1995

Cloning and characterization of two floral-specific genes of sunflower (Helianthus annuus L.).

Joggari T. Reddy
University of Windsor

National Library of Canada

Acquisitions and Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada

Direction des acquisitions et des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Cloning and characterization of two floral-specific genes of sunflower (*Helianthus annuus* L.)

by
J.T. Reddy

A Dissertation Submitted to
the Faculty of Graduate Studies and Research
Through the Department of Biological Sciences
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
at the University of Windsor

Windsor, Ontario, Canada
1995

ISBN 0-612-01482-7

Name JAGGARI T REDDY

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

MOLECULAR BIOLOGY — SUBJECT TERM

0307 — SUBJECT CODE

U·M·I

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
- Architecture 0729
- Art History 0377
- Cinema 0900
- Dance 0378
- Fine Arts 0357
- Information Science 0723
- Journalism 0391
- Library Science 0399
- Mass Communications 0708
- Music 0413
- Speech Communication 0459
- Theater 0465

EDUCATION
- General 0515
- Administration 0514
- Adult and Continuing 0516
- Agricultural 0517
- Art 0273
- Bilingual and Multicultural 0282
- Business 0688
- Community College 0275
- Curriculum and Instruction 0727
- Early Childhood 0518
- Elementary 0524
- Finance 0277
- Guidance and Counseling 0519
- Health 0680
- Higher 0745
- History of 0520
- Home Economics 0278
- Industrial 0521
- Language and Literature 0279
- Mathematics 0280
- Music 0522
- Philosophy of 0998
- Physical 0523
- Psychology 0525
- Reading 0535
- Religious 0527
- Sciences 0714
- Secondary 0533
- Social Sciences 0534
- Sociology of 0340
- Special 0529
- Teacher Training 0530
- Technology 0710
- Tests and Measurements 0288
- Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
- Language
  - General 0679
  - Ancient 0289
  - Linguistics 0290
  - Modern 0291
- Literature
  - General 0401
  - Classical 0294
  - Comparative 0295
  - Medieval 0297
  - Modern 0298
  - African 0316
  - American 0591
  - Asian 0305
  - Canadian (English) 0352
  - Canadian (French) 0355
  - English 0593
  - Germanic 0311
  - Latin American 0312
  - Middle Eastern 0315
  - Romance 0313
  - Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
- Philosophy 0422
- Religion
  - General 0318
  - Biblical Studies 0321
  - Clergy 0319
  - History of 0320
  - Philosophy of 0322
- Theology 0469

SOCIAL SCIENCES
- American Studies 0323
- Anthropology
  - Archaeology 0324
  - Cultural 0326
  - Physical 0327
- Business Administration
  - General 0310
  - Accounting 0272
  - Banking 0770
  - Management 0454
  - Marketing 0338
- Canadian Studies 0385
- Economics
  - General 0501
  - Agricultural 0503
  - Commerce-Business 0505
  - Finance 0508
  - History 0509
  - Labor 0510
  - Theory 0511
- Folklore 0358
- Geography 0366
- Gerontology 0351
- History
  - General 0578
  - Ancient 0579
  - Medieval 0581
  - Modern 0582
  - Black 0328
  - African 0331
  - Asia, Australia and Oceania 0332
  - Canadian 0334
  - European 0335
  - Latin American 0336
  - Middle Eastern 0333
  - United States 0337
- History of Science 0585
- Law 0398
- Political Science
  - General 0615
  - International Law and Relations 0616
  - Public Administration 0617
- Recreation 0814
- Social Work 0452
- Sociology
  - General 0626
  - Criminology and Penology 0627
  - Demography 0938
  - Ethnic and Racial Studies 0631
  - Individual and Family Studies 0628
  - Industrial and Labor Relations 0629
  - Public and Social Welfare 0630
  - Social Structure and Development 0700
  - Theory and Methods 0344
- Transportation 0709
- Urban and Regional Planning 0999
- Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
- Agriculture
  - General 0473
  - Agronomy 0285
  - Animal Culture and Nutrition 0475
  - Animal Pathology 0476
  - Food Science and Technology 0359
  - Forestry and Wildlife 0478
  - Plant Culture 0479
  - Plant Pathology 0480
  - Plant Physiology 0817
  - Range Management 0777
  - Wood Technology 0746
- Biology
  - General 0306
  - Anatomy 0287
  - Biostatistics 0308
  - Botany 0309
  - Cell 0379
  - Ecology 0329
  - Entomology 0353
  - Genetics 0369
  - Limnology 0793
  - Microbiology 0410
  - Molecular 0307
  - Neuroscience 0317
  - Oceanography 0416
  - Physiology 0433
  - Radiation 0821
  - Veterinary Science 0778
  - Zoology 0472
- Biophysics
  - General 0786
  - Medical 0760

EARTH SCIENCES
- Biogeochemistry 0425
- Geochemistry 0996
- Geodesy 0370
- Geology 0372
- Geophysics 0373
- Hydrology 0388
- Mineralogy 0411
- Paleobotany 0345
- Paleoecology 0426
- Paleontology 0418
- Paleozoology 0985
- Palynology 0427
- Physical Geography 0368
- Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
- Environmental Sciences 0768
- Health Sciences
  - General 0566
  - Audiology 0300
  - Chemotherapy 0992
  - Dentistry 0567
  - Education 0350
  - Hospital Management 0769
  - Human Development 0758
  - Immunology 0982
  - Medicine and Surgery 0564
  - Mental Health 0347
  - Nursing 0569
  - Nutrition 0570
  - Obstetrics and Gynecology 0380
  - Occupational Health and Therapy 0354
  - Ophthalmology 0381
  - Pathology 0571
  - Pharmacology 0419
  - Pharmacy 0572
  - Physical Therapy 0382
  - Public Health 0573
  - Radiology 0574
  - Recreation 0575
  - Speech Pathology 0460
  - Toxicology 0383
- Home Economics 0386

PHYSICAL SCIENCES

Pure Sciences
- Chemistry
  - General 0485
  - Agricultural 0749
  - Analytical 0486
  - Biochemistry 0487
  - Inorganic 0488
  - Nuclear 0738
  - Organic 0490
  - Pharmaceutical 0491
  - Physical 0494
  - Polymer 0495
  - Radiation 0754
- Mathematics 0405
- Physics
  - General 0605
  - Acoustics 0986
  - Astronomy and Astrophysics 0606
  - Atmospheric Science 0608
  - Atomic 0748
  - Electronics and Electricity 0607
  - Elementary Particles and High Energy 0798
  - Fluid and Plasma 0759
  - Molecular 0609
  - Nuclear 0610
  - Optics 0752
  - Radiation 0756
  - Solid State 0611
- Statistics 0463

Applied Sciences
- Applied Mechanics 0346
- Computer Science 0984
- Engineering
  - General 0537
  - Aerospace 0538
  - Agricultural 0539
  - Automotive 0540
  - Biomedical 0541
  - Chemical 0542
  - Civil 0543
  - Electronics and Electrical 0544
  - Heat and Thermodynamics 0348
  - Hydraulic 0545
  - Industrial 0546
  - Marine 0547
  - Materials Science 0794
  - Mechanical 0548
  - Metallurgy 0743
  - Mining 0551
  - Nuclear 0552
  - Packaging 0549
  - Petroleum 0765
  - Sanitary and Municipal 0554
  - System Science 0790
- Geotechnology 0428
- Operations Research 0796
- Plastics Technology 0795
- Textile Technology 0994

PSYCHOLOGY
- General 0621
- Behavioral 0384
- Clinical 0622
- Developmental 0620
- Experimental 0623
- Industrial 0624
- Personality 0625
- Physiological 0989
- Psychobiology 0349
- Psychometrics 0632
- Social 0451

ABSTRACT

Since most developmental processes in plants are post-embryonic, the flower is an attractive model system to study plant developmental events. Sunflower (*Helianthus annuus*) being an important oil seed crop was chosen for this study, because genes isolated and characterized from sunflower will shed more light on plant development, and can also be useful in biotechnological applications (such as engineering artificial male sterile lines for hybrid seed production). The study involved the isolation and characterization of two sunflower floral-specific genes: SF15 (an anther-specific gene) and SF17 (a pollen-specific gene).

SF15: Northern blot analysis (with a 526 bp putative floral cDNA probe) revealed that the probe hybridized to 1.0 kb mRNA transcripts present in the young unopened disc florets, whose development was delineated into seven stages (1,2,3,4,5,6 and 7). SF15 transcripts begin to accumulate in stage 1 and reach higher levels in stages 2,3 and 4, while no expression was observed in late developmental stages 5 to 7, and the expression is confined only to the anthers of male-fertile disc florets. *In situ* hybridization experiments revealed that SF15 is expressed in a single layer of epidermal cells of the anther. Screening of the floral cDNA library resulted in the isolation of 4 near full-length cDNA clones (SF15-2, SF15-3, SF15-45 and SF15-49). These cDNA clones have inserts of 785 to 934 bp. Screening of a genomic library resulted in the isolation of a 10.7 kb genomic clone (SF15G). Primer extension analysis revealed that there are two major and, few minor transcription initiation sites. Southern blot analysis showed that SF15 is present as a member of a multigene family in the sunflower nuclear genome and as a single copy in the nuclear genomes of corn and tomato. The DNA sequence information revealed that the genomic and four cDNA clones represent five distinct cognate genes, with nucleotide similarities ranging from 89.9% to 99.7% in the coding regions. The deduced amino acid sequence of the proteins encoded by these transcripts have 81.9% to 99.6% identity. The open reading frame of genomic clone SF15G, encodes a putative polypeptide of 280 amino acids with a molecular mass of 31.7 kDa. SF15 proteins, contain the N-terminal transit signal peptides indicating that these proteins enter the secretory

pathway. Interestingly, only the SF15G protein contains a putative vacuolar targeting sequence, and not the proteins encoded by four cDNA clones. SF15 multigene family encodes both vacuolar and non-vacuolar forms of proteins. The role of SF15 gene in the anther dehiscence, needs to be confirmed.

SF17: In the Northern blot analysis the 329 bp probe, hybridized to the 2.0 kb transcripts present exclusively in the mature pollen grains. Screening of phase III floral cDNA library resulted in the isolation of 1915 bp cDNA clone. A 15 kb genomic clone (SF17G) was isolated from the genomic library, with the 1915 bp cDNA insert as a probe. A 5.5 kb EcoRI fragment of the genomic clone was subcloned into pUC19 vector for sequence analysis. The DNA sequence information revealed that the cDNA and genomic clones are identical, except for the presence of two small introns in the genomic clone. Primer extension analysis showed that there are two major and a few minor transcription initiation sites. Southern blot analysis showed that SF17 is represented in one or two copies, in the sunflower nuclear genome. The cDNA encodes a leucine-rich-protein of 540 amino acid residues with a molecular mass of 61.7 kDa. SF17 protein did not contain a transit signal sequence, but the SF17 protein consists of five structural features: a hydrophilic domain, a leucine-rich-region (LRR), an acid domain and two trans-membrane domains. It is postulated that SF17 protein resides in a membrane with two membrane-spanning segments and its N-terminal side of LRR being disposed in the cytoplasm. The LRR region of SF17 protein shows a striking resemblance to the LRR domains of a variety of proteins from different organisms. During pollen grain germination there is a rapid initiation of protein synthesis and a shift in the increase of polyribosomes and a decrease in single ribosomes. LRR domains of various proteins have been implicated in mediating tight protein-protein interactions. The LRR domain of SF17 protein shows a strong homology to the rat ribosome binding protein p34 located on the rough ER. Whether the SF17 protein (like p34) mediates the association of ribosomes to the rough ER during pollen germination needs to be demonstrated.

Dedicated
to

My Parents

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my supervisor Dr. D. T. N. Pillay for his constant support, excellent guidance and encouragement during the course of this study. I am also profoundly thankful to him for allowing me to take novel approaches and experiment with new ideas. I also wish to thank him for his moral support and confidence in me during the graduate work. I am immensely benefited by my close interaction with him during fast few years.

