博士論文番号　0081040

Characterization of Novel Microsporogenesis-associated Genes from *Lilium longiflorum*

Bayarmaa Gun-Aajav
Laboratory of Plant Molecular and Cellular Biology
Graduate School of Biological Sciences
Nara Institute of Science and Technology
(Prof. Takashi Hashimoto)
平成１５年１月７日提出

Contents

Abbreviations

AD	activation domain
BD	binding domain
CTAB	cetyl trimetyl ammonium bromide
DAPI	4', 6-diamino-2-phenylindole
EF1α	elongation factor 1α
EtBr	ethidium bromide
GCP	gamma –tubulin complex protein
GFP	green fluorescent protein
LIM	lily messages induced at meiosis
MAPK	mitogen activated protein kinase
NLS	nuclear localization signal
NPK1	nucleus and phragmoplast localized protein kinase 1
PAGE	polyacrylamide gel electrophoresis
PMC	pollen mother cell
RACE	rapid amplification of cDNA ends
RFP	red fluorescent protein
rRNA	ribosomal RNA
Rpf2	ribosome production factor 2
RT-PCR	reverse transcription-polymerase chain reaction
SPC97/98	spindle pole body component 97/98

Preface

Microsporogenesis begins with the division of a diploid sporophytic cell, giving rise to the tapetal initial and the sporogenous initial (pollen mother cell). The sporogenous cell undergo meiosis, giving rise to a tetrad of haploid cells, that are released as free microspores by the action of enzyme produced by the tapetum layer of the anther. These uninucleate microspores undergo an asymmetric mitotic division, resulting in a pollen grain containing two cells, a larger vegetative cell and a smaller generative cell that is enclosed within the vegetative cell and finishes when pollen grain is released from the anther. All these processes involve a complex program of spatially and temporally regulated genes in the anther. It was estimated that about 25,000 diverse genes are expressed in the tobacco anther, and approximately 10,000 of these genes encode mRNA species that are anther specific (Koltunow et al., 1990). Studying the abundance and the presence of anther specific mRNA at various stages of anther development has provided valuable information on the temporal and spatial expression of anther genes (Koltunow et al., 1990, Hihara et al., 1996, Oldenhof et al., 1996). Anther development can be divided into two general phases. During phase I, the morphology of the anther is established, cell and tissue differentiation occur and microspore mother cells undergo meiosis to generate tetrads of microspores. During phase II, pollen grains and sperm cells differentiate and tissue degeneration, dehiscence, and pollen grain release occur.

Recently, using Arabidopsis system, genetic approach for the isolation of genes involved in microsporogenesis has been conducted successfully. Young et al., reported the *SPOROCYTELESS (SPL)* gene, also known as the *NOZZLE* gene (Schiefthaler et al., 1999), that is essential for sporogenesis in both male and female organs in Arabidopsis (Young et al., 1999). *spl* mutation blocks the differentiation of primary sporogenous cells into microsporocytes and anther wall formation. The *SPL* gene encodes a nuclear protein related to MADS box transcription factor. *SPL* expression is restricted to sporogenous cells and microsporocytes in anthers and to megasporocytes in the ovules. From these findings, they suggested that development of anther walls and the tapetum

and microsporocyte formation are tightly coupled and *SPL* gene product likely functions as a transcriptional regulator involved in early germline specification.

Also several genes that affect meiosis have been identified. These include Arabidopsis sterile mutants defective in chromosome pairing and segregation: (*asy1*) mutant, defective in meiotic synapsis (Caryl et al., 2000), *dif1/syn1* mutant, a homologue of the yeast *REC8/RAD21* cohesin gene (Bai et al., 1999) and *dmc1* mutant, defective in bivalent stabilization and chromosome segregation (Couteau et al., 1999).

One of the most effective approaches for understanding the molecular mechanisms involved in microsporogenesis is to isolate and characterize genes that are expressed specifically in meiocytes. Previously, 18 meiosis-associated cDNAs were isolated from microsporocytes of a monocotyledonous plant lily (*Lilium longiflorum*) and their corresponding genes designated as *LIM* genes (Kobayashi et al., 1994). Among the gene products, LIM15 that takes part in meiosis-specific homologous chromosome recombination has been well characterized (Kobayashi *et al.*, 1993; Terasawa *et al.*, 1995). However, it has been shown that genes that are expressed in somatic cells play an important role in meiotic progression. For example, a meiotic recombination checkpoint was recently reported to be controlled by mitotic checkpoint genes in yeast (Lydall et al., 1996). Although, the studies of specific genes provided much important knowledge, these results suggest that it is also important to study genes expressed in somatic cells for the elucidation of the processes involved in microsporogenesis.

Random cDNA sequencing can be used to isolate many functional genes, which are expressed in a particular tissue or under a specific environmental condition. Recently, a large number of cDNA clones were identified by random sequencing from plants including maize (Keith et al., 1993), *Arabidopsis thaliana* (Höfte et al., 1993; Newman et al., 1994), *Brassica napus* (Park et al., 1993) and rice (Uchimiya et al., 1992; Sasaki et al., 1994).

In order to identify genes expressed at microsporogenesis, random sequencing was carried out using a lily cDNA library made from zygotene stage microsporocytes. To avoid redundancy of cDNA clones and to obtain novel sequence information efficiently, a simple and versatile "self-hybridization" strategy was developed and single-run DNA

sequence data were obtained from 418 cDNA clones. Deduced amino acid sequences of 171 cDNA clones, which comprised 41% of the selected cDNA clones, showed significant similarity to amino acid sequences registered in the database. These included cDNA sequences for previously identified meiosis-associated genes. However, 59% of the selected clones showed no significant homology to known gene products. Interestingly, compared with the previously described random sequencing projects, the ratio of cDNAs encoding novel protein was relatively high and some of the predicted amino acid sequence encoded by the cDNA exhibited interesting features (Morohashi et al., 2000).

My major purpose of this thesis work is to identify and characterize novel components that may be involved in microsporogenesis or meiosis in plants. In this work, I chose five cDNA clones that contain putative nuclear localization signal(s) and conducted a series of studies for the characterization of novel proteins that may be involved in nuclear or chromosomal events during microsporogenesis in lily.

Part 1. Initial characterization of novel genes from *L. longiflorum* microsporocyte cDNA library

Introduction

Gene expression is a multistep process involving chromatin remodeling, transcription, RNA processing, RNA export and translation. The sequestering of genetic material in the nucleus by eukaryotic cells provides a powerful mechanism for the regulation of gene expression through the selective translocation of proteins between the nucleus and cytoplasm. The best-characterized mechanism for translocation across the nuclear envelope is protein import, which depends on the classical nuclear localization signal (NLS). Meiosis is characterized by two nuclear divisions following a single round of DNA replication. Thus, the dynamics of nuclear function and chromosome structure and behavior are markedly different than that observed during mitosis. Although it is known that the nuclear behavior during microsporogenesis or meiosis is unique, factors and genes involved in the nuclear event are not necessarily clear.

In the aim to better understanding meiosis and mechanisms involved in meiotic nuclear events in plants, I examined several genes expressed during microsporogenesis identified by self hybridization strategy from a cDNA library of zygotene stage microsporocytes in *L. longiflorum* (Morohashi et al., 2000). In this section I conducted studies of four cDNAs encoding novel gene product with deduced amino acid sequences containing a cluster of basic amino acid residues that may constitute a nuclear localization signal. The developmental stage and tissue specific expression patterns of these genes show that they are predominantly or specifically expressed during microsporogenesis. I also investigated intracellular localization of these genes by transient assay in onion epidermal cells using green fluorescent protein (GFP) tag.

Materials and methods

Plant materials

L. longiflorum cv. Hinomoto was used in this study. Flower buds were categorized according to their length and determination of developmental stage was carried out cytologically as previously described (Erickson, 1948). Anthers were dissected from buds, frozen in liquid nitrogen and kept at -80°C. Developing microsporocytes were extruded from corresponding buds, pooled and washed in White's medium to remove sporophytic contamination and kept at -20°C. All samples were frozen in liquid nitrogen and kept at -80°C.

RNA isolation, 5' RACE and RT-PCR

Total RNA was isolated from various tissues and anthers of corresponding stages by using RNeasy kit (QIAGEN), and treated with RNase free DNase. 5' RACE was carried out with 1.0 μg of zygotene to pachytene stage anther total RNA by using SMART RACE cDNA amplification kit (Clontech). Primers used for each clone were: M333R (5'-GGACGGAGAGGCAAGAATCAG-3'), M355R (5'-AACACATCGGTCATCTGT GC-3'), M404R (5'-GCATTCCCCGCATAACACCAACA-3'), M404R2 (5'-CAGCCT GCATC TGCATCACCTGT-3'), M411R (5'-ACACTCACACATCCTACTCGC-3'), M411R2 (5'-GTTCACCCACATGATCCACCTGT-3').

For RT-PCR following primers were used: M333F (5'-GACATCAACCCCACATG TTGAT-3')M355F (5'-AGAGATGGACTCTGCT GCTAGG-3'), M355R, M404F (5'-GTGGAGGATGAGAGGAGGGAGA-3'), M404R, M411F (5'-ACTAAAGGATACA CCACAGC-3'), M411R, EF1αF (5'-GAGGCAGACTGTTGCTGTCGG-3') and EF1αR (5'-AGCAGACTGAAATGAAGATGC-3'). PCR conditions were 94°C, 30 sec; 55°C, 30 sec and 72°C, 1 min; 20 cycles. PCR products were separated on 1.2% agarose gel, transferred to Hybond-N^+ membrane (Amersham). Labeling and hybridization were carried out according to manufacturer's protocol (Gene Images AlkPhos Direct labeling and hybridization system, Amersham) and detected with chemiluminescent detection reagent CDP-Star™ (Amersham). Hybridization was carried out overnight at 55°C.

Construction of GFP fusion and particle bombardment

pBlueScript/M333, pBlueScript/M355, pBlueScript/M404 were excised by *Pst*I-blunt/*Kpn*I, pBlueScript/M411 was digested by *Sac*I-blunt/*Kpn*I and the isolated fragments were inserted into *Bgl*II-blunt/*Kpn*I site of 221-EGFP-C1 vector. Onion epidermal cells were bombarded with 1.6 μM gold particles coated with each construct or 221-EGFP-C1 as a control using PDS-1000/He particle delivery system (BioRad). Samples were kept at room temperature for 8 hours and GFP localization was visualized using a fluorescent microscope and then images were captured by 3CCD camera device (Hamamatsu Photonics).

Results

Sequence analysis of M333, M355, M404 and M411 cDNA clones

The full length cDNA sequence was obtained by 5'-RACE using mRNAs obtained from the whole anther at meiotic prophase.

The M333 cDNA (Fig. 1) encodes a predicted protein of 663 amino acids. The N-terminal half (amino acid residues 71 to 330) of the predicted protein shows strong homology to catalytic domain of the serine/threonine protein kinases. Database searches revealed high homology to a tobacco NPK1 (nucleus- and phragmoplast- localized protein kinase 1), which regulates lateral expansion of the cell plate at cytokinesis (Nishihama et al., 2001). They share 48% identity and 60% similarity over entire length with most conserved region within kinase domain (Fig. 2). Thus, M333 cDNA was named LlNPKL for *Lilium longiflorum* NPK1 like protein. Also significant homology was found with putative protein kinases from Arabidopsis and rice.

