Chooi WY (1984) Mol Cell Biol 4:2643–2652
Chen-Schmeisser U, Garrett R (1976) Eur J Biochem 69:401–410
Christopher DA, Cushman JC, Price CA, Hallick RB (1988) Curr Genet 14:275–286
Covarrubias L, Cervantes L, Covarrubias A, Soberón X, Vichido I, Blanco A, Kupersztoch-Portnoy YM, Bolivar F (1981) Gene 13:25–35
Dean C, Leech RM (1982) Plant Physiol 69:904–910
Dudov KP, Perry RP (1984) Cell 37:457–468
Faliks D, Meyuhas O (1982) Nucleic Acids Res 10:789–801
Gantt JS, Key JL (1983) Biochemistry 22:4131–4139
Gantt JS, Key JL (1985) J Biol Chem 260:6175–6181
Gantt JS, Thompson MD (1990) J Biol Chem 265:2763–2767
Gatermann KB, Teletski C, Gross T, Käufer NF (1989) Curr Genet 16:361–367
Gausing K (1982) Trends Biochem Sci 7:65–67
Gausing K (1987) Planta 171:241–246
Gausing K, Barkardottir RB (1986) Eur J Biochem 158:57–62
Gausing K, Jensen CB (1990) Gene 94:165–171
Hatakeyama T, Hatakeyama T, Kimura M (1988) FEBS Letts 240:21–28
Hatakeyama T, Kaufmann F, Schroeter B, Hatakeyama T (1989) Eur J Biochem 185:685–693
Hiratsuka J, Shimada H, Whittier R, Ishibashi T, Sakamoto M, Mori M, Kondo C, Honji Y, Sun C-R, Meng B-Y, Li Y-Q, Kanno A, Nishizawa Y, Hirai A, Shinozaki K, Sugiura M (1989) Mol Gen Genet 217:185–194
Homann HE, Nierhaus KH (1971) Eur J Biochem 20:249–257
Joshi CP (1987) Nucleic Acids Res 15:6643–6653
Kim CH, Warner JR (1983) Mol Cell Biol 3:457–465
Kimura M, Arndt E, Hatakeyama T, Hatakeyama T, Kimura J (1989) Can J Microbiol 35:195–199
Köpke AKE, Wittmann-Liebold B (1988) FEBS Lett 239:313–318
Köpke AKE, Wittmann-Liebold B (1989) Can J Microbiol 35:11–20
Larkin JC, Hunsperger JP, Culley D, Rubenstein I, Silflow CD (1989) Genes Develop 3:500–509
Mager WH (1988) Biochim Biophys Acta 949:1–15
Mankin AS (1989) FEBS Lett 246:13–16
Ohkubo S, Muto A, Kawauchi Y, Yamao F, Osawa S (1987) Mol Gen Genet 210:314–322
Ohyama K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S-i, Inokuchi H, Ozeki H (1986) Plant Mol Biol Reporter 4:148–175
Otaka E, Kobata K (1978) Mol Gen Genet 162:259–268
Otaka E, Osawa S (1981) Mol Gen Genet 181:176–182
Otaka E, Higo K-i, Itoh T (1984) Mol Gen Genet 195:544–546
Pearson WR, Lipman DJ (1988) Proc Natl Acad Sci USA 85:2444–2448
Pichon J, Marvaldi J, Marchis-Mouren G (1975) J Mol Biol 96:125–137
Pierandrei-Amaldi P, Campioni N (1982) Cell 30:163–171
Rafti F, Gargiulo G, Manzi A, Malva C, Grossi G, Andone S, Graziani F (1988) Nucleic Acids Res 16:4915–4926
Salinas J, Matassi G, Montero LM, Bernardi G (1988) Nucleic Acids Res 16:4269–4285
Shapiro HS (1970) Distribution of purines and pyrimidines in deoxyribonucleic acids. In: Sober HA (ed) Handbook of biochemistry. Selected data for molecular biology, 2nd edn. CRC Press Inc., Cleveland, Ohio, pp H90–91
Short JM, Fernandez JM, Sorge JA, Huse WD (1988) Nucleic Acids Res 16:7583–7600
Strobel O, Köpke AKE, Kamp RM, Böck A, Wittmann-Liebold B (1988) J Biol Chem 263:6538–6546
Tanaka M, Wakasugi T, Sugita M, Shinozaki K, Sugiura M (1986) Proc Natl Acad Sci USA 83:6030–6034
Warner JR (1989) Microbiol Rev 53:256–271
Wittmann-Liebold B (1986) Ribosomal proteins: their structure and evolution. In: Hardesty B, Kramer G (eds) Structure, function and genetics of ribosomes. Springer Series in Molecular Biology, pp 326–361
Wittmann-Liebold B, Greuer B (1980) FEBS Lett 121:105–112
Zhou D-X, Quigley F, Massenet O, Mache R (1989) Mol Gen Genet 216:439–445

Communicated by H. Kössel