Thanks are also due to my dissertation committee members, Drs. K. A. Adeli, D. Cotter and D. Thomas, for their support, guidance and encouragement. I am thankful to Dr. R. J. Cedergren, University of Montreal, for serving as the external examiner and to Dr. T. K. Christopolous for representing Dr. K. Adeli, in his absence. I gratefully acknowledge Dr. N. Dudareva, Research associate in our laboratory for many helpful discussions, suggestions and software analysis.

I also thank our collaborators Drs. A. Steinmetz, R. Krauter and J. L. Evrard (Institute of Plant Molecular Biology, Strasbourg, France), for providing probes and help in *in situ* hybridization. I am thankful to the Faculty of Graduate studies and Research for financial support and Dr. L. K. Smedick, for chairing the defense examination.

Financial assistance from NSERC of Canada (Grant No: A-1984 to Dr. D. T. N. Pillay) is gratefully acknowledged. I would like to thank my parents, members of my family and friends for their moral support.

TABLE OF CONTENTS

SF15

RESULTS

DISCUSSION

SF17

RESULTS

DISCUSSION

LIST OF FIGURES & TABLES

SF15

Figure **Page**

LIST OF FIGURES & TABLES

SF15

Table **Page**

SF17

Figure **Page**

LIST OF ABBREVIATIONS

ATP	adenosine triphosphate
cDNA	complimentary DNA
CMS	cytoplasmic male sterility
CsCl	cesium chloride
dCTP	deoxycytidine 5'-triphosphate
DNA	deoxy ribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphates
DTT	dithiothreitol
EDTA	disodum ethylenediamine tetraacetate
Kb	kilobase pair
kDa	kilo dalton
LRR	leucine rich repeats
mRNA	messenger RNA
ORF	open reading frame
PCR	polymerase chain reaction
PEG	polyethylene glycol
pfu	plaque forming unit
RNA	ribonucleic acid
SDS	sodium dodecyl sulphate
TE	Tris-EDTA
tRNA	transfer RNA

INTRODUCTION

1. Plant reproduction and flowering

From zygote to zygote the life cycle of a flowering plant involves complex morphogenetic events leading to the production of both vegetative organ systems and reproductive organ systems. The major feature of the plant life cycle is alternation of generations. In angiosperms the life cycle alternates between a diploid sporophyte (reproducing through spores), and a haploid gametophyte (reproducing through gametes).

Higher plants have an indeterminate program of growth and development (Raven et al., 1986; Goldberg, 1988). In plants, embryogenesis specifies the shoot and root meristems, and these meristems are packaged in the dormant seed. But most of the morphogenetic events in higher plants occur post-embryonically, from the continuous division and differentiation of the shoot and root meristematic cells. Meristematic cells are a group of actively dividing cells which are analogous to stem cell populations of humans. Upon germination of the seed, the apical meristems undergo division and differentiation and give rise to a mature plant consisting of root, stem, leaf and flower. The various organ systems of the flower are derived from the cells originally present within the vegetative shoot meristem (Steeves and Sussex, 1989). Because no separate germ line is set aside during plant embroyogeneis, sexual reproduction in angiosperms involves the conversion of vegetative meristems into floral meristems.

Before the transition to the reproductive phase, vegetative meristems differentiate and produce leaves, internodes and auxiliary

buds. Following transition to reproductive growth, floral bracts, inflorescence branches and flowers are produced (Huala and Sussex, 1983). In angiosperms, reproductive processes take place in the flower. A typical hermaphrodite flower consists of four organ systems arranged in four whorls. Whorl 1 consists of sepals and whorl 2 contains the petals. Whorls 3 and 4 contain stamens and pistils respectively. Results of genetic and molecular investigations on floral development in *Arabidopsis* and *Antirrhinum* have shown that the number and identity of these floral organ systems is controlled by the combinatorial action of floral homeotic genes (reviewed by Coen and Carpenter,1993; Weigel and Meyerowitz, 1994).

The initiation of reproductive process involves the perception of appropriate environmental stimuli by the plant. The environmental stimuli that potentially control floral initiation are perceived by different parts of the plant. Thus, the apical meristems should integrate one or more long-range signals when 'deciding', whether to switch over from the vegetative to the reproductive phase of development. Physiological studies in *Sinapis alba*, and genetic studies in *Arabidopsis thaliana*, show that molecules such as carbohydrates, auxins, gibberellins, cytokinins, and calcium can control the transition to the reproductive state (Bernier et al; 1993).

In recent years many events of higher plant reproduction have been studied at the physiological, biochemical, genetic and morphological levels. These studies reveal new insights into such processes as the transition to flowering, initiation of floral primordia and floral organs, the development of flowers, sporogensis, gametogenesis and gamete development, fertilization and embryogenesis (Chasan and Walbot, 1993).

2. Anther and pollen development

The major role of a flower is the production of male and female gametes. In flowering plants, male reproductive processes take place within the stamen (Esau, 1977). Stamens are differentiated from the primordia that are specialized within the floral meristem, following transition from a vegetative to a flowering pathway (Esau, 1977; Raven et al., 1986). The stamen consists of diploid meiocytes which undergo meiosis to produce haploid microspores. Haploid microspores undergo mitotic divisions and produce tricellular male gametophytes (pollen grains) containing the male gametes (sperm nuclei). The stamen consists of two morphologically distinct parts, the filament and an anther. The filament contains vascular tissue (serving as a conduit for the transport of water and nutrients) and anchors the stamen at its base. The anther contains both reproductive and nonreproductive tissues. The role of the anther is the production and successful release of pollen grains (Goldberg et al., 1993).

The developmental events leading to anther formation and pollen release have been studied in detail in tobacco (Goldberg et al., 1993). Anther development has been divided into two general phases. During phase 1, the morphology of the anther is established, cell and tissue differentiation occur, and microsporocytes undergo meiosis. At the end of phase 1, the anther contains most of its specialized cells and tissues, and tetrads of microspores are present within the pollen sacs. During phase 2, pollen grains differentiate, the anther enlarges and is pushed upward in the flower by filament extension, anther tissues degenerate, and the pollen grains are released after dehiscence.

Stamen initiation is controlled by floral homeotic gene interactions, and its development is influenced by nuclear and

mitochondrial genome ineteractions (Goldberg et al., 1993). The gene expression program is temporally and spatially regulated during anther differentiation and dehiscence. RNA-excess DNA/RNA hybridization experiments with single copy DNA, indicated that there are about 25,000 diverse genes expressed in stage 6 tobacco anther (Kamalay and Goldberg, 1980). The vast majority of these genes encode rare mRNA transcripts that constitute less than 0.001% of the mRNA mass when averaged over the entire anther mRNA population. Comparison of RNA populations of floral and vegetative organ systems showed that about 10,000 diverse anther mRNAs are undetectable in the nuclear RNA and mRNA populations of other organs and are anther specific (Kamalay and Goldberg, 1980, 1984). Several laboratories have identified cDNA clones representing cognate mRNAs expressed exclusively, or at elevated levels, in the anther (Koltunow et al., 1990; Smith et al., 1990; Evrard et al., 1991; McCormick et al., 1991, 1993; Nacken et al., 1991; Scott et al., 1991a, b; Shen abd Hsu, 1992). Tobacco anther-specific mRNAs have been shown to encode lipid transfer proteins, protease inhibitors, thiol endopeptidases, and glycine and proline-rich polypeptides with properties of cell wall proteins (Goldberg et al., 1993).

Pollen development offers a unique opportunity to study cell fate, cell patterning, cell polarity and cell signaling. Various aspects of pollen development have been reviewed (Bedinger, 1992; Mascarenhas, 1989, 1990, 1993; and McCormick, 1991,1993). Diploid sporogenous cells known as microsporocytes, are the progenitors of pollen grains. Meiosis of microsporocytes, leads to the formation of tetrads of microspores. The four products of meiosis are surrounded by a callose layer. The individual microspores of the tetrad are released into the anther locule by the action of β-1,3-glucanase (a callase enzyme produced by the tapetum). The uninucleate microspores undergo asymmetric mitotic division, resulting in a pollen grain with two cells: a larger vegetative and a smaller generative cell.

The generative cell is enclosed entirely within the vegetative cell. In about 70% of plant families (e.g., Solanaceae and Liliaceae), the pollen grain is released from the anther when it consists of just those two cells. The subsequent mitotic division of the generative cell occurs during pollen tube growth through the female stylar tissue, and gives rise to two sperm cells. In other plant families (e.g., Cruciferae and Graminae), this second mitotic division occurs before the pollen is shed from the anther (McCormick, 1993).

The function of the mature pollen grain is to effect the delivery of its two sperm cells into the female embryo sac (macrogametophyte) in the ovule. In order to accomplish the delivery, the pollen grain must germinate on the receptive stigma, quickly sending a pollen tube out through the pore of the pollen wall and down through the stylar tissue. The germination and rapid growth requires a considerable amount of cell wall synthesis and possibly breakdown of stylar components (Pressey and Reger, 1989). Many of the pollen-specific genes reported to date appear to have homology to cell wall degradation enzymes such as pectin esterase (Albani et al., 1991), pectate lyase (Rogers et al., 1992; Wing et al., 1989), and polygalacturonase (Allen and Lonsdale, 1993; Brown and Crouch, 1990; Niogret et al., 1991). About 20,000 genes are expressed during pollen development, of which 10-20% are thought to be pollen specific (Mascarenhas, 1990). Depending on their expression pattern these genes are divided into two groups: the early genes and late genes. The early genes are expressed soon after the completion of meiosis and include rRNA genes, tRNA genes and a set of other protein genes (actin, alcohol dehydrogenase and β-galactosidase). The late genes become active around the time of microspore mitosis and remain active until anthesis.

3. Inflorescence and florets of *Helianthus annuus* L.

Sunflower (*Helianthus annuus* L.) has a special type of racemose inflorescence, called the head or capitulum (Fig. 2). The capitulum is subtended by a green involucre of bracts. The capitulum consists of two types of florets (flowers), ray and disc. These florets are located on the top of the succulent and flattened receptacle. The ray florets which occupy the periphery are incomplete, and at maturity contain bright yellow-orange petals. They are sterile, play no direct role in reproduction, but serve as signals for bees and other insect pollinators (Knowles, 1978).

The florets over the remainder of the capitulum are called disc florets. Each disc floret is a hermaphrodite, and can develop into a seed. Individual disc florets are arranged in arcs (whorls) radiating from the center of capitulum. Each disc floret contains chaffy bract, pappus scales (modified sepals), an inferior ovary, and a tubular corolla of five petals, united except for the tips. Five anthers are united to form a tube, with separate filaments attached to the base of the corolla tube. Inside the anther tube is the style, terminating in a bilobed stigma. When the disc floret is fully developed, the style elongates and the bilobed stigma curls outward, the two lobes often making a full circle (Knowles, 1978)

Anthesis (opening) of the disc florets, proceeds sequentially and in a centripetal or acropetal direction (from the periphery towards the center). After the ray florets spread out from their folded position, the outer whorl of disc florets opens first. As a result, flower heads carry disc florets of all the developmental stages (Knowles, 1978).

4. Flower-specific genes of *Helianthus annuus* L.

The sunflower is an important oil seed crop. Detailed understanding of the molecular mechanisms regulating floral gene expression, will shed more light on sunflower reproduction. Moreover, the knowledge can be used advantageously in biotechnological applications (Everett et al., 1987) e.g. to engineer new types of male sterility by introducing cytotoxic genes whose expression is controlled by anther-specific regulatory elements (Mariani et al., 1990). Herdenberger et al.,(1990), isolated 31 incomplete floral cDNA clones, by differential screening of a cDNA library constructed using poly(A)$^+$ RNA from the sunflower capitulum (inflorescence) at anthesis.