The M355 cDNA is 1201 bp in length with an open reading frame encoding 308 amino acids that shows significant homology to a putative protein predicted by the genomic sequences of Arabidopsis and rice (Fig. 3). Because of the low abundance of the M355 transcripts, I could not identify the exact size of the mRNA by RNA gel blot analysis. Although its 5'-noncoding region of M355 mRNA did not contain in-frame

```
ACGCGGGGAT TCAGACCCAT TGAGATGCAG GACATCTTCG GCTCCGTCCG CCGTTCCCTT TCCCTCCGCC CAACCACTGC     80
                        M  Q  D  I  F  G  S  V  R  R  S  L  S  L  R  P  T  T  A              19
CGGCGACGCT GCCGTCGGCA CCTCCTCCGG CAGCATAGCC GACCGGATCG GATCTCGCCT CAGAAAATCG CGGGTGGGGC     160
 G  D  A  A  V  G  T  S  S  G  S  I  A  D  R  I  G  S  R  L  R  K  S  R  V  G  L          46
TCGGCCTAGG GTTAGGGCTC CACCCCAAAC TCCACTCCCC GCCATCCCCA ATCGCTGTCG CCCACCCGAT TCGCTGGCGC     240
  G  L  G  L  G  L  H  P  K  L  H  S  P  P  S  P  I  A  V  A  H  P  I  R  W  R            72
AAAGGGGAGC TCATCGGCTG TGGTGCCTTC GGTCATGTCT ATATGGGCCT GAATCTCGAC TCCGGCGAGC TCCTCGCCGT     320
K  G  E  L  I  G  C  G  A  F  G  H  V  Y  M  G  L  N  L  D  S  G  E  L  L  A  V          99
CAAACAGGTG ATGATTGGAA GCAGCGATGC TTCGAAGGCG AAAGCCAGAG ATCACATACT GGAGCTAGAG GAAGAAATTA     400
 K  Q  V  M  I  G  S  S  D  A  S  K  A  K  A  R  D  H  I  L  E  L  E  E  E  I  K         126
AGCTTTTGAA CAACCTTTCT CATCCCAATA TCGTTAGGTA CTTGGGGACA GCGAGGGAGG AAGATACCTT GAACATCCTG     480
  L  L  N  N  L  S  H  P  N  I  V  R  Y  L  G  T  A  R  E  E  D  T  L  N  I  L           152
CTGGAGTTTG TAACCGGTCG GATCGATATC ATCACTGCTC GGGAAATTCG GGTCTTTCCC GGAGGCGGTA ATAAGAATGG     560
L  E  F  V  T  G  R  I  D  I  I  T  A  R  E  I  R  V  F  P  G  G  G  N  K  N  G          179
ATACAAACAA CTTCTGCTGG GCTTGGAATA TCTCTACAAC AAGATCATGC ACAGAGATAT TAAGGGGGCG AACATTCTTG     640
 Y  K  Q  L  L  L  G  L  E  Y  L  Y  N  K  I  M  H  R  D  I  K  G  A  N  I  L  V         206
TGATAACAAG GTGCATAAGC TGCGACTCGG AAGCATCTAA GCAAGTTGCC CAGTTGGCAA CTATACCAGC TGCCAAGTCA     720
  I  T  R  C  I  S  C  D  S  E  A  S  K  Q  V  A  Q  L  A  T  I  P  A  A  K  S           232
ATGAAAGGTA CTCCATATTG GATGGCACCA GAAGTTATTG TTCAGAGTGG ACATAGCTTC TCTGCTGATA TATGGAGTGT     800
M  K  G  T  P  Y  W  M  A  P  E  V  I  V  Q  S  G  H  S  F  S  A  D  I  W  S  V          259
TGGTTGTACT GTCATTGAGA TGGCTACTGG CAAACCTCCC TGGAGCCAGC AATACCAAGA GGTTGCTGCT CTATTCCATG     880
 G  C  T  V  I  E  M  A  T  G  K  P  P  W  S  Q  Q  Y  Q  E  V  A  A  L  F  H  V         286
TCGGGACAAC AAAGTCACAT CCACCCATTC CAGAATATCT GTCTTCTGAG GCCAAGGATT TTCTTTTGAA ATGCTTGCAT     960
  G  T  T  K  S  H  P  P  I  P  E  Y  L  S  S  E  A  K  D  F  L  L  K  C  L  H           312
AGGGAACCGA ACTCGAGGGC AACAGCTGCT GATTTACTGC AGCATCCCTT TGTTACTGGG GATTATCAGG ATCCACATCG    1040
R  E  P  N  S  R  A  T  A  A  D  L  L  Q  H  P  F  V  T  G  D  Y  Q  D  P  H  R          339
ACTCCATTGC TCTTCAGCTG TGAAACCATC AGGAAACATG CTTCCGACAT CAAGATCGGA CAAAAACTCG CCCTCGATTA    1120
 L  H  C  S  S  A  V  K  P  S  G  N  M  L  P  T  S  R  S  D  K  N  S  P  S  I  M         366
TGTGTGAAGG CTCAGACATT GTAGATTATG AAACTGCAAA CTGCTCAACT GCATACTCTG AGAAGCTTAG GACCAAAACT    1200
  C  E  G  S  D  I  V  D  Y  E  T  A  N  C  S  T  A  Y  S  E  K  L  R  T  K  T           392
ATGTGGCAGA AGAGTACGAG TATTGATATG TGTCTTTGGG ATGACAATGA TGACATGTGT CAGTTCGATG ACAATGATGA    1280
M  W  Q  K  S  T  S  I  D  M  C  L  W  D  D  N  D  D  M  C  Q  F  D  D  N  D  D          419
TTTTCCTGTA GTTGGACCAA GTTTTAATCC AATGTCTGAT CCCTTAGACG ACTGGACATC TAAGTTTGAT GCTAGTCCAG    1360
 F  P  V  V  G  P  S  F  N  P  M  S  D  P  L  D  D  W  T  S  K  F  D  A  S  P  E         446
AGCAAAGAAC CAACTCTCAG AATTTTGGTG AATCGACTGG GGAAGTCAAA GATGACTTCA GACTCCCCTG TGAGCCAGCG    1440
  Q  R  T  N  S  Q  N  F  G  E  S  T  G  E  V  K  D  D  F  R  L  P  C  E  P  A           472
GCCGAAGGCG GCTGTGAAGT CACTGAGACT AAAATTAGAG CCTTCCTGCA TGAAAAGGCC GTTGATCTAA GGAAGCTGCA    1520
A  E  G  G  C  E  V  T  E  T  K  I  R  A  F  L  H  E  K  A  V  D  L  R  K  L  Q          499
AACACCACTA TTTGAGTTCT ACAACACCTT GAATACCAAC GATTGTCATG TTATTGCTGC TTTATGCAAG GAAAATGCTA    1600
 T  P  L  F  E  F  Y  N  T  L  N  T  N  D  C  H  V  I  A  A  L  C  K  E  N  A  T         526
CAAATAATTC AAATCTATTT CCCAAAATCA ATCAATCACC AAGCAAGATG GTAACTCGAA TGACATCAAC CCCACATGTT    1680
  N  N  S  N  L  F  P  K  I  N  Q  S  P  S  K  M  V  T  R  M  T  S  T  P  H  V           552
GATGTTGTGA ATAAAAGAAG CACCGGAAGT TCAAGCCCGC TAGTTTCTAC TTGTGACCTC GATAACAATA GAGTTTTGAA    1760
D  V  V  N  K  R  S  T  G  S  S  S  P  L  V  S  T  C  D  L  D  N  N  R  V  L  K          579
AGAGATACTT TCTCCACGCC TTAATGAGAG TGAAGGATTA CGTCAAGGTC ATCAGCTAGA ACCAAACGTT GCAAGTGTGT    1840
 E  I  L  S  P  R  L  N  E  S  E  G  L  R  Q  G  H  Q  L  E  P  N  V  A  S  V  S         606
CTGGCTGCGA CAGTGCGAGC TTCTATGAGA GGCAGAAGAA GTGGGAAGAG GAGTTGGATC AAGAGCTTGC TAGATTAAGA    1920
  G  C  D  S  A  S  F  Y  E  R  Q  K  K  W  E  E  E  L  D  Q  E  L  A  R  L  R           632
GAAGAAAGGC GCCAATCAGG TCTAGGTATG AAATCTCCAG CTCTTTTTGG GAACCGAAAG AGACATCGTT CTAGATTTGA    2000
E  E  R  R  Q  S  G  L  G  M  K  S  P  A  L  F  G  N  R  K  R  H  R  S  R  F  D          659
TTCCCCAGCC AAATGACTAC TGAAGACTTG ACAGATTCAT ATATGTATAC AATTTTTGTA GCCCAGGCGA TGCATGGAGA    2080
 S  P  A  K  *                                                                           663
TGATTCAGTA TTCCAAGCAA GCAACGGAAG ATCTTCGAAA ATTGGTTAGT TCGTTTCGTG ATTCTGATTC TTGCCTCTCC    2160
GTCCCTGAAA TCTAAGCTTG TATCATTTTG AGGAGTGACG CATAACTATT TGACATGACG ATAGAGATAT TCCAAGTTCG    2240
AGCCATGCTG TGCAATATTC GAGAGATTCA TTAATGCCGA GACAGTTTTG CTGCATATTA TCCACTTTGT CGGATTCACT    2320
TAGATAAAGG CTTTCACATG TAACAAGCTG ACCATCAGGT ATCTGGAAAG AAAAAATGCT AATTGGTTGG TTCAGCCTAT    2400
ACGGCAAATG AAAAAAAAAA AAAAAAAAAA                                                         2430
```

Fig. 1. M333 cDNA and deduced amino acid sequence. Two putative NLSs are boxed. Putative kinase domain is underlined.

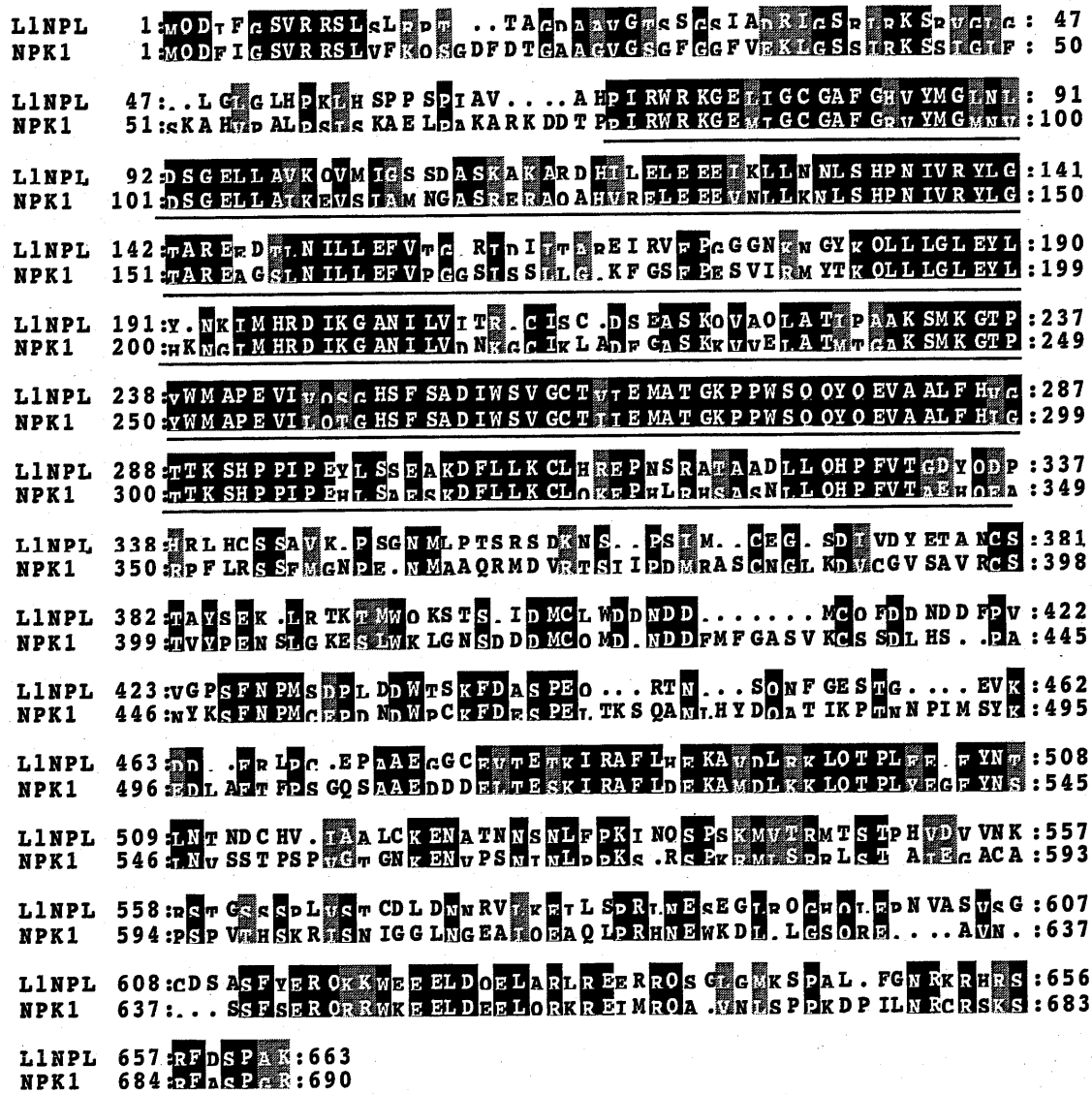

Fig. 2. Amino acid alignment of the LlNPL and tobacco NPK1 proteins. Identical and similar amino acid sequences are boxed in black and gray, respectively. They share 48% identity and 60% similarity over entire length. Sequences corresponding to protein kinase domain are underlined.

```
GGTTAGATTC GTCTCCTCTC CGCCTCTAAC AGCTCACAGC TCGATCGAGA TGGTCGGAAA CGGGAGCCGT    70
                                                      M   V   G   N   G   S   R     7
CGAGGGGATA CTGGAGTTTC CCAGGCGATT AGCAGCCGGA ACGTCTTCGC CGCTCTCGAT ACCCTCAAGA   140
R   G   D   T   G   V   S   Q   A   I   S   S   R   N   V   F   A   A   L   D   T   L   K   K    31
AGAAGAAGAG GTCAGACAAG GAGAAGTGCA AGGTCAAGGA GGAACCGGCA CCGCACGTGT TCTGGGCGCC   210
  K   K   R   S   D   K   E   K   C   K   V   K   E   E   P   A   P   H   V   F   W   A   P    54
CGCGCCCCTC ACCGTCAAAT CCTGGGCCGA TGTCGAAGAC GACGACGACG ATGACTACTA TGCTACCACC   280
 A   P   L   T   V   K   S   W   A   D   V   E   D   D   D   D   D   Y   Y   A   T   T    77
GCGCCGCCGT CAGCGTGGGG TGTGCCTGAG CAGAAGAAGA GCAAGGAGGT CCCTGTTGCT ATAGAGGAGG   350
A   P   P   S   A   W   G   V   P   E   Q   K   K   S   K   E   V   P   V   A   I   E   E   E   101
AAAGTGCAAG TGAAGATGAA GGTCTTGATG AACCTGATGA TGAGGCTGTT GATGAACAAG CTGACCATCA   420
  S   A   S   E   D   E   G   L   D   E   P   D   D   E   A   V   D   E   Q   A   D   H   Q   124
ACTCGAGGTT CCTGCCCCAG CTGAGCCTGA TGTGAAGAAA CCTCCCGTCC CTTCGGTGCC CAAAGATACA   490
 L   E   V   P   A   P   A   E   P   D   V   K   K   P   P   V   P   S   V   P   K   D   T   147
GAAAGGCAAC TCTCGAAAAA GGAGCTGAAG AAAAAGGAGT TGGCAGAATT AGAAGCTCTT TTAACTGAGC   560
E   R   Q   L   S   K   K   E   L   K   K   K   E   L   A   E   L   E   A   L   L   T   E   L   171
TTGGAATGGC AGCTGAAGAT AACCAAGCTG CAAAGCAAAA TGGGACCACA GGTGTTGGCG AGGAGAAACT   630
  G   M   A   A   E   D   N   Q   A   A   K   Q   N   G   T   T   G   V   G   E   E   K   L   194
GGAGGAGCAT AGTGGAGCGA ATGAGAAAAA GGAGAATGTT CCAGTGGAGA GCAAAACATC AAAAAAGAAG   700
 E   E   H   S   G   A   N   E   K   K   E   N   V   P   V   E   S   K   T   S   K   K   K   217
AAGGCGAAAA AGGAGAAACT GTCCAAGGAG GCTAAAGATT CTAATGAACA TCCGAACGGC AATGGTGCCG   770
K   A   K   K   E   K   L   S   K   E   A   K   D   S   N   E   H   P   N   G   N   G   A   D   241
ATGCAGGTAC TAGTGCAGAG GTGACCGCTC GTCCAGTTGA TATGAAAGAG CGGCTGAAGA AGATGGCTTC   840
  A   G   T   S   A   E   V   T   A   R   P   V   D   M   K   E   R   L   K   K   M   A   S   264
AACGAAGAAG AAATCTAGCA AAGAGATGGA CTCTGCTGCT AGGGCTGCCG CGGTGGAGGC AGCTGCCAGG   910
 T   K   K   K   S   S   K   E   M   D   S   A   A   R   A   A   A   V   E   A   A   A   R   287
AGTGCAAAGC TGGCCGCTGC GAAGAAAAAA GAGAAGAGTC ATTACAATCA ACAGCCAATG CGGTAGGCTG   980
S   A   K   L   A   A   A   K   K   K   E   K   S   H   Y   N   Q   Q   P   M   R   *       308
CTGTTGCTAT CTGTGGCTGG CTCATGTTGG GACATCGGTG CTAAATAGTG AGGAAAATAT CTTCTTTTAT  1050
TGGTTGAATG GGAAATAAAC TGAGTTGTCG CACAGATGAC CGATGTGTTC TGTGTGGTAC TAGTTTTTCC  1120
TTAGTCTTGT CAATGCCTTG CATGAATGTG TAGAAGATAC ATGTGACCAT GGTTTTCTTA AAAAAAAAAA  1190
AAAAAAAAAA A                                                                 1201
```

Fig. 3. M355 cDNA and deduced amino acid sequence. Putative NLSs are underlined.

stop codons, high similarity to Arabidopsis homologue suggests that the M355 cDNA encodes a full-length protein. Secondary structure prediction revealed that it contains a coiled coil region (153-184 aa residues).

The M404 cDNA is 1138 bp in length with an open reading frame encoding 219 amino acids that shows no similarity to previously reported proteins deposited in the database (Fig. 4). Like M355, we could not confirm the size of M404 mRNA by RNA gel blot analysis, but the presence of stop codons in all three ORFs suggests that the cDNA encodes a full-length protein.

The M411 cDNA encodes a predicted protein of 540 amino acids (Fig. 5). The putative gene product of M411 cDNA showed high homology to Arabidopsis putative protein. No significant homology was found between M411 and previously known proteins, however, searches for functional motifs revealed sequence similarity to the conserved region of a SPC97/SPC98 protein family (Knop et al., 1997), which are spindle pole body components that form a complex with gamma-tubulin (Fig. 6). A putative signal peptide sequence is also present at N-terminus of the protein.

In addition, searches for functional motifs within the amino acid sequence revealed the presence of clusters of basic amino acid residues that may constitute nuclear localization signals for all four clones.