Some of these incomplete cDNA clones were used as hybridization probes for the isolation of full-length cDNA clones and genomic clones. These floral cDNAs can be grouped into four categories based on the organ specificity of the respective clones (anther-, pollen-, pistil-, and corolla-specific). Seven cDNA clones (SF1, SF2, SF6, SF7, SF9, SF18 and SF19) are anther specific. Two cDNA clones SF3 and SF16 are expressed in the mature tricellular pollen grains. SF21 is pistil-specific and SF28 is a corolla-specific clone. The details of molecular aspects of these clones have been reviewed recently (Steinmetz et al., 1992 and 1993) and a brief summary of these findings is presented below.

Anther-specific cDNA clones: The seven anther-specific cDNA clones represent the cognate mRNA transcripts of three families (based on the homogy of DNA sequence information). The deduced amino acid sequences of these cDNA clones reveal that they encode proline-rich proteins with N-terminal signal peptides. The presence of a signal peptide indicates that they may be

extracellular proteins with a putative structural role (e.g.,toughening of the cell wall). DNA sequence analysis of SF2 and SF18 indicates that they originated by exon shuffling (Domon and Steinmetz, 1994).

Pollen-specific cDNA clones: SF3 and SF16 are expressed in the mature pollen grains. SF3 codes for a 219 amino acid long polypeptide with putative zinc finger domain and a C-terminal pentapeptide repeat (Baltz, 1992a). These fingers have been recently identified as two LIM domains (the term LIM is derived from the initials of the animal developmental regulatory proteins: **L**IN-11, **I**SL-1, **M** EC-3 in which this domain was first identified) (Baltz, 1992b). It is hypothesized that PLIM-1 (rechristened designation of SF3 i.e., **P**ollen-specific **LIM** protein 1) might act via its two LIM domains (DNA binding) as well as via it's C-terminal pentapeptide repeat (protein binding). SF3 protein could be a pollen-specific transcription factor, required for the expression of late pollen-specific genes. Late pollen-specific genes are implicated to encode proteins or enzymes needed for late pollen functions, such as pollen grain germination, pollen tube penetration and elongation through stylar tissue leading to fertilization. The other pollen-specific clone SF16 encodes a 331 amino acid long basic polypeptide with a pI value of 11.3 (Dudareva, 1994 a, b). Its basic nature suggests that it might be a nucleic-acid binding protein.

Pistil-specific cDNA clone: Dot blot hybridization experiments revealed that SF21 is expressed in styles and stigma but not in the ovary. SF21 encodes a 352 amino acid long polypeptide. The deduced protein of SF21, with many short hydrophobic domains, might be a putative membrane protein with unknown function (Steinmetz et al., 1992, 1993)

Petal-specific cDNA clone: SF28 is corolla specific and codes for cytochrome p450. It shows little homology with three described plant cytochromes of avocado, periwinkle and jerusalem artichoke.

5. Objectives of the project

Molecular cloning and characterization of new anther and pollen-specific genes will undoubtedly shed more light on the molecular and cellular mechanism of sunflower reproduction. This investigation was undertaken with two incomplete floral cDNA clones (SF15 and SF17) with the following objectives:

- Demonstration of spatial and temporal expression of the cDNA clones.
- Dissection of the development of capitulum and disc floret, to provide a reference point for future investigations, and for possible comparison of morphological or anatomical changes with temporal gene expression patterns.
- Isolation of full-length cDNA clones and genomic clones.
- Characterization of the cDNA and genomic clones and comparison of the sequence information with previously published clones.
- Demonstration of the cell and tissue specificity of gene expression.
- Mapping of the transcription initiation sites of the mRNA transcripts.
- Delineation of the copy number of the genes and any homology with genes of other crop species

MATERIALS AND METHODS

1. Plant material

Both male-fertile strain (HA401B) and male-sterile strain (HA401A) of sunflower (*Helianthus annuus* L.) supplied by CARGILL, France were used in the project. The plants were grown either in the greenhouse or in the field during summer. For heterologous hybridization experiments, corn (*Zea mays*), pea (*Pisum sativum*), safflower (*Carthamus tinctorius*), tomato (*Lycopersicon esculentum*,), cabbage (*Brassica oleracea capitata*) were used. The varieties of the above plants used were, Early Golden Bantam, Laxton's Progress, Lasting Orange, Ultrasweet and Paragon respectively.

2. Bacterial host strains

Escherichia coli strains NM522 (for transformation), ED8767 (for growing λCharon40 recombinant clones) and C600 hfl (for growing λgt10 recombinant clones) were used as hosts.

3. Ligation of DNA fragments

After digestion of DNA with appropriate enzyme(s), inserts were purified from low melting agarose gels, using Geneclean II kit (Bio/Can Scientific). DNA was eluted in 20 µl TE buffer. Ligations were performed at 16 ^{0}C overnight, with the T4 DNA ligation kit according to the protocol supplied by the vendor (US Biochemical). The ratio of vector to insert DNA was usually 1:2.

4. Transformation of bacterial Cells

The ligated DNA was transformed into competent NM 522 host cells according to the protocol of Hanahan (1983).

5. Isolation of Plasmid DNA

A single colony of recombinant clone was grown overnight in 500 ml of LB medium with ampicillin. Plasmid DNA was isolated according to the alkaline lysis method (Birnboim and Doly 1979). The DNA was gradient purified in OptiSeal polyallomer tubes (Beckman) by ultra centrifugation overnight at 55,000 rpm in VTi 65 rotor. The DNA band was collected and ethidium bromide was removed with water saturated butanol. The DNA solution was diluted with one volume of distilled water and precipitated with ethanol. DNA was pelleted, rinsed twice with 70% ethanol, vacuum dried and dissolved in TE buffer.

6. Construction of cDNA library

As the titer of the previous library (constructed by Herdenberger et al., 1990) was found to be very low, a new cDNA library was constructed. The cDNA library was constructed from the poly(A)$^+$ RNAs into the EcoR I restriction site of the λgt10 vector (an insertion vector). Disc florets of phase III capitulum (see below, the developmental staging of disc florets and capitulum, in the results section) were used for the isolation of total RNA. cDNA synthesis was performed according to the kit and protocol supplied by the vendor (Pharmacia). Synthesis of first strand cDNA was catalyzed by Moloney Murine Leukemia virus (MMLV) reverse transcriptase. Poly adenylated RNA and oligo (dT) were used as template and primer respectively. Second-strand cDNA synthesis was carried out by the procedure of Gubler and Hoffman (1983), in

which RNase H nicks the mRNA strand of the RNA: cDNA duplex formed after the synthesis of first strand cDNA. Following the second-strand synthesis, Klenow fragment was added to ensure that the ends of the cDNA duplex are blunt. The blunt ended, duplex cDNA was purified on a spun column, and ligated to the EcoRI/NotI adapters. The EcoRI-ended cDNA inserts (with internal NotI site) were ligated to λgt10 arms. The products of the ligation reaction (concatemers) were in vitro packaged, with Packagene in vitro packaging system (Promega), according to the vendor's protocol. The size of the preamplified cDNA library was 1.16 million clones. The original library was amplified (titer 6 x 10^6), aliquoted, stored at -80 ^{0}C and was used for further screenings.

7. Screening of cDNA library

To isolate full length cDNAs, 0.3 to 1.0 x 10^6 pfu (plaque forming units) were infected with C600 hfl bacteria (2 ml of overnight culture) and plated on 24.5 x 24.5 cm Nunc plates. After overnight growth at 37 ^{0}C, the plaques were transferred to Hybond-N$^+$ membrane (Amersham), as described by Benton and Davis (1977). The filters were subsequently hybridized in 25 ml hybridization solution (5x SSC, 5x Denhardt's, 0.5% SDS, with α-^{32} P-labeled probe at a concentration of 2 ng/ml) for 15 h at 65 ^{0}C. Filters were washed twice, at 65 ^{0}C in 2x SSC for 30 min. Pre-hybridizations, hybridizations and washings were done in a Turbo Speed hybridization oven (Bio/Can Scientific). Filters were dried and autoradiographed by exposing to Fuji RX X-ray film at -80 ^{0}C, using intensifying screens. Plaques showing strong hybridization signals were removed and purified via two to four further rounds of screening at low density.

8. Isolation of insert DNA from recombinant λ gt10

50 ml of LB medium (10 g/l bactotryptone, 5 g/l yeast extract, 10 g/l NaCl, 0.2% maltose and 10 mM $MgSO_4$) was inoculated with 200 µl of C600 hfl (restrictive host) and infected with recombinant phage from a single plaque. After overnight incubation at 37 ^{0}C under vigorous shaking, bacteria and bacterial debris were eliminated by centrifugation. Bacterial chromosomal DNA of the supernatant was digested by DNAse I (150 µg) for 1 h at 37 ^{0}C. Phage was precipitated by the addition of 2.5 g PEG (8000) and 3 g NaCl, at 4 ^{0}C for 4 hours or overnight. Following centrifugation, phage were suspended in 500 µl of TE, and phenol: chloroform extracted. After centrifugation the supernatant fluid was recovered and nucleic acids were precipitated with the addition of 0.1 vol. of 4 M NaCl and 2 vol. of ethanol. The precipitated nucleic acids were then pelleted by centrifugation, rinsed twice with 70% ethanol and dissolved in 200 µl of TE. Thirty microliters were used for insert analysis by restriction digestion .

9. Screening of genomic DNA library

Infection doses of 1- 2 x 10^6 pfu in ED8767 were plated on Nunc plates (24.5 x24.5 cm), grown overnight at 37 ^{0}C and transferred to Hybond-N+ membrane (Amersham), as described by Benton and Davis (1977). The filters were subsequently hybridized in 25 ml hybridization solution (5x SSC, 5x Denhardt's, 0.5% SDS, with α-^{32}P-labeled probe at a concentration of 2 ng/ml) for 15 h at 65 ^{0}C. Filters were washed twice at 65 ^{0}C in 2x SSC for 30 min. Pre-hybridizations, hybridizations and washings were done in a Turbo Speed hybridization oven (Bio/Can Scientific). Filters were dried and autoradiographed by exposing to Fuji RX X-ray film, at -80 ^{0}C, using intensifying screens. Plaques showing strong hybridization signals were removed and purified via four to six further rounds of screening at low density.

10. Isolation of insert DNA from recombinant λ Charon40

For large scale growth of phage λCharon 40 (Dunn and Blattner 1987) or its recombinant derivatives, 1 ml of overnight culture of ED8767 was infected with 2.0 x 10^6 pfu (15 min at 37 ^{0}C) in the presence of infection salts (10 mM $MgCl_2$, 10mM $CaCl_2$) and subsequently transferred into 1.0 liter of NAZM growth medium in a 2 l Erlenmeyer flask. After vigorous shaking at 37 ^{0}C for 17 h, 60 g NaCl, 5ml of chloroform and 125 µl DNase (10mg/ml) were added and shaking was continued (for 15 min) until all salts were dissolved. Un-lysed cells and cell debris were removed by centrifugation and supernatant was collected carefully without any traces of chloroform. Phage was precipitated overnight from the supernatant by the addition of 70g polyethylene glycol (PEG 8000, Sigma). The precipitated phage was pelleted by 20 min centrifugation at 5000 X g, and resuspended in 2 ml of lambda dilution buffer (10 mM Tris-HCl pH 7.4, 10 mM $MgSO_4$) and chloroform extracted. The phage was further purified by two runs on CsCl block gradients in Ultra-Clear centrifugation tubes, through ultra-centrifugation with a SW50.1 rotor (Davis et al., 1980). DNA was isolated from the purified recombinant phage using the formamide lysis method (Davis et al., 1980). The pellet was dissolved in an appropriate volume of TE buffer. Appropriate restriction fragment(s) of the insert was subcloned into plasmid vectors.