Expression analysis of LlNPL, M355, M404 and M411 mRNA

The expression pattern of each gene was examined with respect to temporal accumulation during microsporogenesis as well as tissue specificity. Because I could not detect any discrete band by RNA gel blot analysis, I exploited RT-PCR for the detection of mRNAs (Fig. 7). For M355 and M411, corresponding transcripts were detected in anthers of all stages of microsporogenesis and mature pollen. Although higher expression levels are detectable from whole anther and PMC samples, the M355 and M411 mRNA are also detected from somatic tissues including callus, leaf, stem and root. Therefore, the tissue specificity of M355 and M411 mRNA accumulation does not appear to be tightly regulated at the level of transcription. On the other hand, LlNPL and

```
TGCAGTGCTG CAAGACTGCT TCAGAGTCCA GACTCGAGTC GCTATCCTCA TCCCCTCCCG CCGCCGGTGA      70
TCGCCGGCAC TCCAATGAAG CGCAAAAAAT GGACCGGAGT ACGAAGAATC CACCCTCATT ACCGAATACT     140
CCTCCCTCCT CTCATCGGGC GCCCTCTCCA AGCTCCGCTC CCGTGAGAAA AAGTTTCTCC CCATCGCCGA     210
CCGCGTCAAC GCCCTCCACC ACCTCCGCGA CCCCGACTCC TTCCCCTTCC TCTGGTCCTG GCGCGACGTC     280
TCCGTCAAGA TCCAGAACAT GCGCCACCAG TACCTTGGCG TGAAGCAAAA AATCCTCCGC CCCTCTTCCT     350
                  M   R H Q   Y L G V   K Q K   I L R   P S S S      18
CCCCTTCCGC CGCCGCCGAC TTCGACTGGG ACGACGGCGT CAATATCTGG CCCAACTTCC TCCTCTACAA     420
  P S A   A A D   F D W D   D G V   N I W   P N F L   L Y K        41
GCAGGTCTTC GGCGATCTCG ACCTCGATCC CGATCTCGAT TCTGACGATG AGGATGAGGA ATTGGGGGTG     490
 Q V F   G D L D   L D P   D L D   S D D E   D E E   L G V         64
GAAACTGAGG AGGAGACGCG GATGCTGGCG AGGAGGGTGT TGGAGCAGGG GGAGGAGGCG GCGAGGAGGA     560
E T E E   E T R   M L A   R R V L   E Q G   E E A   A R R R       88
GGGTGAGGGA GGAGGAGGAG GAGGATTGGG CGAGGGACTG GCGGCGGAGG GAGCGGGAGA CGGCGGTGGA     630
 V R E   E E E   E D W A   R D W   R R R   E R E T   A V E        111
GGATGAGAGG AGGGAGAGGT GGAGGAGGAG GGAGGAGAGG AGGGAGGAGG AGGAGATGGA GTGGAGGGAG     700
 D E R R E R W   R R R   E E R   R E E E   E M E   W R E          134
AGGATGCTGG TGATGCAGAT GGAGCATGAG AAACAGGTGA TGCAGATGCA GGCTGAGGTG TGTCAAGGTC     770
R M L V   M Q M   E H E   K Q V M   Q M Q   A E V   C Q G Q       158
AGATGCAGAT AGCGGGGATA CTCGCACGGC TGGCGTGCCA GTTTTTCGGG GCCGGGGCTG GCGGGGGTGA     840
  M Q I   A G I   L A R L   A C Q   F F G   A G A G   G G D      181
TGGGGGGATG CCGAGCCAGG TTCTGCAAGG TTTGCAGCAG GAGCAGGAGC ATTCACCGGG AAGCTTGGTG     910
 G G M   P S Q V   L Q G   L Q Q   E Q E H   S P G   S L V       204
GGGGATGATG GGAAGAATGG AGCTCATTCA AGTGAGCACT ATCTCTAGAA TCACTGTAAG TTTGTGTCTC     980
G D D G   K N G   A H S   S E H Y L   *                          219
ATCTAACTTT ATCTTGTTTA GAATCTTGTA CGGCATAGAA ATGTCTGAAT GCTTGGATTT TCAATGATGA    1050
ATCAATTTGA ACATCTGAAA TAAGATAATT GTTGGTGTTA TGCGGGGAAT GCAAATTATA TAGTTATTTC    1120
AAAAAAAAAA AAAAAAAA                                                          1138
```

Fig. 4. M404 cDNA and deduced amino acid sequence. Two putative NLSs are underlined.

```
GCAGTTGGTT AACATGTATC GAGATGAGTT GCCGTATCAT TTGATAGGAG TAATGATTGC ACTCCTTGTG     70
ATATTGGAGA ATATTTCGTG TCACAATCTG GCTATGAGTC AACTAACCAA AGACTTGAAG TGGAAGTTAA    140
CTCTCCAAAA TCCGATGACA GACAACCAAG ACACAATGAT TCAAATGTAA AGCCAACAAG TCATGTAGAT    210
CGTTCGCGGA TAATGGGCGC GTTGACCTTG TCTGAGGTAT TTTTGATTTC CTTGTTGGGT TTAATAGGTG    280
                M  G  A  L  T  L  S  E  V  F  L  I  S  L  L  G  L  I  G  D     20
ATGGCAGTCA TATATATCAG TACTTCAAGA TGTCCTCACC TCCAATCTCT CGACTACACA AAATTGTCAT    350
  G  S  H  I  Y  Q  Y  F  K  M  S  S  P  P  I  S  R  L  H  K  I  V  M      43
GTGTACTGAT TCAGCTAAGG AAACTGGTGG GACTACACAG ACCCCATTGA CTTCCGGCAA CATCTGGTTC    420
 C  T  D  S  A  K  E  T  G  G  T  T  Q  T  P  L  T  S  G  N  I  W  F       66
AAGTTCTTAG CAGATACCAT ATTGGGGAAG ACAAATGTGG ACATTATGAA AGATGTTTCT GTTAAATGTG    490
K  F  L  A  D  T  I  L  G  K  T  N  V  D  I  M  K  D  V  S  V  K  C  A      90
CAAATAGGGA CTCAGATTCT GTTGACTTAA AAGAAATGGA AGACACTGCA GAACATCATA ACCCTAAATT    560
  N  R  D  S  D  S  V  D  L  K  E  M  E  D  T  A  E  H  H  N  P  K  L      113
GGAAGATAGC CGTGCTGAGA ATATACTAGA GACGAGTCTC ATTAGAAGTT CTTGCCTTGG AAACCCTGTT    630
 E  D  S  R  A  E  N  I  L  E  T  S  L  I  R  S  S  C  L  G  N  P  V       136
GTTACTGTAT CCAGGGAACT ACTTGAAAAG AACATATCTT GTTGGAATGA ACTCAACATA TCTAGTAATT    700
V  T  V  S  R  E  L  L  E  K  N  I  S  C  W  N  E  L  N  I  S  S  N  F     160
TCCGCCTTCC TCCACTAAAT GATGGAAATT TACGAGAGGC AATTTTTGGG AAAAATTGTT GGAGTTCAGA    770
  R  L  P  P  L  N  D  G  N  L  R  E  A  I  F  G  K  N  C  W  S  S  D      183
TTTGTTGAAC TACAAAACAG CTCCTGGATA TAATGGAACA GACTATATTT TTGGTTTTCA ATTTAATGAA    840
 L  L  N  Y  K  T  A  P  G  Y  N  G  T  D  Y  I  F  G  F  Q  F  N  E       206
CTAGAACATC TTCGAGTAGA AGATAACACA AAAACCTTGG AAAAATTGTA TCCGTTTCCA ACTATTCTTC    910
L  E  H  L  R  V  E  D  N  T  K  T  L  E  K  L  Y  P  F  P  T  I  L  P     230
CAGGTTTTCA GGATAATGTA CCTATGTCAG AGTTTCTTCC TTTCCAGAAA AATAGCACCC TTGCCTCAAG    980
  G  F  Q  D  N  V  P  M  S  E  F  L  P  F  Q  K  N  S  T  L  A  S  R      253
AATCCTTTGC TGGATTGATA ATATCAAACT AAAGGATACA CCACAGCCTG CAATTATCAT CCAGGAATGC   1050
 I  L  C  W  I  D  N  I  K  L  K  D  T  P  Q  P  A  I  I  I  Q  E  C       276
CTTGCTTTTT ATATTAAAAA ACAGGTGGAT CATGTGGGTG AACATATTTT GTTGAAGCTG ATGAGTGACT   1120
L  A  F  Y  I  K  K  Q  V  D  H  V  G  E  H  I  L  L  K  L  M  S  D  W     300
GGAAATTGAT GGACGAACTT GGTGTGTTAC ATGCTGTTTA TTTACTGGGC TCAGGTGACC TACTGCAGCA   1190
  K  L  M  D  E  L  G  V  L  H  A  V  Y  L  L  G  S  G  D  L  L  Q  H      323
CTTCCTGGCA GTTATTTTTG AGAAGCTGGA TAAGGGAGAA TCCTGGGATG ACGAGTTTGA ATTGAATACT   1260
 F  L  A  V  I  F  E  K  L  D  K  G  E  S  W  D  D  E  F  E  L  N  T       346
GTGTTACAGG AATCCATCAG AAACTCTGCT GATGTACTTC TAAGTGCCCC GGATTCTTTG GTTGTATCGG   1330
V  L  Q  E  S  I  R  N  S  A  D  V  L  L  S  A  P  D  S  L  V  V  S  V     370
TCGCCAAGCA AAATGAATTC GACGACGAGG CAAATACAAC CAGCAATGTT ACTCCCCGCA AAATTCAAAA   1400
  A  K  Q  N  E  F  D  D  E  A  N  T  T  S  N  V  T  P  R  K  I  Q  K      393
GCAATTTCTT GGAATTGATG CCCTTGATTT GATGAAATTC ACTTACAAAG TTCCGTGGCC TCTTGACCTT   1470
 Q  F  L  G  I  D  A  L  D  L  M  K  F  T  Y  K  V  P  W  P  L  D  L       413
ATCGCCAATA TGGAAGCACT AAAGAAGTAT AATCAGGTCA TGGGCTTCTT ATTGAAGGTC AAGCGTGCAA   1540
I  A  N  M  E  A  L  K  K  Y  N  Q  V  M  G  F  L  L  K  V  K  R  A  K     436
AATTTGTTCT TGACAAAGCT CGGAGGTGGA TGTGGAAGAG CAAAAGCAGT ACGGGACACA GTTATAAGCA   1610
  F  V  L  D  K  A  R  R  W  M  W  K  S  K  S  S  T  G  H  S  Y  K  H      460
TCACCTGCTA GTGGAACAGA AGCTCCTCCA CTTTGTTGAT GCATTTCATC AATATGTCAT GGATAGGGTA   1680
 H  L  L  V  E  Q  K  L  L  H  F  V  D  A  F  H  Q  Y  V  M  D  R  V       483
TTTCACAGTG CGTGGCTTGA ACTCTGCACA GGCATGGCGT CAGCAAGCTC CCTTGATGAG GTCATTGAAG   1750
F  H  S  A  W  L  E  L  C  T  G  M  A  S  A  S  S  L  D  E  V  I  E  V     506
TTCATGAGGC CTACCTTTTG TCTATTCAGC GCCAGTGCTT TGTGGCTCCC GACAAGCTGT GGGCCTTGAT   1820
  H  E  A  Y  L  L  S  I  Q  R  Q  C  F  V  A  P  D  K  L  W  A  L  I      530
AGCCAGTTTG ATGAATGCAT TGCCTTTCTA CTAAGAATCC TGTCTTTCAA GCTCAATGTG GGGCACTTTC   1890
 A  S  L  M  N  A  L  P  F  Y  *                                           540
CTCATTTGGC TGATTTGGTG ACCAGAATCA ACTACAACTA CTTCTACATG TCTAGCAGTG GCAACTTGCT   1960
GACTGTCCCA AGCTTTGAAG CTGCAGCTAA ACGTCGACCG ATACTTTAGA TATTACCTTT CTGGGGGCGT   2030
AAATTCAAAG CGAGTAGGAT GTGTGAGTGT TGTGCCTGTT TGTGTCCCGT ACGTCACTGA TGAAAAAAAA   2100
AAAAAAAAA                                                                    2109
```

Fig. 5. M411 cDNA and deduced amino acid sequence. N-terminus signal peptide is double underlined and bipartite NLS is underlined.

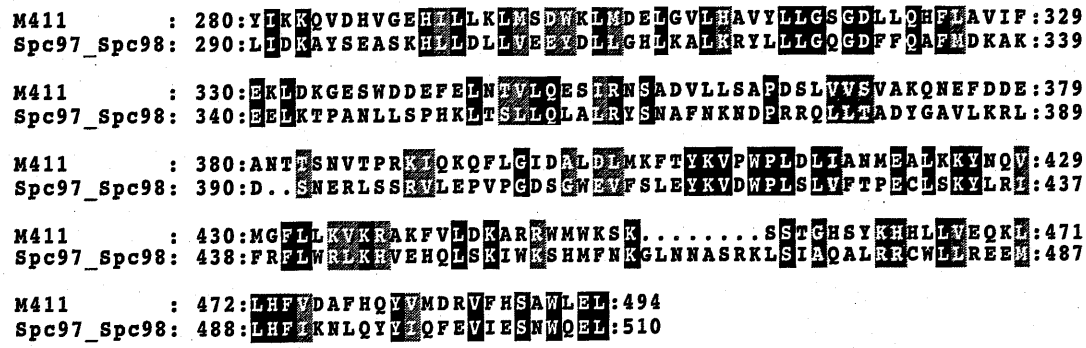

Fig. 6. Sequence comparison of the M411 to the to conserved domain present in Spc97/Spc98 family γ-tubulin complex proteins. Identical and similar residues are boxed in black and gray, respectively.

M404 mRNA expression shows a discrete tissue- and stage-specific expression pattern. As shown in Fig. 7, M404 mRNA accumulates at high levels during early stages of microsporogenesis. LlNPL mRNA expression is restricted to PMC and anther. Although low levels of mRNA accumulation can be detected from somatic tissues, a PMC-specific expression is also evident for M404. These results indicate that the LlNPL and M404 gene expression is associated with male meiosis and is predominantly expressed in PMC.

Intracellular localization of LlNPL, M355, M404 and M411 proteins tagged with GFP

Because amino acid sequences of four proteins indicated the presence of clusters of basic amino acid residues that may constitute NLS, I tested the nuclear localizing activity by fusing GFP and expressed in plant cells. Plasmid vectors expressing GFP fusion proteins were constructed by fusing GFP at the amino terminus of each protein. Plasmids were introduced into onion epidermal cells by microprojectile bombardment and expressed under control of the CaMV 35S promoter. A plasmid construct that expressed GFP alone, designated 221-EGFP-C1, was used as a control. As shown in Fig. 8, fluorescent microscopic observation of GFP signals revealed that GFP::M355 and

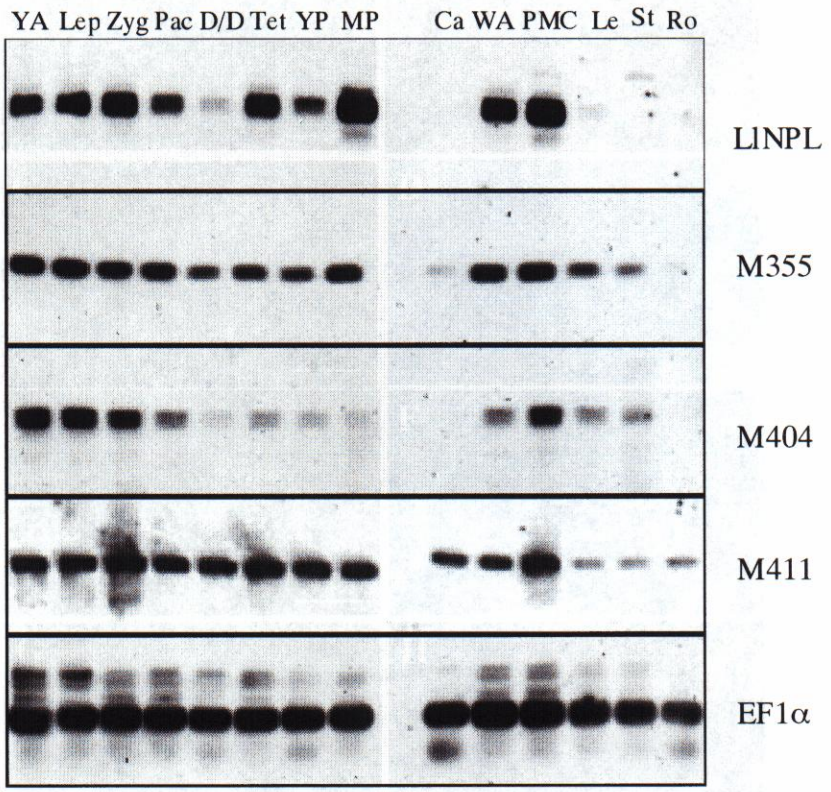

Fig. 7. RT-PCR analysis for mRNA accumulation pattern of LlNPL, M355, M404 and M411. Reverse transcription was conducted with 0.5 μg total RNA from anthers of indicated developmental stages and different tissues. PCR was performed with specific primers for each clone. EF1α was used as control. The PCR products were separated on a 1.2% agarose gel, transferred to nylon membrane, and probed with cDNA fragments corresponding to each of the amplified products. YA, young anther; Lep, leptotene stage; Zyg, zygotene stage; Pac, pachytene stage; D/D, diplotene-diakinesis stages; Tet, tetrads; YP, young pollen; MP, mature pollen; Ca, callus; WA, whole anther; PMC, pollen mother cell; Le, leaf; St, stem and Ro, root.