11. Preparation of DNA inserts for use as hybridization probes

The DNA inserts (used for the preparation of radioactive probes) were amplified by PCR, using specific synthetic oligonucleotides as primers. A Perkin-Elmer 480 thermocycler was used for the PCR reactions. All PCR amplifications were performed

in 50 or 100 µl of 1x PCR buffer (50 mM KCl, 10 mM Tris-HCl pH 8, and 1.5 mM $MgCl_2$) and were layered with 50 µl of mineral oil. The reaction also contained 0.15 mM of each dNTP, 1-2 units of Taq DNA polymerase, 1.0 uM of each primer and 10 ng of recombinant plasmid DNA (template). Amplification cycles were as follows: 92 ^{0}C for 65 sec, 52 ^{0}C for 65 sec, and 72 ^{0}C for 120 sec (with 10 sec for each additional extension). Twenty seven amplification cycles were preceded by a primary denaturation step (95 ^{0}C for 5 min) and followed by a final extension step (72 ^{0}C for 5 min), after the last cycle. After amplification and electrophoresis, inserts were visualized by ethidium bromide staining in low melting agarose gel. Recovery of insert DNA was done according to standard protocol (Sambrook et al.,1989).

12. Radiolabeling of DNA probes

Insert DNA was radioactively labeled using (α-^{32}P-labeled)-dCTP (3000 Ci/mmol, Amersham). Radiolabeling was done by random priming (Feinberg and Vogelstein, 1984). Radiolabeled probes were prepared, using 50 ng of DNA and Random Primed DNA labeling kit (US Biochemical). The protocol suggested by the vendor was used. The specific activity of the probe was greater than 1 x 10^9 dpm/ug. Unincorporated nucleotides were removed by precipitating the labeled reaction product with yeast tRNA. The radioactively labeled probes were denatured at 95 ^{0}C for 10 min, before use for hybridization.

13. Isolation of nuclear genomic DNA

Nuclear DNA was isolated from hypocotyls, using a modification of a protocol described by Dellaporta et al. (1983). Seeds were surface sterilized with 15% sodium hypochlorite and grown in the dark, between Whatman 3MM papers for 1-2 weeks. Hypocotyls were collected and frozen in liquid nitrogen and stored at -80 ^{0}C.

Thirty grams of hypocotyls were ground in liquid nitrogen (using mortar and pestle) to a fine powder. The powder was suspended in 200 ml homogenization buffer (100 mM Tris-HCl pH 8, 50 mM EDTA, 500 mM NaCl and 10 mM β- mercaptoethanol). The suspension was further homogenized by a Kinematica homonizer at setting 7 for 2 min and lysed with SDS (1% final concentration). After 30 min incubation at 65 ^{0}C, 0.3 volume of 3M potassium acetate, pH 4.8, was added and after gentle shaking the mixture was put on ice for 30 min. The dense, white precipitate was pelleted by centrifugation (GSA rotor, 10,000 rpm) and the nucleic acids in the supernatant were precipitated with 2 volumes of ethanol (100 %). Nucleic acids were recovered by low speed centrifugation, washed twice with 70% ethanol, and dissolved in 20 ml TE buffer (10 mM Tris-HCl pH8, 1 mM EDTA). Twenty grams of CsCl and 500 ul ethidium bromide (10 mg/ml) were added. The DNA was gradient purified in OptiSeal polyallomer tubes (Beckman) by ultra centrifugation overnight at 55,000 rpm in VTi 65 rotor. The DNA band was collected and ethidium bromide was removed with water saturated butanol. The DNA solution was diluted with one volume of distilled water and precipitated with ethanol. DNA fibers were spooled with a pasteur pipette, washed twice with 70% ethanol, air dried for 10 min and dissolved in 1 ml TE buffer.

14. Isolation of total and poly(A)$^+$ RNA

RNA was isolated using guanidium chloride extraction (Arrand, 1985). Plant material (50 g) was ground to a fine powder, in liquid nitrogen, using a mortar and pestle, and transferred to 250 ml of GK solution on ice (6 M guanidium chloride, 100 mM potassium acetate pH 4.8). The mixture was further homogenized by Kinematica homonizer at setting 7, for 2 min. Following a first centrifugation (20 min at 14,000 x g) and filtration through Miracloth to eliminate cellular debris, the filtered supernatant fraction was allowed to

precipitate for 1 h at - 20 ^{0}C, following the addition of 0.5 vol. of ethanol. After centrifugation the pellet was dissolved in 15 ml of TES (10 mM Tris-HCl pH7.5, 10 mm EDTA, 0.5% SDS) and extracted with water saturated phenol/chloroform (1:1 by vol). After centrifugation, the aqueous phase was ethanol precipitated by the addition of 0.1 vol. of 4M NaCl and 2 vol. of ethanol. The RNA pellet was collected after centrifugation and dissolved in 250 µl of distilled water.

Poly(A)$^+$ RNA was isolated from the total RNA, using the polyATract mRNA isolation system and the protocol recommended by the vendor (Promega), and stored as an ethanol precipitate at -80 ^{0}C.

15. Northern hybridization

RNA samples were separated in a vertical urea-agarose gel (Locker, 1979) for 5 h at 20 watts at 4 ^{0}C, blotted onto hybond-N$^+$ membrane (Amersham). Hybridizations and washings were carried out as indicated previously for genomic library screening.

16. Southern hybridization

Plasmid/genomic DNAs were digested with appropriate restriction enzymes. Digested fragments were size fractionated on 0.8% agarose gels, denatured, neutralized and transferred to Hybond-N$^+$ membrane. Hybridizations and washings were carried out as indicated previously for genomic library screening.

17. DNA sequencing and analysis

cDNA fragments and genomic DNA subfragments were cloned into appropriate restriction sites of either pUC19 or pBluescript II SK$^+$ vectors. Both strands of the plasmid DNA were used as

templates for DNA sequencing. Double stranded plasmid DNAs were sequenced according to Sanger et al., (1977), using synthetic oligonucleotide primers, Sequenase version 2.0 DNA sequencing kit (US Biochemical) and ^{35}S-dATP (Amersham). For sequencing of double stranded DNA, the denaturation and annealing conditions described by Zhang et al., (1988) were used. Analysis of the DNA and polypeptide sequence was carried out using Microvax II (UWGCG programs) and Macintosh LC II (Mac Vector^{TM}, version 4.1 and BLAST programs) computers.

18. Primer Extension Analysis

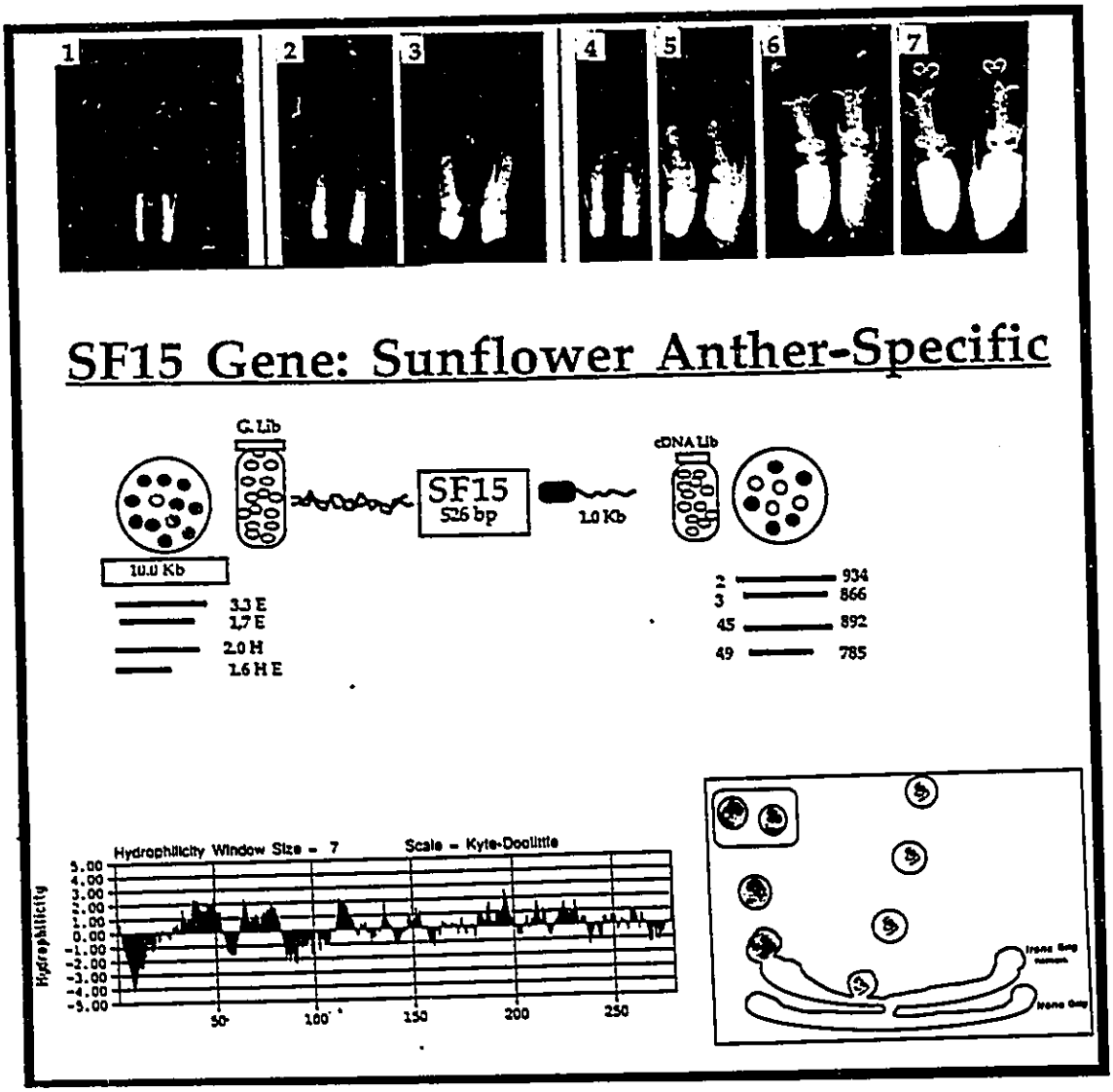
Oligo-nucleotide primers complementary to the 5' end of the template strand were used to determine the transcription initiation site(s). Primer extension reactions were carried out according to Sambrook et al. (1989). Briefly the oligomer was radiolabelled by 5'-phosphorylation using T4 Polynucleotide kinase and (γ-^{32}P) ATP. Labeled primer was annealed to 35-100 ug of freshly isolated total RNA (from appropriate organ system), in annealing buffer (20 mM Tris-HCl, pH 8.0; 200 mM NaCl; 100 mM EDTA and 1mM DTT) for 3h at 55 ^{0}C. The cDNA was made by extending the annealed primer with AMV reverse transcriptase (Bio/Can Scientific). The same oligonucleotide was also used as the primer in the DNA sequencing reaction. The reaction products of primer extension and DNA sequencing were run in the adjoining lanes on a sequencing gel (6% polyacrylamide and 8M urea).

19. *In situ* hybridization and light microscopy

Florets of stage 4 and 5 were fixed at 10 ^{0}C, in 2% paraformaldehyde, 0.2% glutaraldehyde, 0.1% Triton X-100 in 0.1 M sodium phosphate pH 7.0 buffer. And 16 μm transverse sections were cut with a Reichert-Jung 2800N cryomicrotome. Sections were

treated as described by Meyerowitz (1987) and hybridized with ^{35}S- labeled SF15 probe. Overnight hybridizations were performed , in 5X SSPE and 50% formamide in the presence of 10 ng labeled probe (4 x 10^6 cpm/slide) at 37 ^{0}C. The sections were washed 10 min in 5X SSPE at room temperature and twice in 0.2X SSPE at 65 ^{0}C. The sections were coated with an Ilford K5 nuclear track emulsion, to allow simultaneous observation of sections and autoradiograms. Alternatively sections were autoradiographed by Hyperfilm-βmax (Amersham) which were then processed separately. For histological studies, cross sections of disc florets were stained with 0.1% toludine blue and photographed with a light microscope (Cox and Goldberg, 1988).

SF15


1
2
3
4
5
6
7
SF15 Gene: Sunflower Anther-Specific
G. Lib
SF15
526 bp
1.0 Kb
cDNA Lib
10.0 Kb
3.3 E
1.7 E
2.0 H
1.6 H E
2
3
45
49
934
866
892
785
Hydrophilicity Window Size = 7
Scale = Kyte-Doolittle
Hydrophilicity
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-5.00
50
100
150
200
250

RESULTS

1. SF15 is specific to young unopened disc florets

Differential screening of the cDNA library constructed in λgt10 with poly(A)$^+$ RNA from a sunflower inflorescence, has resulted in the isolation of several flower-specific cDNA clones (Herdenberger et al., 1990). SF15 was one of these clones. The SF15 cDNA however, was incomplete and estimated to be about 500 bp in size, based on agarose gel electrophoresis. This clone was chosen for further analysis and subcloned into pUC19 and sequenced completely. It has a size of 526 bp and an open reading frame (ORF). Northern blot analysis was performed to determine the organ specificity of SF15 cDNA. Total RNA was isolated from vegetative (roots and leaves), and reproductive organs (ovaries, corolla, styles+ stigmas, pollen depleted anthers , mature pollen , immature seeds), as well as from closed disc florets. Individual organs such as ovaries, corolla, anthers, styles+ stigmas, mature pollen and immature seeds were harvested from disc florets 2 or 3 days after anthesis. After size fractionation, the total RNA was blot-hybridized with SF15 cDNA insert as a probe (Fig. 1). The SF15 cDNA probe, hybridized with 1000-nucleotide-long RNA transcripts detected exclusively in closed young disc florets. Results demonstrate that SF15 is specifically expressed in young closed florets, but the level of expression is very low (a barley visible band was observed after 5 days of exposure of the film). No expression was detected in other reproductive and vegetative organs, even after prolonged exposure of the autoradiogram.

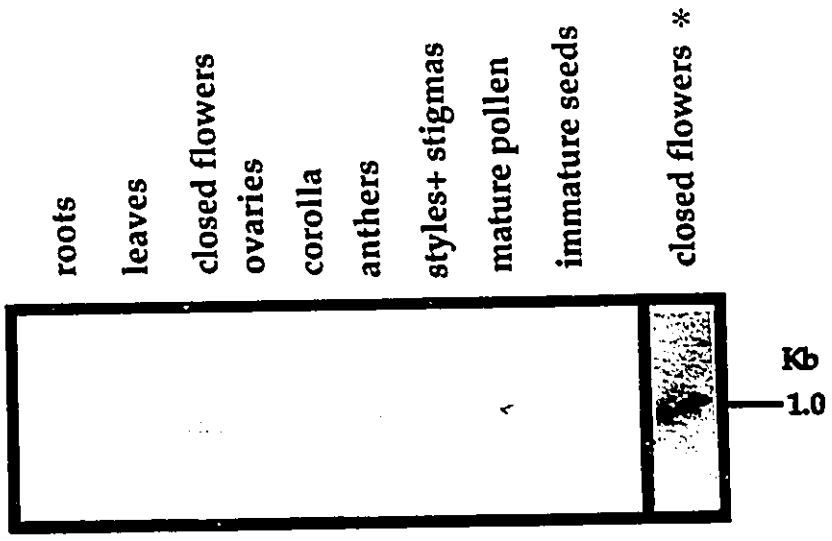


Figure 1. Floral specificity of sunflower SF15 cDNA clone. RNA gel-blot analysis was performed with 30 µg of total RNA from different tissues indicated. The RNA samples were size-fractionated on a vertical urea-agarose gel, transferred to a Hybond-N^+ membrane and probed with ^{32}P-labeled SF15 probe. The autoradiogram was exposed for 18 h (left panel with lanes not marked by *) and 5 days (right panel with a lane marked by *) with intensifying screen, at -80 ^0C.

2. Developmental staging of capitulum and disc florets

Initial studies on the pattern of expression of SF15 indicated that the gene is expressed in the young unopened disc florets before anthesis (and its expression in the floral organs harvested after anthesis, was not detected). In order to investigate the temporal expression of SF15 in detail, the development of the capitulum (inflorescence) was divided into three phases (I, II and III), and the development of the disc floret into seven stages (1, 2, 3, 4, 5, 6 and 7). The staging was based on the diameter of the capitulum, opening of the disc florets and capitulum, colour of the disc florets and so on. Figure 2 shows these phases and stages, and the morphological and anatomical features used to delineate each phase and stage are given below and summarized in Table 1.

The diameters of the capitulum at phases I, II and III were 6.0 9.0 and 13.5 cm, respectively. We designated these phases as immature capitulum (phase I), closed capitulum (phase II) and open capitulum (phase III). Phase I head was harvested from a 7-week-old plant. Phase II and III heads were harvested one and two weeks later, respectively. In the immature capitulum (phase I) the involucre of bracts completely masks the green ray florets (Fig.2 A-I). Developmentally the next phase (phase II) is one to two days prior to the opening of the capitulum and the beginning of anthesis. At this phase, petals of ray florets turn completely orange-yellow and their tips barely touch each other (Fig.2 A-II). In the open capitulum (phase III), one third (1/3) of disc floret whorls have completed anthesis (Fig. 2 A-III). Seven stages of florets were obtained from the heads of these three phases: the florets of phase I capitulum represented stage 1; florets of stage 2 and 3 were picked from phase II capitulum. Florets of stages 4, 5, 6, 7 were picked from phase III capitulum (Fig. 2A.1-7, and 2B).

Figure 2. Floret and inflorescence development in sunflower.
The development of capitulum and disc florets was divided into three phases (I, II and III) and seven stages (1, 2, 3, 4, 5, 6 and 7) respectively on the basis of size and other morphological criteria (see table 1).
(A) The three developmental phases of the sunflower capitulum with the corresponding developmental stages of disc florets. In the sequential development of the disc florets, stage 4 represents the midway between stages 2 and 3.
(B) The line diagrams of a sunflower capitulum indicating the locations of the seven floret stages of the three capitulum phases.

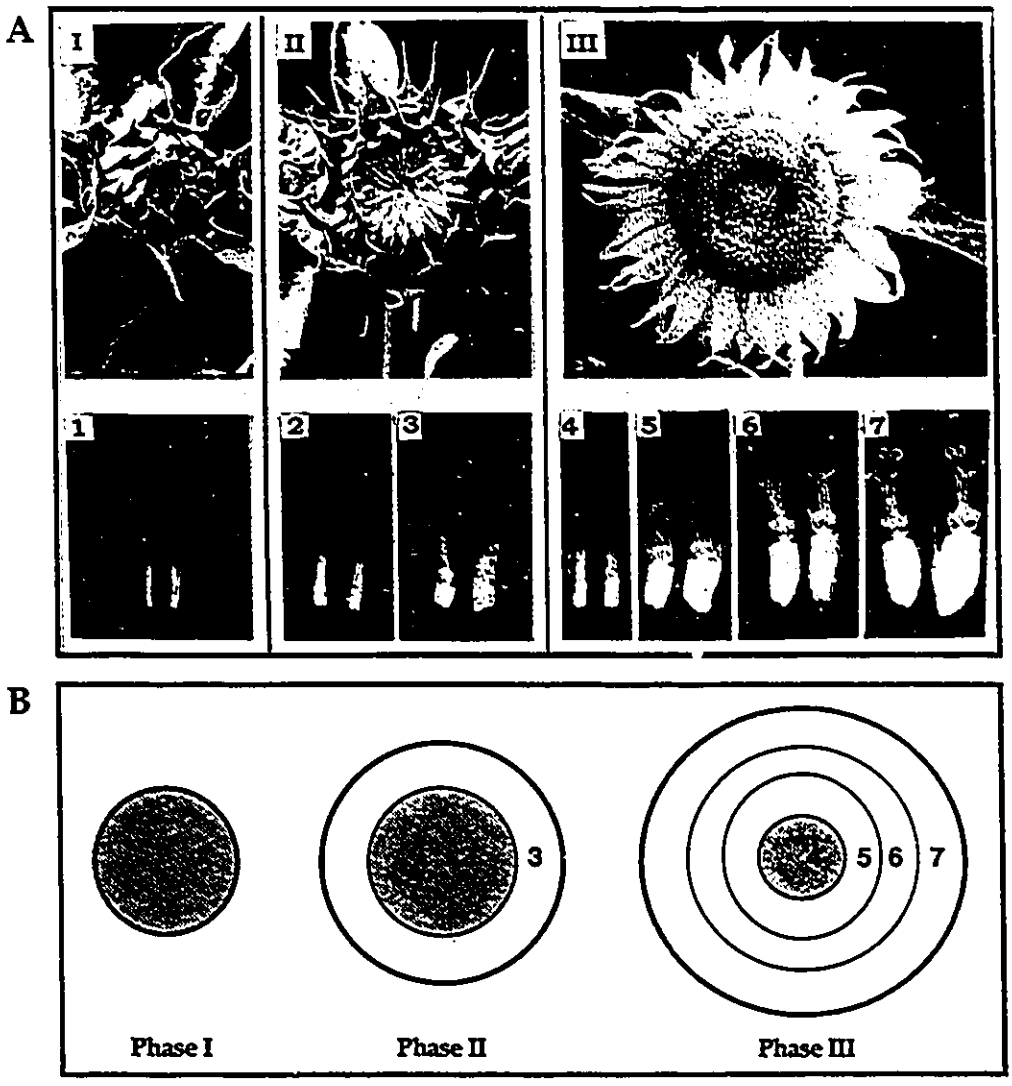

A
I
II
III
1
2
3
4
5
6
7
B
3
5
6
7
Phase I
Phase II
Phase III

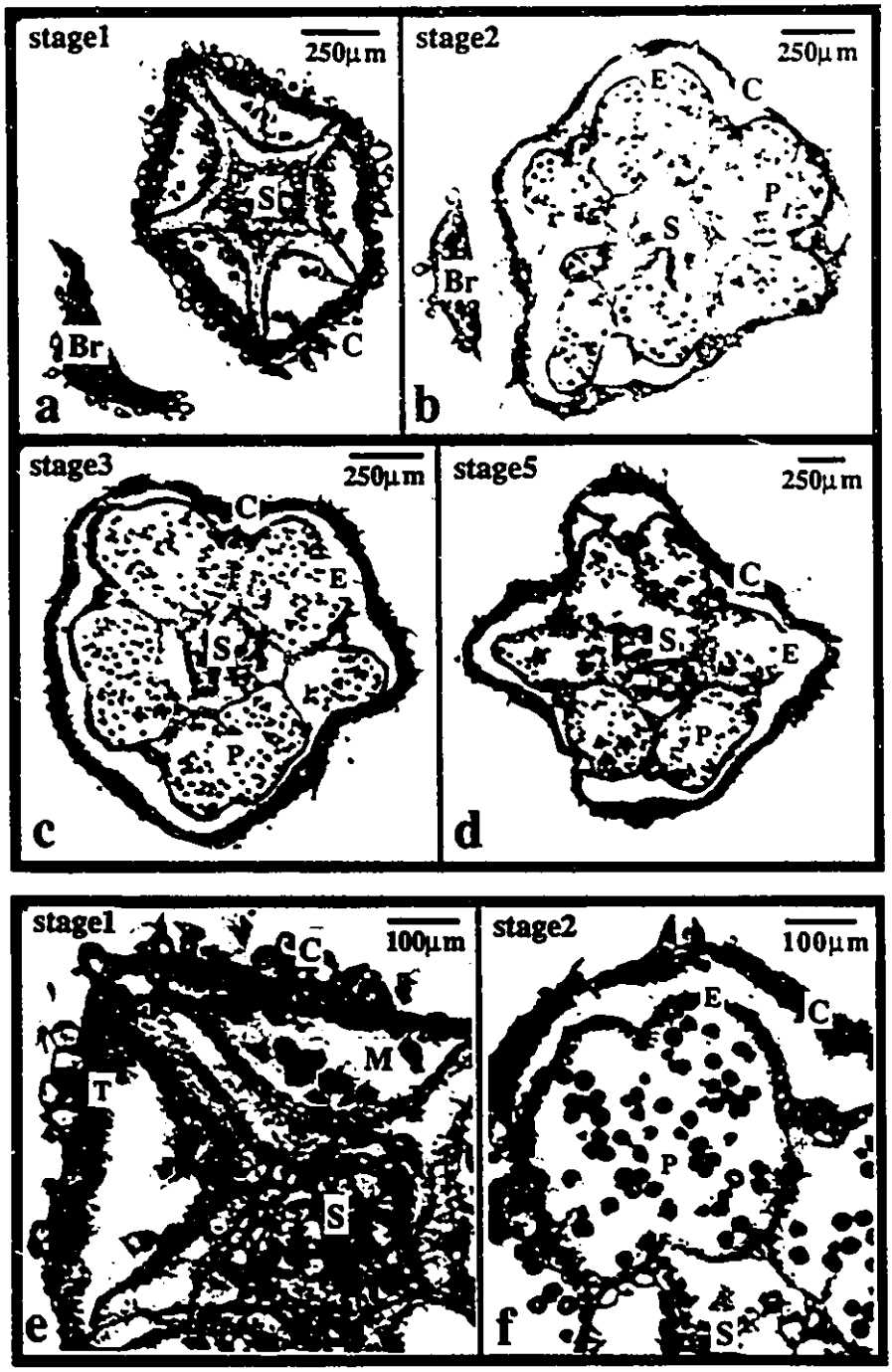
Table I. Characteristic features of capitulum and disc floret development in sunflower