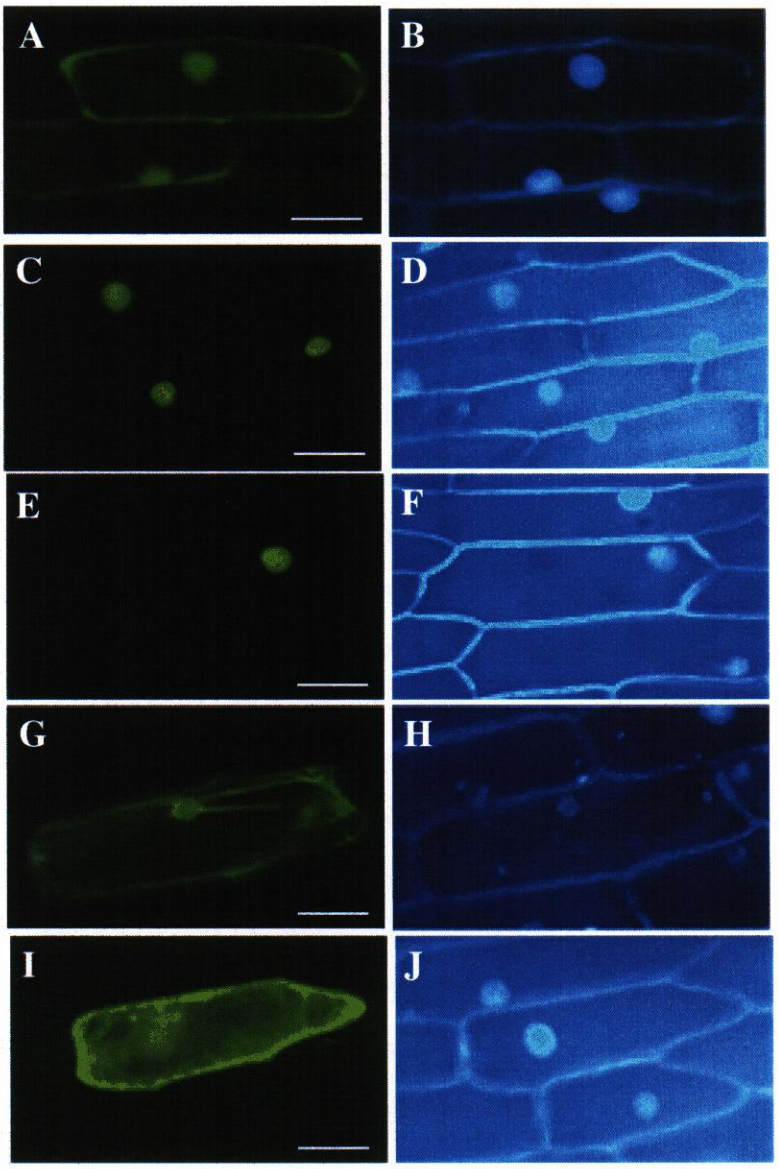

Fig. 8. Intracellular localization of GFP::LlNPL, GFP::M355, GFP::M404 and GFP::M411 fusion proteins. Onion epidermal cells transiently expressing (A) GFP::LlNPL, (C) GFP::M355, (E) GFP::M404, (G) GFP::M411. and (I) GFP alone (221-EGFP-C1 plasmid). GFP localization is visualized by fluorescent microscopy 8 hours after bombardment. (B), (D), (F), (H), (J) are DAPI staining of each panel on the left. Scale bar, 100 μm.

GFP::M404 fusion proteins are targeted to the nucleus, whereas GFP::LlNPL and GFP::M411 fusion proteins are localized in the nucleus and also in the cytoplasm (Fig. 8). These results indicate that M355 and M404 proteins are localized to the nucleus when fused to GFP.

Discussion

M333, M355, M404 and M411 clones were identified as cDNA clones from *L. longiflorum* zygotene stage cDNA library without any selection for specific expression.

M333 cDNA encodes LlNPL, a putative protein kinase of 663 amino acids. The deduced sequence is closely related to a member of MAPK kinase kinase encoded by *NPK1* from tobacco. NPK1 protein accumulates in the equatorial region of the tobacco phragmoplast, suggesting involvement in formation of the cell plate. It was shown that the NPK1 regulates lateral expansion of the cell plate at cytokinesis (Nishihama et al., 2001). Because RNA gel blot analysis of LlNPL showed meiosis specific expression and based on the similarity to NPK1 gene, it is possible that LlNPL also function in the cell plate formation at cytokinesis during meiosis.

Database searches of M411 deduced amino acid sequence revealed no homology except putative protein in Arabidopsis. Some low degree of similarity was found with γ-tubulin complex proteins (GCP). The γ-tubulin complex is necessary for microtubule nucleation at the centrosome. M411 shows 47% similarity and 23% identity over 280 amino acid residues with GSP5 from human and mouse (Murphy et al., 2001). This region also corresponds to the part of the conserved domain in Spc97/Spc98 family proteins. Members of this family are spindle pole body components such as Spc97 and Spc98 in the yeast that form a complex with γ-tubulin. Although the similarity is low, it is possible that M411 might function in the γ-tubulin complex structure and function in *L. longiflorum* during microsporogenesis.

The predicted amino acid sequence of M355 exhibited significant similarity to previously reported putative plant gene products from Arabidopsis and rice whose functions are unknown. On the other hand, M404 shares no homology to previously

described proteins deposited in database. The results of RT-PCR analysis show that the M355 and M404 mRNAs are preferentially expressed during prophase I of microsporogenesis. M355 mRNA can be detected at high levels throughout the progression of microsporogenesis while M404 mRNA accumulation is more specific to PMCs at early stages. These results suggest that the M404 is involved in early meiotic events during microsporogenesis. On the other hand, because low levels of mRNA accumulation are also observed in vegetative tissues, it is possible that both genes are also expressed in vegetative tissue and involved in cellular events other then microsporogenesis or meiosis. Together with the fact that the GFP::M355 and GFP::M404 fusion proteins localize to the nucleus, data obtained in this study suggest that M355 and M404 are nuclear proteins that may play a role in meiotic progression or microsporogenesis. Since the mechanisms of nuclear protein transport during the process of meiosis may be different from those of mitotic cells, it may be necessary to test NLS activity of M355 and M404 in a cell undergoing meiosis. However, because of our inability to express high levels of GFP fusion proteins enough to be visualized by fluorescent microscopy, I could not carry out GFP-tag experiments using meiotic cells. On the other hand, the presence of typical NLSs of meiosis-associated proteins suggests that the NLS sequence capable of mediating nuclear transport in mitotic cells is also active in the meiotic nuclear transport system.

Because nuclear events during meiosis are unique, it is possible that specifically expressed genes such as *LIM* genes (Kobayashi et al., 1994) are involved in nuclear or chromosomal events during meiosis. Previous studies on *LIM* genes suggested that the LIM15 protein that shows similarity to an *E. coli* DNA recombination protein, RecA, is involved in meiosis and plays an important role in meiotic progression (Kobayashi et al., 1993; Terasawa et al., 1995). Also, we have identified the nuclear localizing activity of *LIM5* and *LIM13* gene products both of which harbor NLS and accumulate specifically during microsporogenesis (Ogata et al., 1999). Since LIM5 and LIM13 show no homology to previously described amino acid sequences deposited in the database, it is speculated that they might be involved in the nuclear event unique to lily microsporogenesis. Similarly, because no homology was discovered by the database

searches of the M404 amino acid sequence, available evidence suggests the possibility that the M404 protein is unique to lily and that it is involved in events specific to lily microsporogenesis or meiosis. On the other hand, the presence of putative homologous genes in Arabidopsis and rice indicates that the M355 is a member of a ubiquitous protein that may play a role in microsporogenesis.

The *LIM* genes have been identified as lily meiosis-specific genes specifically expressed during prophase I of microsporogenesis (Kobayashi et al., 1994). The most critical difference in the gene expression pattern of *LlNPL, M355, M404* and *M411* compared to previously described *LIM* genes is the expression of mRNA in the young anther stage. Because LIM cDNAs were selected from zygotene stage cDNA by subtracting cDNAs from young anther, a certain portion of cDNAs for meiosis-associated genes, including *LlNPL, M355, M404* and *M411,* would have been lost during the process of the cDNA selection. Therefore, some of the meiosis-associated genes involved in early stages of meiosis or microsporogenesis would not be identified by the strategy exploited for the isolation of *LIM* genes. I consider it necessary to carry out large-scale sequencing without the subtraction process for the identifying novel meiosis associated genes.

Part 2. Characterization of microsporogenesis-associated nucleolar protein LlRpf2.

Introduction

M532 cDNA was identified from microsporocyte cDNA library as one of the weak signal clones (Morohashi at al., 2000). Although several cDNA clones exhibit similar characters, I chose to work on M532 for further characterization of the protein because of its relatively high expression levels during microsporogenesis that may allow further studies at protein levels. Initial sequence homology search indicated that the M532 may be a member of an evolutionarily conserved protein, whose yeast homologue is required for ribosome biogenesis: in pre-rRNA processing and assembly and localizes to the nucleolus.

Ribosome biogenesis is a complex and tightly regulated process. The complexity of ribosome biogenesis gives rise to many steps and opportunities where regulation could take place. The regulation of pre-rRNA synthesis is a major control point in ribosome biogenesis. Although the steps and mechanisms involved in processing pre-rRNA have been extensively investigated, the overall regulation of and coordination between the processing and assembly of pre-ribosomal particles is not well understood. Many of the steps in ribosome biogenesis have been elucidated by studies in *S. cereviciae*, and most of the ribonucleoprotein and protein factors required for these steps are conserved throughout evolution (Venema and Tollervey, 1999).

Nucleolus is the site for ribosome biogenesis. In addition to being a site dedicated for ribosome biogenesis, a new studies define mitotic and meiotic roles for the nucleolus in cell cycle control and gene regulation acting as a sequestering compartment for regulatory complexes (San-Segundo and Roeder, 1999, Shou et al., 1999).

In this part, I show that M532, named LlRpf2 for *Lilium longiflorum* Rpf2, is a nuclear/nucleolar protein highly expressed during microsporogenesis. In order to obtain information for protein functions, I conducted yeast two hybrid screening and identified a novel lily zinc finger protein, as an interacting protein with LlRpf2.

Materials and Methods

Plant materials

L. longiflorum samples were prepared same as described in Part 1.

5' RACE and Northern analysis

For 5' RACE total RNA from anther of pachytene stage was isolated using a RNeasy kit (QIAGEN), and treated with RNase free DNase. 5' RACE was carried out with 0.5 μg of total RNA by using 5'RACE system (GIBCO BRL) with LlRpf2 specific primers GSP1 (5'-GAATGTAGATCCTCC-3'), GSP2 (5'-CCACCAAGTCCATCGAG GGTCC-3') and GSP3 (5'-GTCGGAGAGCACAGTGCGTGAAG - 3').

For RNA gel blot total RNAs from anthers of corresponding stages and tissues were isolated according to the ATA (Aurine Tricarboxylic Acid) method (Kuhlemeier et al. 1988). 20 μg total RNA was separated on 1% agarose gel containing 18% formaldehyde and 1xMOPS. After electrophoresis, the RNA was transferred to Hybond-N^+ membranes (Amersham) and crosslinked. Labeling of LlRpf2 cDNA and hybridization were carried out as described in Part1. Hybridization was carried out overnight at 55°C.

For AtRpf2 cDNA 3'and 5' RACE the following primers were used: AtM532/1 (5'-GGTTGAAACTGGGAAGAAGACG-3'), MDB1 (5'-CTCTGTAATGACGGAACTC TAC-3'), MDB2 (5'-CTTTGTTCCTTCATGAGGTGC-3'), NMDB2 (5'-CGCCCGCA AGGATTTAAAGTTC-3'). RNA gel blot for AtRpf2 was carried out as described above.

Preparation of antibody and Western blot

The C-terminus portion of the LlRpf2 (amino acids 179-312) was cloned into pET-30b expression vector (Novagen). The construct was transformed into *E.coli* BL21 (DE3) pLysS. The His-tag fusion protein was prepared and purified according to the manufacturer's protocol (pET system manual, Novagen). Purified recombinant protein was used to raise mouse polyclonal antibody. For the preparation of total protein extracts, samples were powdered in liquid nitrogen, homogenized in 2xSDS sample

buffer (0.1 M Tris-HCl, pH 6.8, 1.8% SDS, 18% (w/v) glycerol, 9% (w/v) β-mercaptoethanol). The homogenate was then boiled for 5 min, centrifuged to remove insoluble materials and supernatants were kept at –70°C. The amount of protein was estimated by Bradford's method using Bio-Rad protein assay kit. Protein extracts were separated on 10% SDS-PAGE and electroblotted onto PVDF membrane (Millipore). The membrane was blocked in TBST buffer containing 5% dry milk, washed 3 times with TBST buffer and then incubated overnight with LlRpf2 antibody (1:1000) in TBST buffer containing 1% dry milk. The membrane was then washed three times for 10 min, incubated for 40 min in secondary antibody, anti-mouse IgG conjugated with alkaline phosphatase (1:3000) in TBST containing 1% dry milk. The bound antibody was visualized by using Immune-lite II chemiluminescent protein detection system (BioRad).

Construction of fluorescent fusion proteins and particle bombardment

The *Nco*I-blunt/*Sal*I digested fragment of LlRpf2 was inserted into 221-EGFP-C1 vector, which was digested with *Xho*I-blunt/*Sal*I. Onion epidermal cells were bombarded with 1.6 μM gold particles coated with 2 μg EGFP::LlRpf2 using PDS-1000/He particle delivery system. For colocalization experiment, *Nco*I-blunt/*BamH*I digested LRZ cDNA was fused to 221-Red1-C1 vector digested with *Bgl*II-blunt/*BamH*I. 2 μg of each GFP::LlRpf2 and RFP::LRZ were bombarded into onion epidermal cells. The samples were kept overnight at room temperature and the GFP and RFP were visualized as described in part one.

Cell fractionation experiment

L. longiflorum zygotene stage microspores were used to prepare nuclear and cytosol fractions. Zygotene microsporocytes from about 50 buds were washed three times with buffer B (20 mM Tris-HCl (pH 7.5), 10% glycerol, 0.2 mM ethyleneglycol bis (2-aminoethyl ether) tetraacetic acid, 0.1 mM phenylsulfonylfluoride, 0.5 mM dithiothreitol, 0.3 M sucrose and 10 mM $MgCl_2$) and homogenized in 4 ml buffer B with a glass homogenizer at 1000 rpm. The cell suspension was centrifuged at 4000 g for 10 min at 4°C and the supernatant was used as cytosol fraction. The pellet was

washed two times and resuspended in 1 ml buffer B, overlaid onto 30% percoll in buffer B, centrifuged at 1500 g for 10 min at 4°C, washed two times and resuspended in buffer B without sucrose and $MgCl_2$. The resultant cell suspension was used as nuclear fraction. The amount of loaded nuclear and cytosol protein extracts were equilibrated to the number of fractionated cells. The samples were separated on 15% SDS-PAGE and immunostained as described above. Anti-LlRpf2 polyclonal antibody, anti-LIM10 polyclonal antibody (Takase et al., unpublished) and anti-histone H2B antibody (generous gift from Prof. Tanaka, Yokohama City University) were used at dilution 1:1000.

Immunolocalization

Lily zygotene stage microsporocytes were fixed in 3% paraformaldehyde in 1xPBS for 40 min. After several washes in PBS, microsporocytes incubated in cell wall digesting solution (1% Macerozyme R-10, 0.1% Pectolyase Y-23, 0.1% Zymolyase in 0.3% sucrose) for 10 min. Then samples were blocked with 3% BSA in 1xPBS for one hour and incubated with anti-LlRpf2 and preimmune serum diluted 1:100 in 0.1% BSA in PBS. Primary antibody incubations were conducted overnight at 4°C. The fluorescein isothiocyanate (FITC) labeled goat anti-mouse IgG (Biosource) serum diluted 1:100 was used as a secondary antibody and incubated for one hour at room temperature. The preparations were counter stained with 1 μg/ml DAPI. FITC and DAPI fluorescences were viewed under UV and blue light irradiation using a fluorescence microscope.

Yeast two-hybrid screen and quantitative assay

Yeast two-hybrid screening procedure was carried out as described in Matchmaker library construction and screening kit (Clontech). The C-terminal region of LlRpf2 (amino acids 139-312) was fused to DNA binding domain vector and used as a bait for screening a cDNA library prepared from *L. longiflorum* pachytene stage anther. Yeast cells were plated on selective media lacking histidine, leucine and tryptophan and incubated at 30°C for 7 days. His^+ colonies were assayed for quantitative β-galactosidase activity by luminescent β-galactosidase reporter system 3 (Clontech).

Genomic DNA isolation and DNA gel blot analysis.

Genomic DNA was isolated from whole plant using CTAB method (Rogers et al., 1988). 3.0 μg of genomic DNA was digested with *EcoR*I, *Pst*I and *Xba*I. The DNA fragments were separated on 1% agarose gel, blotted for two hours in 0.4N NaOH to Hybond-N$^+$ membrane. The labeling and hybridization was carried out same as described in Northern blot analysis. Hybridization was carried out overnight at 55°C.

Construction of transgenic plants expressing antisense AtM532.

Full length AtLlRpf2 cDNA was inserted in antisense orientation into pBI121 vector and transformed by agrobacterium mediated transformation using the floral dip method (Clough et al., 1998). Transgenic plants were selected on MS medium containing 30 μg/ml kanamycin and 250 μg/ml carbenicillin. After two weeks, resistant seedlings were transferred to soil and grown in growth chambers under long day condition 16 hours light and 8 hours dark at 22°C.

Results

Analysis of M532 protein sequence

M532 cDNA was identified from *L. longiflorum* zygotene stage microsporocyte cDNA library by self hybridization strategy. Full length of M532 was obtained by 5'-RACE and it is 1284 bp in length with an open reading frame encoding 312 amino acids (Fig.9). The calculated molecular mass of the encoded protein is 35.4 kDa with a predicted pI of 10. Database searches of the predicted M532 amino acid sequences showed no strong homology to known proteins. However, a significant homology was found with a protein family consists of evolutionarily conserved proteins from yeast to human. The highest homology was found with rice and arabidopsis homologues with 69% and 62% identity, respectively. *Rpf2* (ribosome production factor), is a recently described yeast essential gene and shares 54% similarity and 31% identity with M532

```
CCTGAAGCAA GTAAAACCCT AGCCGCCGCT CCCCAACGAG CACCGGGCCG TCGACTAGAC    60
GGCTCCGACA ATCTGGCGAC CGGAGAAGGC GGTTCAGACC TCCGTTCTGC TTCCTCTATA   120
CCTCGCGGCG GCGATCCCAC AAGCTTTAAC AATGATGAAA GTAAAGACTC CCAAGACAAA   180
                                M  M  K   V  K  T  P   K  T  K     10
ACAAGGTGCA AGGGCTCTTG AAAAGCGTGC TCCTAAACTT GTTGAAAATT TCAAGAAGAC   240
 Q  G  A   R  A  L  E   K  R  A   P  K  L   V  E  N  F   K  K  T     30
ACTTATTCTC CATGGTACCA AGACCAGCAA TGTATTGAAT ACCATCTTAA CTCAAATTTA   300
 L  I  L   H  G  T  K   T  S  N   V  L  N   T  I  L  T   Q  I  Y     50
TCACTTGAAA AGAGATAGCG CCATAAAGTA CACCAAGAAG AATGAAAATA TCAATCCTTT   360
 H  L  K   R  D  S  A   I  K  Y   T  K  K   N  E  N  I   N  P  F     70
TGAAAGTGGT GGGGAAACAC GTCTGGAATT TTTCTCCCTA AAGAATGACT GCAGTCTTTT   420
 E  S  G   G  E  T  R   L  E  F   F  S  L   K  N  D  C   S  L  F     90
TGTGTTTGGT TCTCATTCCA AGAAGAGACC CGAAAACCTT GTTCTTGGGA GGATGTATGA   480
 V  F  G   S  H  S  K   K  R  P   E  N  L   V  L  G  R   M  Y  D    110
TCATCATGTC TACGATCTCG TGGAAGTCGG GGTTGAGAAA TTCAAACCTA TGGAGTCATT   540
 H  H  V   Y  D  L  V   E  V  G   V  E  K   F  K  P  M   E  S  F    130
TTCGTACGAG AAGAAGATAG CACCAAAGAT CGGATCAAAA CCTTTCTTTG CTTTTATAGG   600
 S  Y  E   K  K  I  A   P  K  I   G  S  K   P  F  F  A   F  I  G    150
TGAAGGCTTT GAGGGTGCCG CAGAACTAAA GCACTTGAAG GAAGTGTTGC TTGATCTCTT   660
 E  G  F   E  G  A  A   E  L  K   H  L  K   E  V  L  L   D  L  F    170
TAGAGGACAG GTCGTGGAAA ATTTAAACCT CGCTGGGATA GATCGAGTTT TTGTATGCAC   720
 R  G  Q   V  V  E  N   L  N  L   A  G  I   D  R  V  F   V  C  T    190
AGCCATATCC TCGACTACCG TATTCTTCAC GCACTGTGCT CTCCGACTTA AAAGATCAGG   780
 A  I  S   S  T  T  V   F  F  T   H  C  A   L  R  L  K   R  S  G    210
GACAAATATT CCGAGAATGG AGTTAGTGGA AGTTGGACCC TCGATGGACT TGGTGGTTCG   840
 T  N  I   P  R  M  E   L  V  E   V  G  P   S  M  D  L   V  V  R    230
TCGTCATCGC CTCCCTAATG ACAGCTTAAA GAAAGAGGCA ATGAAAACTG TCTCTCACCA   900
 R  H  R   L  P  N  D   S  L  K   K  E  A   M  K  T  V   S  H  Q    250
AGCTAAGAAG ATGAAAAACG TTAACAAGGA TGCAGTGAAA GGAGAAGTAG GGAGGATCTA   960
 A  K  K   M  K  N  V   N  K  D   A  V  K   G  E  V  G   R  I  Y    270
CATTCCAGAC CAACAGGTTG GAGGGATGGC TTTGTCGAAG GATGTTAAAG GACTGAAGAG  1020
 I  P  D   Q  Q  V  G   G  M  A   L  S  K   D  V  K  G   L  K  R    290
GGAGCGACGT GATGCGAAGA ACACTCGGAA GGGGGATCAA TTGAAGAAGA GAAGGACGGT  1080
 E  R  R   D  A  K  N   T  R  K   G  D  Q   L  K  K  R   R  T  V    310
CTCCAAGTGA TGCCTTTGTT GTCTTATCTG GTGCTGTAAT GTGTTGTAGG CTATACTGGT  1140
 S  K  *                                                           312
TTGGTGTTCT GCTTCAGCTT TGGGATATCC AATCTAAGTG AACTAATCAA TTACAGTGAA  1200
GTATATACTT CGTAGCTGAT AATTCCTTAT ATTTTATCAT CCCTTTAGAA GGTTTTTCTT  1260
CTTTTAAAAA AAAAAAAAAA AAAA                                         1284
```

Fig. 9. Nucleotide and deduced amino acid sequence of M532 gene. Putative NLSs are underlined. Putative σ^{70}–like RNA binding motif is written in bold letters.

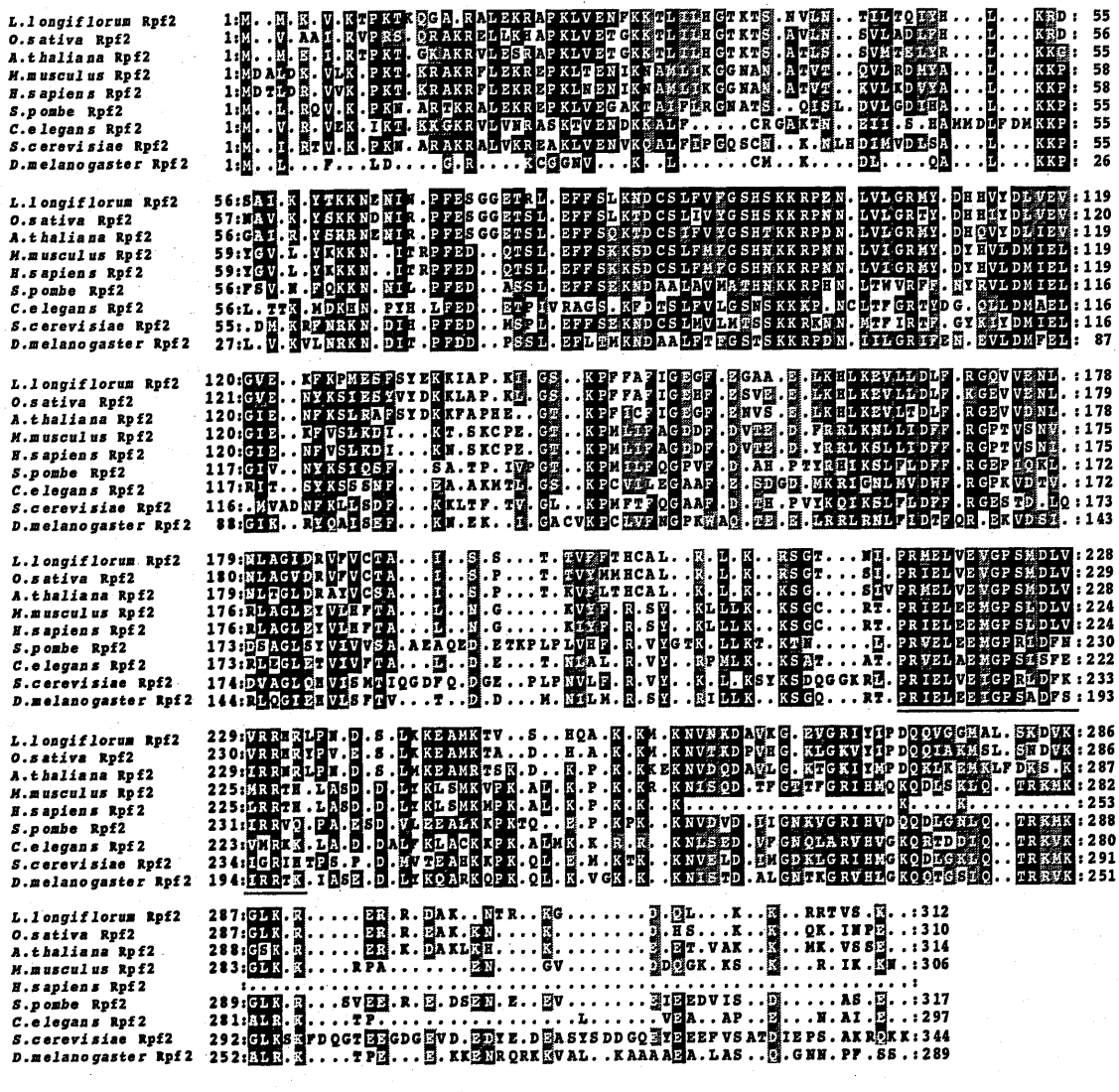

Fig. 10. LlRpf2 is a member of a conserved gene family. Sequence comparison of LlRpf2 with homologous proteins from *O. sativa* Rpf2 (BAB32971), *A. thaliana* Rpf2 (BAB02781), *M. musculus* Rpf2 (BAA95117), *H. sapiens* Rpf2 (BAB14983), *C. elegans* Rpf2 (AAF60765), *D. melanogaster* Rpf2 (AAF55514), *S. pombe* Rpf2 (CAB54156), and *S. cereviciae* Rpf2 (AAF60765). Putative σ^{70}-like RNA binding motif is underlined. Identical and similar residues are boxed in black and gray, respectively.

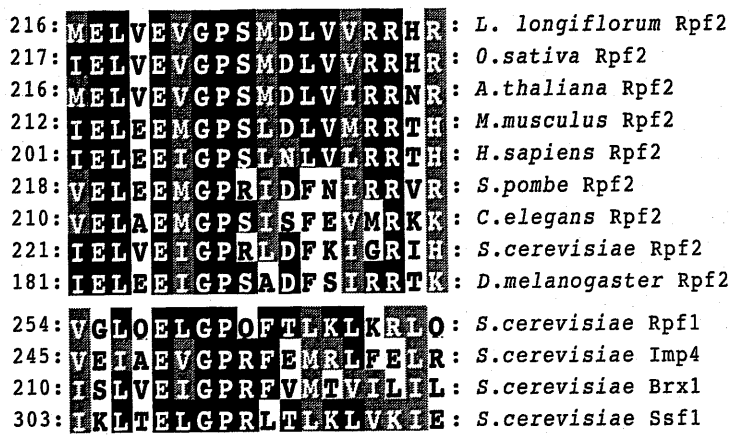

Fig. 11. Sequence comparisons of the σ^{70}-like domains. LlRpf2 and family members, and *S. cerevisiae* homologs of other four conserved proteins that possess σ^{70}-like domains are shown aligned. GenBank accession numbers are *S. cerevisiae* Rpf1 (AAB68926), *S. cerevisiae* IMP4 (AAB95949), *S. cerevisiae* Brx1 (CAA99087), and *S. cerevisiae* Ssf1 (AAB68379).

over the entire length (Zuniga et al., 1999; Wehner et al., 2002). The Rpf2 protein is a member of Imp4 superfamily that possesses a common RNA binding motif and required for ribosome biogenesis (Wehner et al., 2002). This σ^{70}-like RNA binding motif also is well conserved in M532 sequence (Fig. 11). Based on the sequence homology to yeast Rpf2, M532 clone was named as LlRpf2 for *Lilium longiflorum* Rpf2. Sequence analysis revealed potential nuclear localization signals at positions 229 (RRHR) and 305 (KKRR).

Tissue specificity and temporal expression of LlRpf2 during anther development

The expression levels of *LlRpf2* gene in various tissues and anthers at different developmental stages were investigated. The results of RNA gel blot analysis indicated

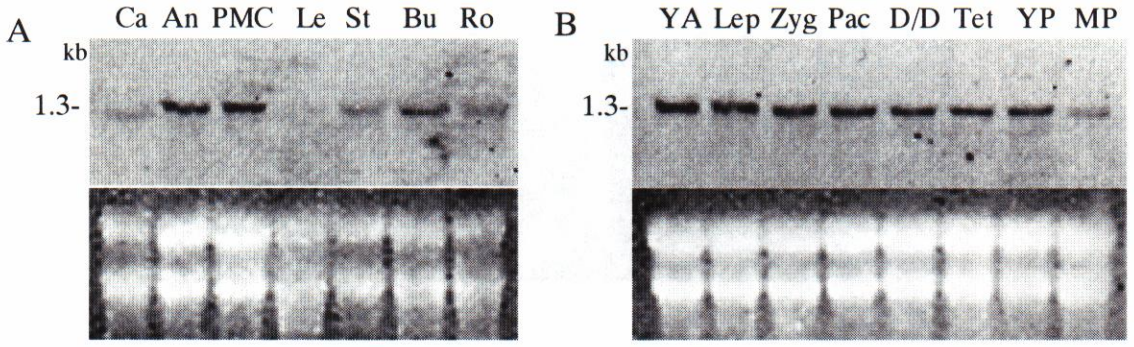

Fig. 12. Expression of LlRpf2 mRNA in different tissues (A) and during anther development (B). 20 μg of total RNA of indicated tissues and anthers at different developmental stages were separated on denaturing agarose gel and hybridized with alkaline phosphatase labeled M532 cDNA as a probe. Ethidium bromide staining shows equal amounts of total RNA were loaded. Ca, callus; An, anther; PMC, pollen mother cell; Le, leaf; St, stem; Bu, bulb; Ro, root; YA, young anther; Lep, leptotene stage; Zyg, zygotene stage; Pac, pachytene stage; D/D, diplotene-diakinesis stages; Tet, tetrads; YP, young pollen; MP, mature pollen.

the presence of LlRpf2 transcripts with a size of approximately 1.3 kb in anther, pollen mother cell and bulb. Although the highest expression level was observed in pollen mother cell and anther, LlRpf2 transcript was also detected in tissues examined: callus, stem and root (Fig. 12A).