Phase (of the head)	Age of plant (weeks)	Diameter of the head (cm)	Stage of disk floret	Size of disc floret (mm)#	Disc florets (open/ unopen)	Color of petal tips of disc florets	Features of capitulum/ disc floret
I. Immature capitulum	7	6.0					Bractioles completely mask green ray florets. Cell specification and histodifferentiation of stamen occur before stage 1.
			1	6.8	Closed	Green	Microsporogenesis is completed.Callose wall of microspores starts disappearing and microspores are released. Tapetal cells begin shrinking, connective tissue and bundles are disintegrated. Anthers are united.
II. Closed capitulum	8	9.0					One to two days prior to opening of capitulum and beginning of anthesis. Petals of the ray florets completely turn into yellow and their tips barely touch each other.
			2	8.0	Closed	Green	Degeneration of tapetum is complete. Pollen grains contain spiny exine. Pollen sacs and anther lobes are clearly visible. Anthers sustained by filaments have a simple structure consisting of a single layer of epidermis.
			3	11.0	Closed	Yellow	Pollen grains continue their development: they have well developed spines and three germinal furrows. Pollen wall contains outer layer exine and inner layer intine.
III. Open capitulum	9	13.5					About one third of disc floret whorls completed anthesis.
			4	9.8	Closed	Green	Florets are four to six days prior to anthesis. Developmentally the florets are a midway between stage 2 and 3.
			5	15.1	Closed	Yellow	Florets are one day prior to anthesis. Yellow tipped corolla tubes are longer than yellow bractioles. Florets are still closed, anther locules are filled with mature viable pollen grains.
			6	18.3	Open	Yellow	Florets are 1-2 hours after anthesis.Corolla tubes open, anthers dehisce and release tricellular pollen grains. Few hours later elongation of lower portion of style occurs and stigma appears.
			7	19.5	Open	Yellow	Florets are one to two days after anthesis. Florets have bilobed pubescent stigma, receptive to pollen. Pollination and fertilization follow. Anther tubes have alredy receded, stigma and style wither and recede 1-3 days later.

Measured from the base of the ovary to the tip of the petal (data is an average of 20 florets).

Some of the features used in the staging are adopted from Dedio and Putt (1980); Horner (1977); Knowles (1978); Nakashima and Hosokawa (1974); Nanda and Gupta (1975); Wodehouse (1935).

Figure 3. Bright-field photographs of cross-sections of disc florets. Disc floret cross sections of 16 µm were obtained as described in Materials and Methods. The sections were stained with toluidine blue and photographed with bright-field illumination. Panels a, b, c, and d are sections of stages1, 2, 3 and 5 respectively.
Panels e and f are higher magnifications of sections of stage1 and 2 respectively.
(a) Stage 1 floret: At this stage sporogenesis is complete, the callose wall of tetrads starts disappearing and tapetal cells start shrinking. Connective tissue and vascular bundles have disintegrated.
(b) Stage 2 floret: Disintegration of the tapetum is complete. Anther harboring the pollen grains, has a simple structure, consisting of a single epidermal cell layer and is sustained by the filament.
(c) Stage 3 floret: Pollen grains continue their development.
(d) Stage 5 floret: Anther locules contain the mature pollen grains.
(e) and (f) are higher magnifications of florets of stage 1 and 2 respectively.
(Br- bractiole; C-corolla tube; E-epidermis; M-microspore; P-pollen grain; S-stigma; T-tapetum)


stage1
250μm
S
C
Br
a
stage2
250μm
E
C
P
S
Br
b
stage3
250μm
C
E
S
P
c
stage5
250μm
C
S
E
P
d
stage1
100μm
C
M
T
S
e
stage2
100μm
E
C
P
S
f

In order to characterize the developmental events that occurred from stage 1 to 5 and to correlate morphological changes of disc florets with the presence of specific tissues, cross-sections of the disc florets of stages 1, 2, 3 and 5 have been analysed. Bright-field photographs of representative sections are shown in Fig. 3. In the sequential development of disc florets, stage 4 is intermediate between stage 2 and 3 (not shown in Fig. 3).

Cell specification and tissue differentiation of stamens occurs before stage 1. At stage 1, sporogenesis is completed, the callose wall of the microspores starts disappearing and the microspores are released; tapetal cells start shrinking, connective tissue and vascular bundles have degenerated (Fig. 3A, and E). At stage 2 disintegration of the tapetum is completed and the pollen grains continue their development; the wall of the anther locules is enclosed by the epidermis (Fig. 3 B and F). At stage 3, the pollen grains have well developed spines (Fig. 3 C). The stage 4 florets are developmentally midway between stage 2 and 3. At stage 5 (one day before anthesis) florets contain viable pollen surrounded by epidermis (Fig. 3 D). Anthesis occurs at stage 6 and the stigma becomes receptive to pollen at stage 7 (Table 1 for details).

3. Temporal expression of SF15 gene

To determine the temporal expression of SF15 transcripts during disc floret development, total RNA was isolated from disc florets of stages 1 to 7. The RNA gel-blot was hybridized with SF15 probe. A high level of expression of SF15 was observed in florets of stage 2, stage 4 and stage 3 (Fig. 4A). Closer inspection of the data indicated that the level of SF 15 transcripts, increases slightly in prevalence between stages 2 and 4 and declined in stage 3. A faint hybridization signal at stage 1, reveals that the accumulation of

SF15 transcripts begins at this stage. The data also indicate that SF15 gene is not expressed in the later stages (5, 6 and 7) of disc floret development .

4. SF15 gene is specific to male-fertile disc florets

To determine the pattern of expression of SF15 in male-sterile disc florets, total RNA was isolated from male-sterile disc florets corresponding to stages 1 to 7. The RNA gel-blot was hybridized with SF15 cDNA probe. In the male-sterile line (HA401 A), anther pigmentation and filament elongation are strongly reduced and pollen formation is aborted, but other organs of the floret develop normally. Northern blot analysis did not reveal expression of SF15 transcripts in male-sterile disc florets (Fig. 4. B). Together, these data indicate that SF15 cDNA clone represents cognate mRNA transcripts, expressed exclusively in young unopened male-fertile disc florets. SF15 gene is regulated temporally and there are timing differences in the expression patterns during disc floret development.

5. Organ specificity of SF15

In order to investigate the organ specific expression of SF15, the male-fertile disc florets were dissected into various organ systems i.e., ovary, bractiole, styles+stigmas and anthers+corolla (it is difficult to isolate anthers free of corolla tubes, in the unopened florets). Total RNA isolated from these organs, was size fractionated and blotted to Hybond-N$^+$membrane. The RNA gel-blot was assayed with SF15 probe. The results (Fig. 5) revealed that the probe hybridized to transcripts expressed in the anthers+corolla tube .

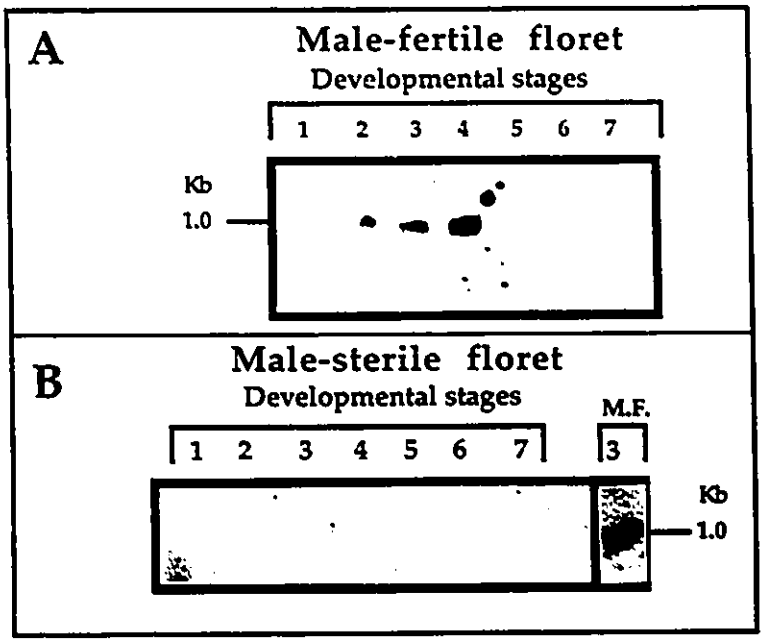


Figure 4. Temporal expression of sunflower SF15 gene.
RNA gel-blot analysis was performed with 30 μg of total RNA of male-fertile (A), and male-sterile (B) disc florets of different stages indicated. The RNA samples were electrophoresed on a vertical urea-agarose gel, transferred to Hybond-N^+ membrane and probed with ^{32}P-labeled SF15 probe. (M.F- Male-fertile disc floret of stage 3)

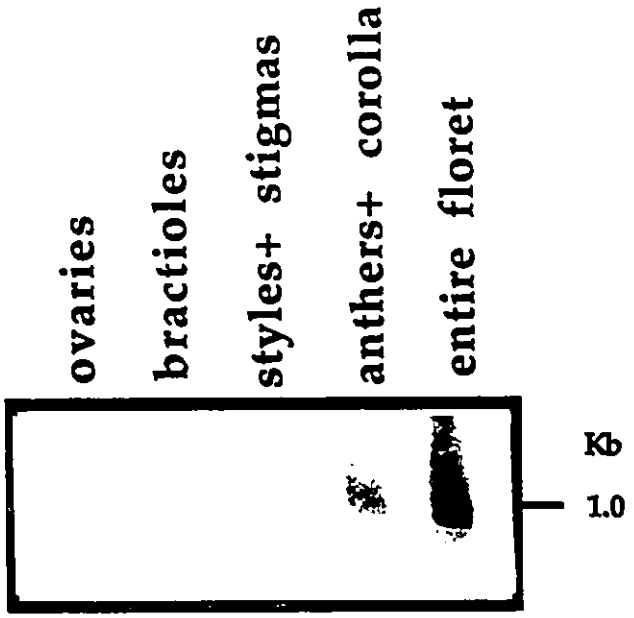


Figure 5. Organ specificity of sunflower SF15 gene. RNA gel-blot analysis was performed with total RNA from different organ systems of male-fertile disc floret as indicated (15 µg of total RNA from the organ systems and 40 µg from the entire stage 3 floret). RNA samples were electrophoresed on a vertical urea-agarose gel, transferred to Hybond-N+ membrane and probed with ^{32}P-labeled SF15 probe.