Analysis of LlRpf2 gene expression during anther development indicated that LlRpf2 mRNA is present at young anther stage, continued to be expressed throughout the prophase I of meiosis with a gradual decrease and almost disappears with pollen maturation (Fig. 12B). These results indicate that LlRpf2 gene is associated with meiosis and predominantly expressed in anther tissue.

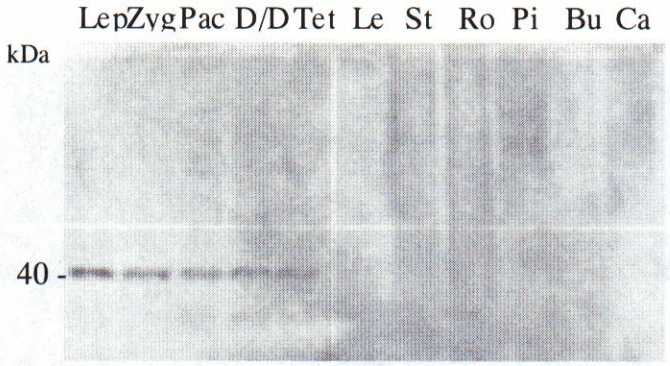

Fig. 13. Immunoblot analysis of LlRpf2 protein. 5 μg protein lysate from microsporocytes of indicated stages of development and different tissues was separated on 10% SDS-PAGE and detected with polyclonal anti-M532 antibody. The molecular weight size markers are indicated in the left in kDa. Lep, leptotene stage; Zyg, zygotene stage; Pac, pachytene stage; D/D, diplotene-diakinesis stages; Tet, tetrads; Le, leaf; St, stem; Ro, root; Pi, pistil; Bu, bulb; Ca, callus

Accumulation pattern of LlRpf2 protein

To investigate the LlRpf2 gene product, I exploited the immunological approach using antibodies against LlRpf2 protein. A plasmid DNA for the overexpression of histidine-tagged LlRpf2 protein was introduced into *E. coli* and induced the expression by the addition of 0.4mM IPTG. The recombinant protein was purified by Ni^{2+}-NTA column chromatography and purified fraction was used as an antigen for anti-LlRpf2 polyclonal antibodies.

The tissue and developmental stage specific expression of LlRpf2 protein was investigated by protein gel blot analysis with using LlRpf2 antibody. The antibody was raised against recombinant LlRpf2 protein expressed in *E.coli*. Immunoblot analysis revealed that the anti-LlRpf2 antibody specifically recognizes a protein of approximately 40 kDa from total protein samples of prophase I of meiosis. No clear

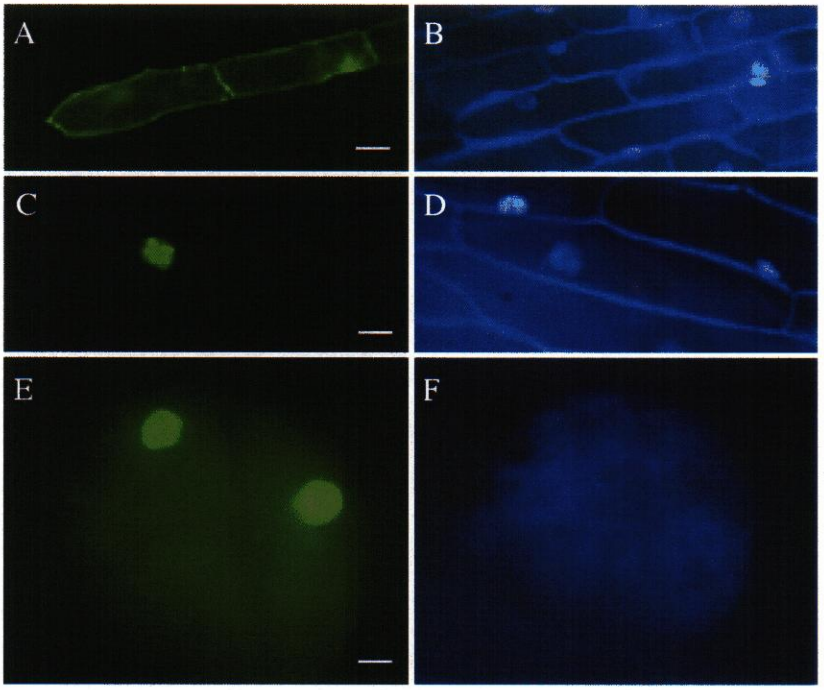

Fig. 14. Intracellular localization of GFP::LlRpf2 fusion protein. Onion epidermal cells were bombarded with GFP::LlRpf2 construct (C and D) and GFP (221-EGFP-C1; A and B). Localization of GFP alone and GFP::LlRpf2 fusion protein were visualized by fluorescent microscopy as shown in A and C. The same cells stained with DAPI are shown in B and D. Larger images of the nucleus of C and D are shown in E and F. Bars in A and C is 100 μm, and the bar in E is 10 μm.

signal was obtained from tissues examined including leaf, stem, root, pistil, bulb and callus (Fig. 13). The calculated molecular weight of deduced amino acid of LlRpf2, 35.4 kDa, was close to the size of detected bands.

Intracellular targeting of the GFP::LlRpf2 fusion protein

In order to confirm the functionality of nuclear localization signals present in LlRpf2 amino acid sequence, translational fusion of the LlRpf2 to GFP was constructed

and tested in transient expression assays using onion epidermal cells. The 221-EGFP-C1 plasmid that expresses GFP alone was used as a control and localized throughout the cytoplasm and in the nucleus (Fig. 14A). Localization of GFP::LlRpf2 fusion protein was targeted to the nucleus (Fig. 14C). Furthermore, a specific subnuclear localization of GFP::LlRpf2 was observed; a high accumulation was detected in the nucleolus (Fig. 14E).

Immunoblotting of LlRpf2 protein in subcellular fractions of microsporocytes

Further I investigated the intracellular localization of LlRpf2 protein during prophase I of meiosis using cell fractionation experiment. Microsporocytes from zygotene stage buds were homogenized and fractionated nuclear and cytoplasmic protein extracts were analyzed by using anti-LlRpf2 polyclonal antibody. The protein of the same size in the gel blot analysis was detected only in the nuclear fraction (Fig. 15, indicated with arrowhead). Also a band with a higher molecular weight was found in a cytosolic fraction, which is considered as a nonspecific binding to polyclonal anti-LlRpf2 antibody (Fig. 15, indicated with asterisk). The LIM10 protein that shows high homology with cytoplasmic small heat shock protein was detected from the cytosolic fraction whereas the histone H2B signals were detected from nuclear protein extracts.

Immunolocalization of the LlRpf2 protein

To examine the spatial distribution of LlRpf2 protein and to confirm the intracellular and/or intranuclear localization obtained by transient expression of GFP::llRpf2 in onion cells, I conducted immunocytochemical experiment. Lily zygotene stage microspoprocytes were treated with anti-LlRpf2 polyclonal antibody and then stained with FITC labeled anti-mouse IgG serum. In zygotene microsporocytes, LlRpf2 protein was highly accumulated in the nucleolus faintly stained with DAPI (Fig. 16 C, D). No immunofluorescence signal was observed with samples stained with preimmune serum (Fig. 16 A).

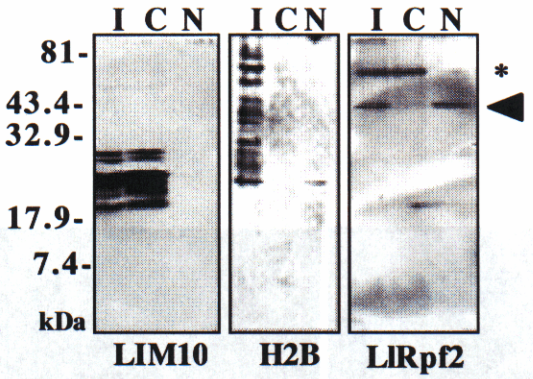

Fig. 15. Subcellular distribution of LlRpf2 protein. Protein samples of total input (I), cytosolic (C) and nuclear (N) fractions from lily microsporocytes were resolved on 15% SDS-PAGE. Zygotene stage microsporocytes were homogenized and fractionated by centrifugation. Protein band signals were detected using specific antibodies against LlRpf2, histone H2B (nuclear protein control) and LIM10 (cytosolic protein control), respectively. The size markers in kDa are indicated on the left.

LlRpf2 interacts with LRZ, a novel zinc finger protein from lily

In order to identify proteins that interact with LlRpf2, C-terminal portion of LlRpf2 was fused to GAL4 DNA binding domain (BD) and used as bait in a yeast two-hybrid screen of pachytene stage anther cDNA library from lily. From a screen of approximately $4x10^7$ cells, one cDNA was identified as a putative clone encoding an interacting protein. The insert sequence encodes a protein of 140 amino acids. Analysis of the protein sequence revealed a C2H2 Zn-finger domain and a nuclear localization signal within Zn-finger domain (Fig. 17). The Zn-finger follows the general pattern C-x2-C-x(12/16)-H-x5-H, and is different from the 'classical' DNA-binding C2H2 Zn-finger. This type of zinc fingers is present in matrin, U1 small nuclear ribonucleoprotein C and other RNA-binding proteins. However, other than the Zn-finger domain, database searches revealed no homology to known proteins. These results indicate that the

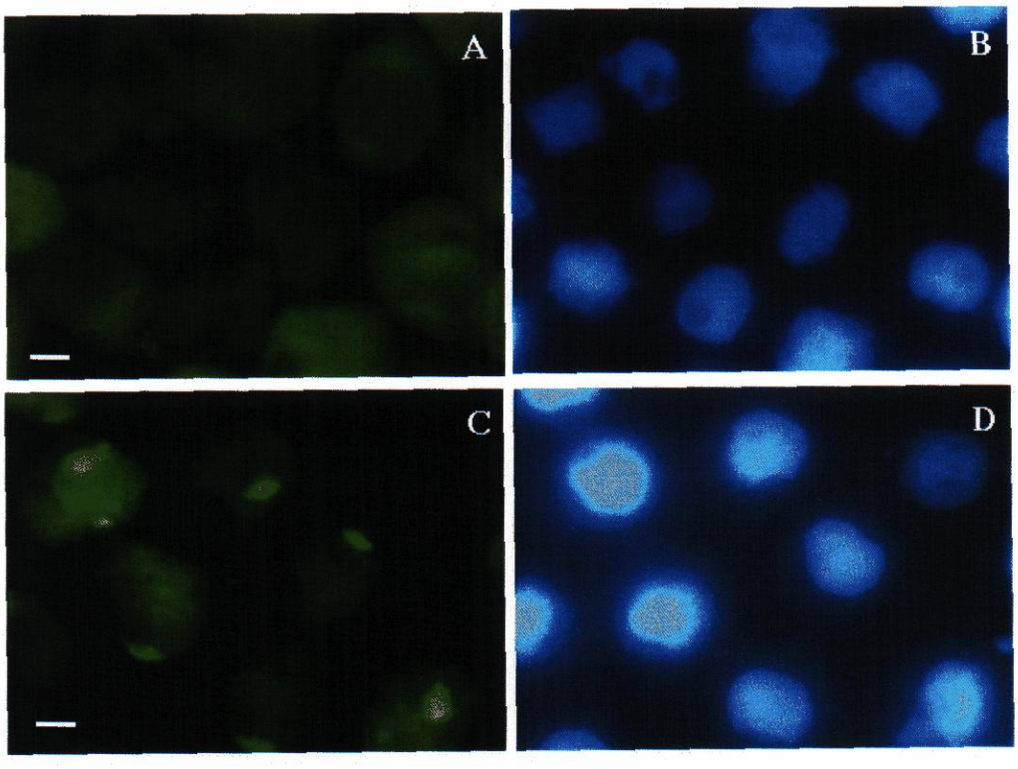

Fig. 16. Immunolocalization of LlRpf2 on lily microsporocytes. Immunofluorescent images of squashed zygotene stage microsporocytes using (A) pre-immune serum as a control and (C) anti-LlRpf2 antibody and secondary antibody conjugated to fluorescein isothiocyanate. (B and D) DAPI staining of (A) and (B). Scale bar, 50 μm.

identified LlRpf2 interacting protein is a novel protein with C2H2 type Zn-finger and NLS, and I named this protein as LRZ (lily Rpf2 interacting zinc finger protein). As shown in the quantitative assay for β-galactosidase activity, the interaction of AD::LRZ with BD vector only was about the same level with the negative control p53 and Lamin C, whereas the level was much higher in the interaction of AD:::LRZ and BD::LlRpf2 (Fig. 18). Results also suggested that the LRZ protein is associated weakly with LlRpf2 comparing to the positive control interaction of p53 and SV40 proteins.

Next, I investigated the intracellular localization of LlRpf2 and LRZ proteins using two different fluorescent tags. GFP::LlRpf2 and RFP::LRZ fusion proteins were transiently expressed by cotransformation in onion epidermal cells. When introduced alone RFP::LRZ fusion protein localizes to the nucleus, and when introduced together with GFP::LlRpf2, LRZ protein colocalizes with LlRpf2 in the nucleoplasm and nucleolus (Fig. 19).

```
GGGAACATGG TGGCAAGGGA GAAAGCAAAG ATGGAGTACA ATGAATTAGT GCGTGCCCTC    60
G  N  M  V   A  R  E   K  A  K   M  E  Y  N   E  L  V   R  A  L     20
GCAGGCTTCG CCAGGGAGAA GGACACGCGG GTGATTGAGA TGGAGGAGAA GAGGAAATTG   120
A  G  F  A   R  E  K   D  T  R   V  I  E  M   E  E  K   R  K  L     40
GAGGAGGAGA AGCGACGAGA GGAGGAGGAA GTACGGAGGG CTGAGAGGGA GGAGGAGACG   180
E  E  E  K   R  R  E   E  E  E   V  R  R  A   E  R  E   E  E  T     60
AGGAAGAGGT TGGAGCGGGT TTTGAGATGC GGAGAGGTGG AGGAGACTGA GGATAGTGAT   240
R  K  R  L   E  R  V   L  R  C   G  E  V  E   E  T  E   D  S  D     80
GAGTTGTACT GTGTCGTGTG TAACCAAACT TTGAAGTCCG AGAAACAATG GAGGAAGCAC   300
E  L  Y  C   V  V  C   N  Q  T   L  K  S  E   K  Q  W   R  K  H     100
AGGAAGTCGA AGAAGCACCA AGATAAGCTT GTTGGGTTGA AGATTTCTTT GAATATAAGG   360
R  K  S  K   K  H  Q   D  K  L   V  G  L  K   I  S  L   N  I  R     120
GCTGAACCTG TCTTTCACCA GACCAGCAGC AATCCATATT GGGCATTGTT GGAAACAAGT   420
A  E  P  V   F  H  Q   T  S  S   N  P  Y  W   A  L  L   E  T  S     140
TAAGATTTGT ATGAATCCTC ATCTCTAATT GATTGTCTAG GTAGTACTTG TTTAAGATGT   480
*
AATAGAAATT AAATTCAGAA AATAGTTATT TTTTTCTATG TTGACTAAAA AAAAAAAAAA   540
AAAAAAAAAA AAAAAAA                                                  557
```

Fig. 17. Nucleotide and deduced amino acid sequence of LRZ cDNA. A C2H2 type zinc finger motif is shown in bold and the Cys and His residues of the zinc finger domain are circled. Putative NLS is underlined.

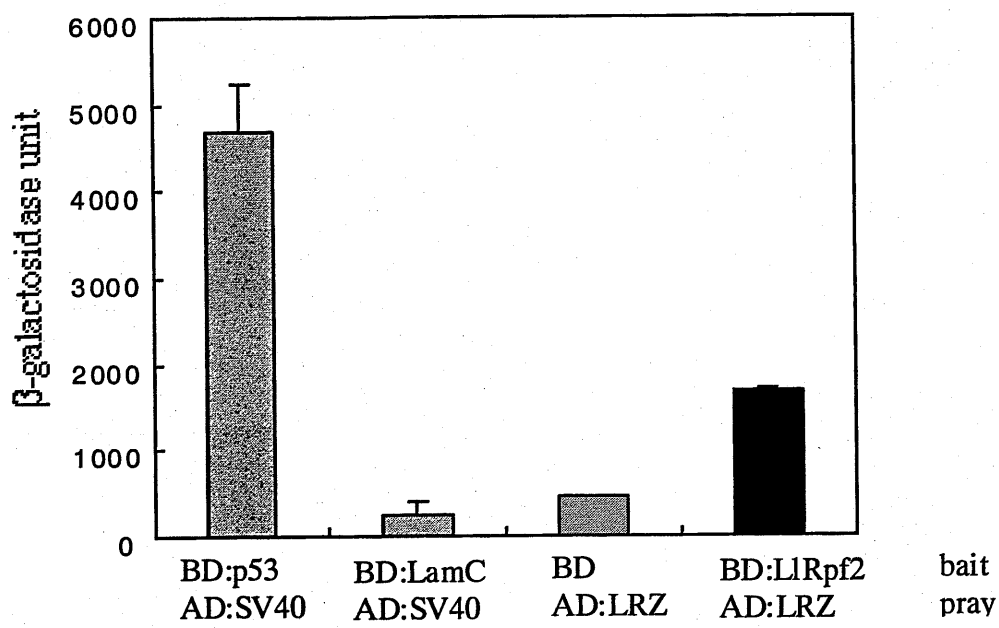

Fig. 18. LRZ interacts with M532 in a yeast two-hybrid assay. The β-galactosidase quantitative assay of p53 and pSV40 (positive control), p53 and pLaminC (negative control), binding domain (BD) only and AD:LRZ, BD:M532 and AD:LRZ proteins. The β-galactosidase activity was quantified in liquid culture assay using chemiluminescent Galacton-Star reaction mixture as a substrate. Error bars show SD (n=3).

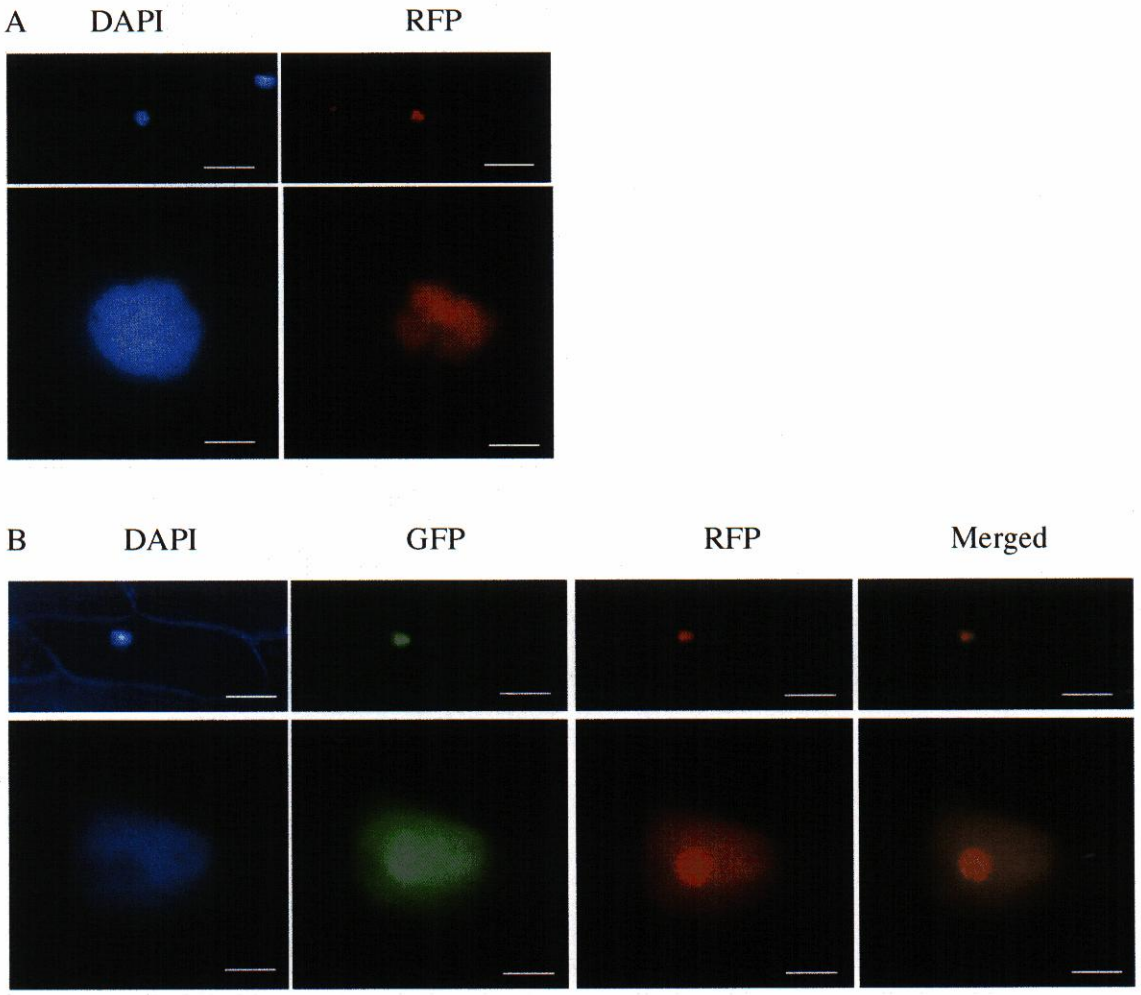

Fig. 19. LlRpf2 and LRZ proteins localize in the nucleus when cotransformed in onion epidermal cells. (A) Intracellular localization LRZ protein fused to red fluorescent protein dsRed1. (B) Colocalization of GFP:: LlRpf2 and RFP::LRZ proteins. Equal amount of plasmids were bombarded into onion epidermal cells and observed under fluorescence microscopy after overnight incubation at room temperature. Under each photo, enlarged view of nucleus is shown. When GFP and RFP colocalize, the merged image appears in yellow. Scale bars are 100 μm and 10 μm for whole cell and nuclear image capture, respectively.

Cloning and characterization of Arabidopsis LlRpf2 homologue

Because LlRpf2 sequence showed high homology to a putative ORF predicted by the Arabidopsis genomic sequence, I carried out an isolation and characterization of the LlRpf2 homologue of Arabidopsis, AtLlRpf2, and cloned a 1.2 kb cDNA corresponding to AtLlRpf2. Predicted amino acid sequence of AtLlRpf2 shows high homology to LlLlRpf2, sharing 77% similarity (Fig. 9). DNA gel blot analysis and search of the completed sequence of the Arabidopsis genome shows AtLlRpf2 gene exists as a single copy gene (Fig. 20). Expression analysis shows that AtLlRpf2 is expressed at high levels in reproductive organs. However, unlike lily, AtLlRpf2 expression was also high in root. In an attempt to reveal possible function of AtLlRpf2 in Arabidopsis growth and development, I generated transgenic plants expressing antisense AtLlRpf2. A full length of AtLlRpf2 cDNA was inserted in antisense orientation in pBI121 vector and expressed under the control of CaMV 35S promoter. From the screening of three independent transformations I obtained about 50 independent transgenic lines. From these, six lines exhibited dwarf phenotype (Fig. 21). The growth of these antisense lines was delayed when compared with wild type. Antisense expression of AtLlRpf2 did not cause any other visible phenotypic changes.

Discussion

LlRpf2 cDNA was isolated from zygotene stage microsporocyte cDNA library by self-hybridization strategy that was efficiently used to isolate many novel meiosis-associated genes (Morohashi et al. 2000). The results of RNA gel blot analysis indicate that LlRpf2 mRNA is expressed at prophase I of meiosis, however low level mRNA accumulation was also observed in vegetative tissues. On the other hand, using anti-LlRpf2 antibody a specific 40 kDa protein was detected from total protein samples of prophase I stage of meiosis but not from vegetative tissues examined. These results indicate that the temporal as well as spatial expression of LlRpf2 is not necessarily

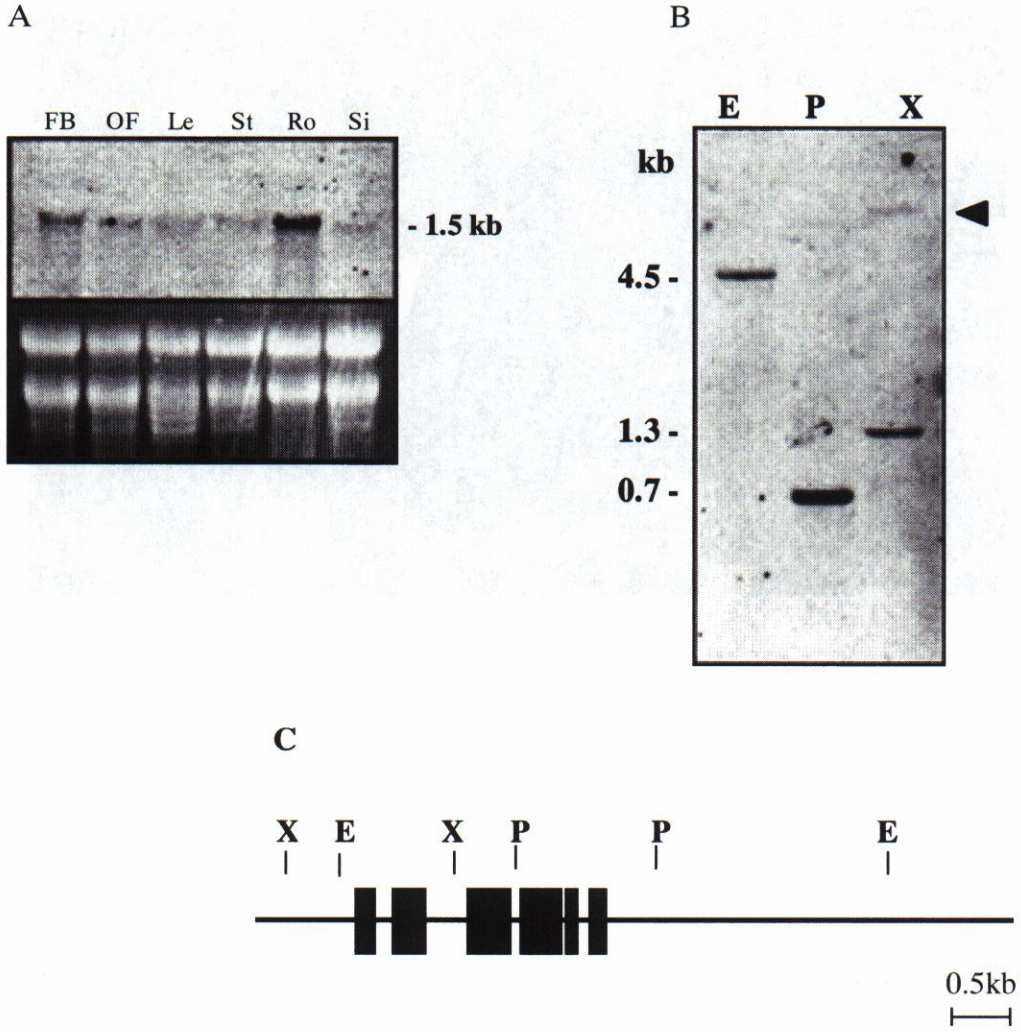

Fig. 20 Characterization of AtRpf2. A. Tissue specific expression of AtRpf2. Total RNA was isolated from floral bud (FB), open flower (OF), leaf (Le), stem (St), root (Ro) and silique (Si). 20 μg of total RNAs were electrophoresed on 1% agarose/formaldehyde gel and probed with alkaline phosphatase labeled AtRpf2 cDNA. EtBr staining of rRNA is assessed as a loading control. B. DNA gel blot analysis of AtRpf2 gene. Genomic DNA was digested with *EcoR*I (E) *Pst*I (P) *and Xba*I (X). An arrowhead shows the position of weak hybridization signal that is consistent with the presence of *Xba*I site within the sequence. C. Genomic structure of AtRpf2 sequence. Exons (dark boxes) and restriction sites are indicated.

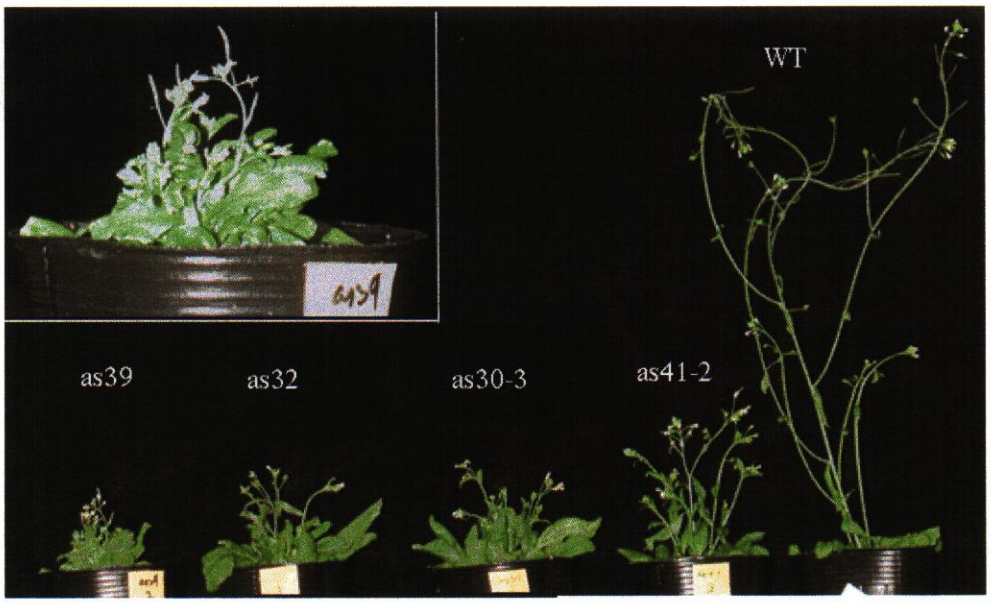

Fig. 21. Antisense expression of AtRpf2 gene results in a dwarf phenotype. Full length AtRpf2 cDNA was cloned in antisense orientation into pBI121 vector and expressed under the control of CaMV 35S promoter. Four antisense lines are shown in comparison to wild type plant. Other two antisense lines showing similar phenotype is not shown.

regulated at the level of transcription and suggest a possible involvement of post-transcriptional regulation in LlRpf2 expression. Previous studies on specific gene expression in plants focused mainly at the level of transcription or mRNA accumulation, however, the results obtained in this study point out the importance of post-transcriptional regulation that regulate a specific control of gene function.

I studied the intracellular localization of LlRpf2 protein by cell fractionation of microsporocytes with anti-LlRpf2 antibody. The LlRpf2 protein was detected in nuclear fraction of microsporocytes. In immunocytochemical experiment using anti-LlRpf2 antibody in microsprocytes and transient expression in onion epidermal cells using GFP tag, LlRpf2 protein was found to be highly accumulated in the nucleolus. These results suggest that LlRpf2 is a nuclear/nucleolar protein associated with meiosis or microsporogenesis.