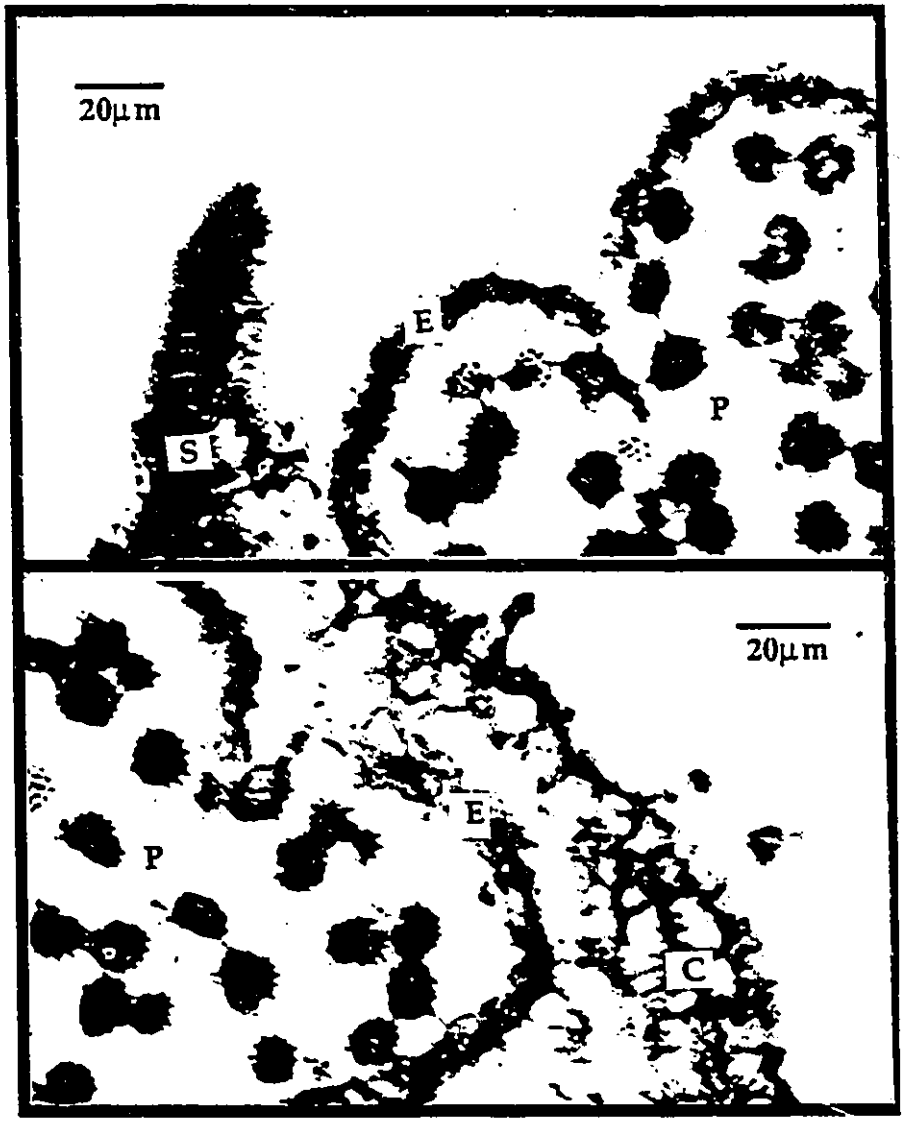


Figure 6. Localization of SF15 mRNA transcripts in sunflower anther. Cross sections (16 µm thick) of stage 4 male-fertile disc floret were hybridized with ^{35}S-labeled SF15 probe. The slides were exposed to radiographic emulsion for 12 days. The probe hybridizes to anther epidermal cells. No hybridization signal was detected on the pollen grains, stigma (top panel) or to the corolla tube (bottom panel). (Br- bractiole; C-corolla tube; E-epidermis; P-pollen grains; S-stigma)

6. *In situ* localization of SF15 transcripts

In order to localize the tissue specificity of SF15, disc floret cross-sections of stages 4 and 5, were hybridized with the ^{35}S-labelled SF15 probe. Floret sections were taken from the upper part of the floret, which consists of bractioles, petals, anthers and stigma. The results showed that the probe hybridized strongly to a single layer of anther epidermal cells of stage 4 disc florets (Fig. 6). The corolla, stigma and pollen grains did not show any hybridization signal. Also there was no hybridization of SF15 probe to any of these tissues including the anther of stage 5 disc florets (data not shown). These results were consistent with the results of RNA gel-blot analysis (Fig. 4A) and confirm the specificity of *in situ* hybridization.

7. Isolation and characterization of SF15 cDNA clones

In order to isolate full length cDNA clone(s), a floral cDNA library (of phase III) was screened with incomplete SF15 cDNA insert as a probe. Out of 300,000 phage recombinants screened, 90 clones hybridized strongly to SF15 probe. Four cDNA clones (SF15-2, -3, -45 and -49) with the longest inserts, were chosen for further investigation. Restriction analysis with EcoRI shows that the four cDNA clones can be grouped into two classes based on the presence or absence of internal EcoRI restriction sites (Fig. 7 and10). Class I cDNA clones (SF15-2 and SF15-3) have two internal EcoRI restriction sites. Class II cDNA clones (SF15-45 and SF15-49) do not contain any internal EcoRI sites. Class I cDNA clones were subcloned into the NotI site of pBluescript-II SK^+, and class II cDNA clones were subcloned into the EcoRI site of pUC19 (Fig. 7).

Ali these subclones were sequenced completely. The cDNA inserts were 785 to 934 bp long and comparable with the size of mRNA transcripts, as estimated by Northern hybridization (1 kb). The differences in size can be accounted for by a longer average poly(A)$^+$ tail at the 3' end in the mRNA population. Coding region analysis indicated that all four cDNAs contained open reading frames (ORFs), but terminated close to the 'ATG' start codon, compared to the ORF of the genomic clone (Fig. 11). The ORF of the genomic clone contains two in-frame 'ATG' initiation codons separated by three nucleotides.

In majority of higher-plant genes described so for, the first in-frame methionine codon of mRNA acts as a translation initiation site. In the case of SF15-49 cDNA, the second methionine codon (ATG) may act as a translation start site (because in this cDNA there is a substitition of T for G, of the first ATG initiation codon). Only 8 bp of the coding region were missing in the case of SF15-2 cDNA, where as 20 bp were found to be missing from first start codon of SF15-3 and SF15-45 cDNAs (Fig. 11). Three cDNAs (SF15-49, -3, and -45) were terminated by short poly(A)$^+$ tails of 7 to 9 residues at different positions, whereas SF15-2 cDNA has a long poly(A)$^+$ tail of 70 residues. The 3'-untranslated regions including poly(A)$^+$ tail were 90, 112 and 147 bp long for SF15-3, SF15-45 and SF15-2, respectively.

Comparison of cDNA sequences indicated that the polyadenylation site is variable. The putative polyadenylation motif 'AATAAA' has been found 36 and 40 nucleotides upstream of the poly(A)$^+$ stretch in SF15-2 and SF15-3 clones, respectively. And these positions are in close agreement with those reported for the majority of plant genes (27$\pm$ 9) (Joshi, 1987b). Two variant hexanucleotide sequences occured nearer to the polyadenylation site

in SF15-45 clone: 'AATTAT' was 29 bp and 'ATTAAA' was 23 bp upstream the poly(A)$^+$ taii, respectively. It is possible that one of these sequences might serve as polyadenylation signal. The SF15-49 clone has an atypical structure, the 'TAA' termination codon is followed sequentially by a poly(A)$^+$ tail without a 3' untranslated trailer.

8. Isolation and characterization of SF15 genomic clone

Screening of sunflower genomic DNA library constructed in λCharon 40 vector (with 526 bp SF15 cDNA as a probe) resulted in the isolation of several putative SF15 genomic clones (out of 1×10^6 plaques, 8 positive clones were detected after first screening). One genomic clone (SF15G) with an insert size of approximately 10.7 kb was selected for further investigation. Restriction mapping was performed and four fragments were subcloned (in pUC19 vector) separately: a 5' 1.7 kb EcoRI and two overlapping 2.0 kb HindIII and 1.6 kb EcoRI/Hind III clones and 3' 3.3 kb EcoRI (Fig. 8, 9 and 10). The coding region was located in the 1.6 kb EcoRI/Hind III and 2.0 kb HindIII fragments (Fig.10). These subclones were sequenced and shown to contain the complete coding region of SF15 gene as well as a 1.8 kb region upstream of the methionine translation start codon and 0.68 kb down stream from translation stop codon. The 5' untranslated region is characteristically A/T-rich (67%) when compared to the translated region (60%). The sunflower SF15G coding region is 837 nucleotide-long starting at the ATG codon (position 9-11) and terminating with TAA stopcodon (position 846-848). No introns were detected after comparison of the cDNA and genomic clones.

The transcriptional initiation site(s) were determined by primer extention analysis with total RNA isolated from disc florets of stage 4. A 31- mer oligonucleotide primer (3' C CTC AGT TAG AAC GGT TTC GAG GTT GTG GAG 5') , complementary to nucleotides 63 to 93 of the ORF of SF15 genomic clone (Fig.11), was extended by AMV reverse transcriptase. The primer extention products were electrophoresed next to the products of a DNA sequencing reaction using the same oligonucleotide as primer and genomic DNA subclone as template. The two major products differed in size by 1 nucleotide locating the transcription starts of the SF15 transcripts on the T nucleotides at position 8 and 6 bp 5' to the ATG start codon (shown as asterisks in Fig. 12 and Fig. 11).

Also there were small amounts of longer transcripts that started at A residue (17 bp), T (21 bp) and T (50 bp) upstream from the translational initiation codon, respectively (Fig. 11). The leader sequence of the main transcripts of SF15 gene is unusually short. The study of 79 plant genes showed that the length of the leader sequence ranged mostly from 40 to 80 nucleotides (Joshi, 1987a). In the 5' region three putative TATA boxes were identified (bold letters in Fig. 11). The first TATA box lies 30 and 32 bp upstream from the two major transcriptional initiation sites, the second TATA box is 33 and 29 bp upstream from the two closed minor transcription start sites. The third TATA box lies 30 bp upstream from the longest transcript initiation site. The position of these three putative TATA boxes is in good agreement with other start site distances from the TATA box observed in plants (32 ± 7) (Joshi, 1987a). Transcription of the SF15 gene might be initiated from different sites, each with a TATA box 29-33 bp upstream. It is possible that the set of transcripts varied in length at their 5'end exist, and are unequally expressed. The fact that each transcriptional initiation site has a TATA box 29-33 bp upstream strongly indicates that all sites might be active at different levels. It is possible that only one of these sites is under developmental control.

Two possible CAAT sequences were identified in the 5' non-coding region of the SF15 gene: the proximal CAAT box lies 31 bp upstream of the first putative TATA box and the distal one lies 359 bp upsream of the same TATA box (shown in bold letters in Fig. 11). Closer observation of data revealed a GATA motif (shown in bold letters in Fig. 11), 34 bp upstream of the proximal CAAT box . This motif is a cis-acting element in several plant light-responsible promoters (Lam and Chua, 1989). 34 bp upstream of the distal CAAT box is located a GTGG motif (shown in bold letters in Fig. 11), that has been shown to modulate promoter activity of the late pollen-specific genes from tomato (Twell et al., 1991).

In addition, numerous direct repeat structures are present in various parts of the 5'-flanking region of SF15 gene (Table 2). 314 and 66 nucleotides upstream of the first TATA box, there is a long 21 bp inverted repeat (underlined by arrows in Fig. 11) which might be involved in the regulation of gene expression by DNA looping. Two short (11 to 9 bp) perfect inverted repeats (underlined by arrows in Fig. 11) are also found in the promoter region of SF15 gene. The significance of all these repeat sequences is not known but they may be involved in gene-specific regulation. In the 3'-flanking region of the gene, three 'AATTAT' sequences were found 36, 46 and 77 nucleotides downstream of the TAA stop codon (shown in bold letters in Fig. 11), which might serve as polyadenylation signal.

A nucleotide sequence comparision of the four SF15 cDNAs and SF15 genomic clone reveal a high level of identity ranging from 89.9% to 99.7% among coding regions of the genes (Table 3 and Fig.11). These four cDNAs and the genomic clone represent five unique cognate mRNAs, and belong to a closely related family of genes.