In a yeast two-hybrid screening, a novel protein with U1-like zinc finger was identified as an interacting protein with LlRpf2. Zinc finger proteins are a class of regulatory proteins that participate in a variety of cellular activities such as development and differentiation. In addition, it is known that many zinc fingers mediate binding to DNA, RNA or protein. The binding properties depend on the amino acid sequence of the finger domains and of the linker between fingers. A specific C2H2 Zn-finger, named U1-like C2H2 zinc finger, with general pattern C-x2-C-x12,16-H-x5-H, is conserved in matrin and several RNA-binding proteins such as human U1 small nuclear ribonucleoprotein (snRNP) C, yeast pre-mRNA splicing factor and mouse spliceosome associated protein. Except matrin, all other proteins with U1-like Zn-finger are considered to be involved in spliceosome assembly and splicing. In U1 snRNP-C, Zn finger like motif is essential for the binding to U1 snRNP particles (Nelissen et al., 1994). The presence of a U1-like C2H2 Zn-finger in LRZ protein points to a possibility of involvement of LRZ protein in basic or alternative splicing of pre-mRNAs. Because no homologous gene was found by the database search, the LRZ may be involved in nuclear events unique to lily microsporogenesis.

LlRpf2 is an evolutionarily conserved protein from yeast ho human. There have been a few reports on *S. cerevisiae* homologue of LlRpf2, Rpf2. Rpf2 is an essential gene in yeast involved in ribosome biogenesis (Zuniga et al., 1999, Wehner et al., 2002). Studies with Rpf2 indicate that it is a nucleolar RNA binding protein required for maturation, processing and assembly of the large ribosomal subunit rRNAs. Based on these data, it was suggested that Rpf2 is required to coordinately regulate assembly of the large ribosomal subunit. Rpf2 protein with other four yeast essential gene products Rpf1 (Wehner et al., 2002), Imp4 (Lee and Baserga, 1999), Ssf1 (Fatica et al., 2002) and Brx1 (Eisenhaber et al, 2001) constitute Imp4 superfamily which possess a common RNA binding motif similar to DNA binding motif of the *E.coli* σ70 transcription factor and plays a central role in many different steps in the production of ribosomes.

Intracellular localization studies of LlRpf2 protein revealed that LlRpf2 localizes to the nucleus with higher accumulation in the nucleolus, indicating LlRpf2, like Rpf2, is a nucleolar protein. The relatively high degree of sequence similarity between LlRpf2 and

Rpf2, the presence of conserved RNA binding motif and the identical intracellular localization suggest a functional similarity between LlRpf2 and Rpf2.

Ribosome biogenesis is a complex and tightly regulated process. Although the steps and mechanisms involved in processing pre-rRNA have been extensively investigated, the overall regulation and coordination between the processing and assembly of preribosomal particles is not well understood. The involvement of a large number of cellular components and of multiple steps provides numerous opportunities for coordination and regulation.

Antisense expression of AtRpf2 is resulted in dwarf phenotype and growth retardation in Arabidopsis. Dwarf phenotype has been exhibited in a wide variety of mutant Arabidopsis plants. Mutants defective in brassinosteroid biosynthesis or perception display severe dwarfism (Mussig et al., 2001). One such mutant is *bin-5*, which is defective in AtSPO11-3. Unlike all the other SPO11 homologs that are involved in meiosis, BIN5 /AtSPO11-3 plays an essential role during somatic development (Yin et al., 2002). Gibberellin (GA) response mutants also comprises dwarf mutants: *gai* (GA insensitive) (Peng et al., 1997), *sly* (sleepy) (Ogas et al., 1997) and *pkl* (pickle) (Steber et al., 1998). Also in Arabidopsis mutant defective in plastid ribosomal protein PRPL11, the overall size and growth rate reduction was observed (Pesaresi et al., 2001).

Mutations in genes encoding the structural components of ribosomes have been described in various organisms and it causes growth arrest and delay. In addition to this, mutations in ribosomal protein result in defects in the development of the organism. In *Drosophila Minute* mutant, defective in ribosomal protein *RPS5* gene, is characterized by semi-dominant phenotypes, such as delay in larval development, smaller body size and recessive embryo lethality (Lambertsson, 1998). In the *Arabidopsis Minute like -1* (*aml-1*) mutant, growth retardation, floral and vascular defects, and the recessive embryo lethality were observed; analogous to defects observed in the *Drosophila Minute* mutants (Weijers et al., 2001).

Although, the fertility was not affected in antisense AtRpf2 plants, the observed dwarf phenotype in several independent lines, might be caused by the defect in

ribosome biogenesis due to the reduced levels of the AtRpf2 protein. Because AtRpf2 is a single gene in Arabidopsis, it is possible that the AtRpf2 is involved in ribosome biogenesis in both vegetative and reproductive organs. On the other hand, lily LlRpf2 is specifically expressed and play a role in restricted tissues.

The results presented in this work and the available evidences suggest that the LlRpf2 protein may be involved in ribosome biogenesis during early stages of microsporogenesis or meiosis in lily.

Acknowledgements

I would like to express my gratitude to Prof. Kazuyuki Hiratsuka, Yokohama National University, for supervising, encouragement throughout this study and critical reading of the manuscript. I would like to thank Prof. Takashi Hashimoto and Dr. Hisabumi Takase for assistance and support. I wish to thank all the members of Laboratory of Plant Biotechnology, Yokohama National University and Laboratory of Plant Molecular and Cellular Biology, Nara Institute of Science and Technology. And I express my gratitude to the Ministry of Education, Culture, Sports, Science and Technology of Japan for supporting and giving me the opportunity to study in Japan.

References

Bai, X., Peirson, B.N., Dong, F., Xue, C., and Makaroff C.A. (1999). Isolation and characterization of *SYN1*, a *RAD21*-like gene essential in meiosis in *Arabidopsis*. Plant Cell 11, 417-430

Caryl, A.P., Armstrong, S.J., Jones, G.H., and Franklin, F.C. (2000). A homologue of the yeast *HOP1* gene is inactivated in the *Arabidopsis* meiotic mutant *asy1*. Chromosoma 109, 62-71

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16, 735-743

Couteau, F., Belzile, F., Horlow, C., Grandjean, O., Vezon, D., and Doutriaux, M.P. (1999). Random chromosome segregation without meiotic arrest in both male and female meiocytes of a *dmc1* mutant of *Arabidopsis*. Plant Cell 11, 1623-1634

Erickson, R.O. (1948). Cytological and growth correlations in the flower bud and anther of *Lilium longiflorum*. Am. J. Bot. 35, 729-739

Eisenhaber, F., Wechselberger, C., and Kreil, G. (2001). The Brix domain protein family-a key to the ribosomal biogenesis pathway? Trends Biochem. Sci. 26, 345-347

Fatica, A., Cronshaw, A.D., Dlakic, M., and Tollervey, D. (2002). Ssf1p prevents premature processing of an early pre-60S ribosomal particle. Mol. Cell 9, 341-351

Hihara, Y., Shoda, K., Liu, Q., Hara, C., Umeda, M., Toriyama, K., and Uchiyama, H. (1997). Expressed sequence tags in developing anthers of rice (*Oryza sativa* L.). Plant Biotechnol. 14, 71-75

Höfte, H., Desprez, T., Amselem, J., Chiapello, H., Caboche, M., Moisan, A., Jourjon, M.-F., Charpenteau, J.-L., Berthomieu, P., Guerrier, D., Giraudat, J., Quigley, F., Thomas, F., Yu, D.-Y., Mache, R., Raynal, M., Cooke, R., Grellet, F., Delseny, M., Parmentier, Y., de Marcillac, G., Gigot, C., Fleck, J., Philipps, G., Axelos, M., Bardet, C., Tremousaygue, D., and Lescure, B. (1993). An inventory of 1152 expressed sequence tags obtained by partial sequencing of cDNA from *Arabidopsis thaliana*. Plant J. 4, 1051-1061

Keith, S.C., Hoang, D.O., Barreth, B.M., Feigelman, B., Nelson, M.C., Thai, H., and Baysdorfer, C. (1993). Partial sequence analysis of 130 randomly selected maize cDNA clones. Plant Physiol. 101, 329-332

Knop, M., Pereira, G., Geissler, S., Grein, K., and Schiebel, E. (1997). The spindle pole body component Spc97p interacts with the gamma-tubulin of *Saccharomyces cerevisiae* and functions in microtubule organization and spindle pole body duplication. EMBO J. 16, 1550-1564

Kobayashi, T., Hotta, Y., and Tabata, S. (1993). Isolation and characterization of a yeast gene that is homologous with a meiosis-specific cDNA from a plant. Mol. Gen. Genet. 237, 225-232

Kobayashi, T., Kobayashi, E., Sato, S., Hotta, Y., Miyajima, N., Tanaka, A., and Tabata, S. (1994). Characterization of cDNAs induced in meiotic prophase in lily microsprocytes. DNA Res. 1, 15-26

Koltunow, A.M., Truettner, J., Cox, K.H., Wallroth, M., and Goldberg, R.B. (1990). Different temporal and spatial gene expression occur during anther development. Plant Cell 2, 1201-1224

Kuhlemeier, C., Fluhr, R., and Chua, N.-H. (1988). Upstream sequences determine the differences in transcript abundance of pea rbcS-3A genes. Mol. Gen. Genet. 212, 405-411

Lambertsson, A. (1998). The minute genes in Drosophila and their molecular functions. Adv. Genet. 38, 69-134

Lee, S.J., and Baserga, S.J. (1999). Imp3p and Imp4p, two specific components of the U3 small nucleolar ribonucleoprotein that are essential for pre-18S rRNA processing. Mol. Cell. Biol. 19, 5441-5452

Lydall, D., Nikolsky, Y., Bishop, D.K., and Weinert, T. (1996). A meiotic recombination checkpoint controlled by mitotic checkpoint genes. Nature 383, 840-843

Morohashi, K., Takase, H., Hotta, Y., and Hiratsuka, K.(2000). Large-scale sequencing of meiosis-associated genes from a cDNA library of lily microsporocytes. Plant Biotechnol. 17, 131-135

Murphy, S.M., Preble, A.M., Patel, U.K., O'Connell, K.L., Dias, D.P., Moritz, M., Agard, D.M., Stults, J.T., and Stearns, T. (2001). GCP5 and GCP6: Two members of the human γ-tubulin complex. Mol. Biol. Cell 12, 3340-3352

Mussig, C., and Altman, T. (2001). Brassinosteroid signaling in plants. Trends Endo. Meta. 12, 398-402

Nelissen, R.L., Will, C.L., van Venrooij, W.J., and Luhrmann, R. (1994). The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. EMBO J. 13, 4113-4125

Newman, T., de Bruijn, F. J., Green, P., Keegstra, K., Kende, H., McIntosh, L.,

Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M., Retzel, E., and Somerville, C. (1994). Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones. Plant Physiol. 106, 1241-1255

Nishihama, R., Ishikawa, M., Araki, S., Soyano, T., Asada, T., and Machida, Y. (2001). The NPK1 mitogen-activated protein kinase kinase kinase is a regulator of cell-plate formation in plant cytokinesis. Genes Dev. 15, 352-363

Ogas, J., Cheng,J.-C., Sung, Z.R., and Somerville, C (1997). Cellular differentiation regulated by gibberellin in the *Arabidopsis thaliana pickle* mutant. Science 277, 91-94

Ogata, S., Takase, H., Hiratsuka, K., and Hotta, Y. (1999). Mutational analysis of the signal for a nuclear localization of proteins which accumulate specifically during meiosis in lily microsporocytes. Plant Cell Rep. 19, 1-5

Oldenhof, M.T., de Groot P.F., Visser, J.H., Schrauwen, J.A., and Wullens, G.J. (1996). Isolation and characterization of a microspore-specific gene from tobacco. Plant Mol. Biol. 31, 213-225

Park, Y. S., Kwak, J. M., Kwon, O-Y., Kim, Y. S., Lee, D. S., Cho, M. J., Lee, H. H., and Nam, H. G. (1993). Generation of expressed sequence tags of random root cDNA clones of *Brassica napus* by single-run partial sequencing. Plant Physiol. 103, 359-370

Sasaki, T., Song, J., Koga-Ban, Y., Matsui, E., Fang, F., Higo, H., Nagasaki, H., Hori, M., Miya, M., Murayama-Kayano, E., Takiguchi, T., Takasuga, A., Niki, T., Ishimaru, K., Ikeda, H., Yamamoto, Y., Mukai, Y., Ohta, I., Miyadera, N., Havukkala, I., and Minobe, Y. (1994). Toward cataloguing all rice genes: large-scale sequencing of randomly chosen rice cDNA from a callus cDNA library. Plant J. 6, 615-624

Peng, J., Carol, P., Richard, D.E., King, K.E., Cowling, R.J., Murphy, G.P., and Harberd, N.P. (1997). The *Arabidopsis GAI* gene defines a signaling pathway that negatively regulates gibberellin responses. Genes Dev. 11, 3194-3205

Pesaresi, P., Varotto, C., Muerer, P., Jahns, P., Salamini, F., and Leister, D. (2001). Knock-out of plastid ribosomal protein L11 in Arabidopsis: effects on mRNA translation and photosynthesis. Plant J. 27, 179-189

San-Segundo, P.A., and Roeder G.S. (1999). Pch2 links chromatin silencing to meiotic checkpoint control. Cell 97, 313-324

Schiefthaler, U., Balasubramanian, S., Sieber, P., Chevalier, D., Wisman, E., and Schneitz, K. (1999). Molecular analysis of *NOZZLE*, a gene involved in pattern formation and early sporogenesis during sex organ development in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. USA. 96, 11664–11669

Shou, W., Seol, J.H., Shevchenko, A., Baskerville, C., Moazad, D., Chen, Z.W., Jang, J., Schevchenko, A., Charbonneau, H., and Deshaies, R.J. (1999). Exit from mitosis is regulated by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. Cell 976, 233-244

Steber, C.M., Cooney, S.E., and McCourt, P. (1998). Isolation of the GA-response mutant *sly1* as a suppressor of *ABI1-1* in *Arabidopsis thaliana*. Genetics 149, 509-521

Terasawa, M., Shinohara, A., Hotta, Y., Ogawa, H., and Ogawa, T. (1995). Localization of RecA-like recombination proteins on chromosomes of the lily at various meiotic stages. Genes Dev. 9, 925-934

Uchimiya, H., Kidou, S., Shimazaki, T., Aotsuka, S., Takamatsu, S., Nishi, R., Hashimoto, H., Matsubayashi, Y., Kidou, N., Umeda, M., and Kato, A. (1992). Random

sequencing of cDNA libraries reveals a variety of expressed genes in cultured cells of rice (*Oryza sativa* L.). Plant J. 2, 1005-1009

Venema, J., and Tollervey, D. (1999). Ribosome synthesis in *Saccharomyces cerevisiae*. Annu. Rev. Genet. 33, 261-311

Wehner KA, Baserga SJ (2002) The σ70-like motif: a eukaryotic RNA binding domain unique to a superfamily of proteins required for ribosome biogenesis. Mol. cell 9, 329-339

Weijers, D., Dijk, M.F., Vencken R.-J., Quint, A., Hooykaas, P., and Offringa, R. (2001). An *Arabidopsis Minute*-like phenotype caused by a semi-dominant mutation in a *RIBOSOMAL PROTEIN S5* gene. Development 128, 4289-4299

Young, W-C., Ye, D., Xu, J., and Sundaresan, V. (1999). The *SPOROCYTELESS* gene of *Arabidopsis* is required for initiation of sporogenesis and encodes a novel nuclear protein. Genes Dev. 13, 2108-2117

Yin, Y., Cheong, H., Friedrichsen, D., Zhao, Y., Hu, J., Mora-Garcia, S., and Chory, J. (2002). A crucial role for the putative Arabidopsis topoisomerase VI in plant growth and development. Proc. Nat. Acad. Sci. 99, 10191-10196

Zuniga , S., Boscovic, J., Garcia-Cantalejo, J.M., Jimenez, A., Ballesta, J.P.G., and Remacha, M. (1999). Deletion of 24 open reading frames from chromosome XI from *Saccharomyces cerevisiae* and phenotypic analysis of the deletants. Gene 233, 141-150