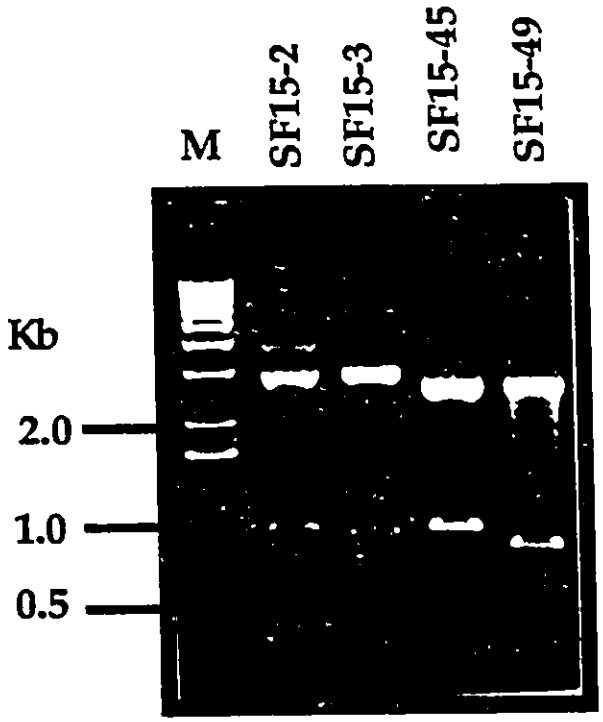


Figure 7. Agarose gel electrophoresis of four SF15 cDNA clones. About 30,000 λgt10 recombinant clones of sunflower floral cDNA library were screened with SF15 probe. Four cDNA clones (SF15-2, -3, -45 and -49) with longest insert sizes were chosen for sequence analysis and subcloned (into the EcoRI site of pUC19 or the NotI site of pBluescript SK+ vector). SF15-2 and -3 subclones were restricted with NotI and SF15-45 and -49 were restricted with EcoRI. The restriction products were electrophoresed on a 0.8% agarose gel. (M-Molecular size marker).

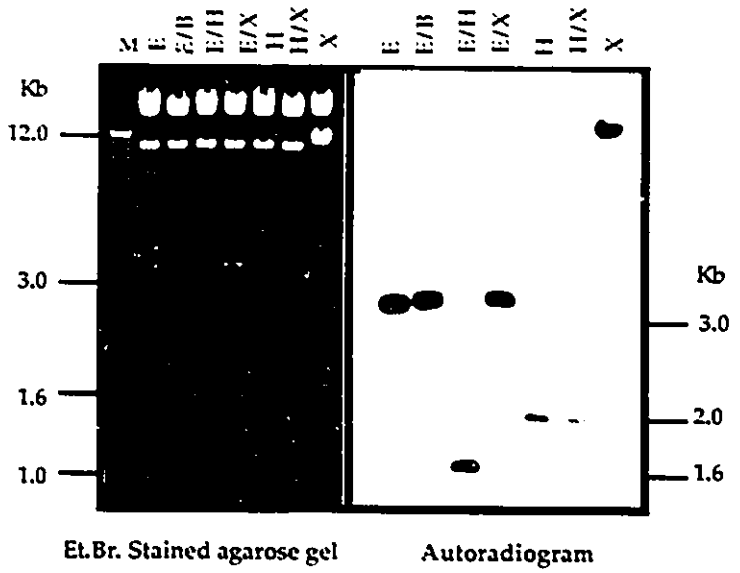


Figure 8. Restriction analysis of sunflower genomic clone SF15G. A 10.7 Kb genomic clone (SF15G) was isolated from the sunflower genomic library by screening with SF15 probe. The genomic DNA clone was restricted with various enzymes and the DNA-gel blot was probed with the SF15 probe. The left panel shows the ethidium bromide stained agarose gel and the right panel shows the autoradiogram. (M-Molecular size marker; B-BamHI; E-EcoRI; H-HindIII, X-XhoI)

Figure 9. Gel electrophoresis of subclones of the SF15 genomic clone. Four restriction fragments of SF15G were subcloned into pUC19 vector i.e., 3.3 Kb EcoRI; 1.7 Kb EcoRI; 2.0 Kb HindIII and 1.6 Kb EcoRI/HindIII. These subclones were electorphoresed after digestion with appropriate enzyme(s). (M-Molecular size marker; B-Bam HI; E-EcoRI; H- HindII).

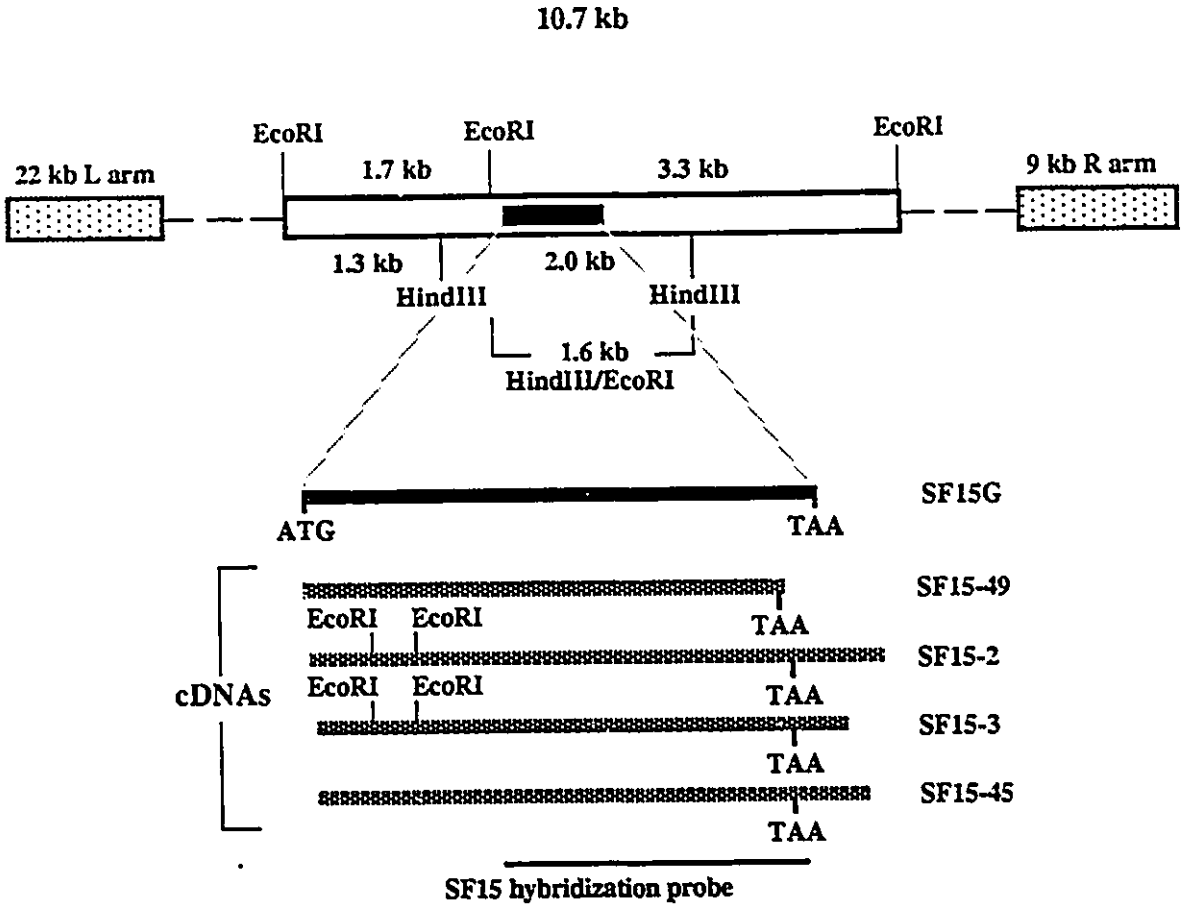
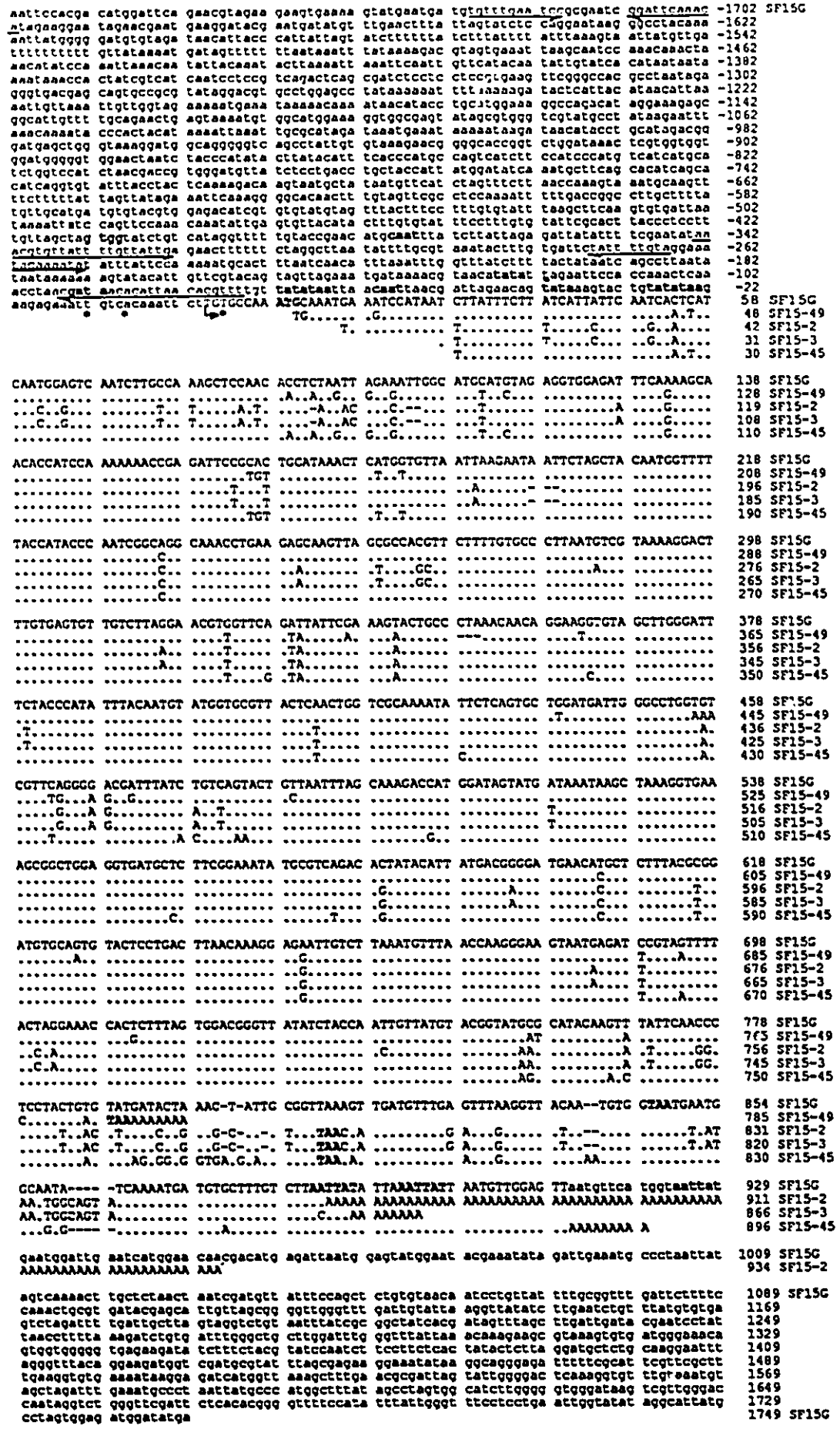


Figure 10. Restriction maps of genomic clone SF15G and SF15 cDNAs. Closed box represents the open reading frame of the genomic clone. Stippled boxes at the ends of genomic clone represent the λ Charon 40 vector arms. The cDNA clones and the SF15 probe are represented with reference to the open reading frame of the genomic clone.

Figure 11. Comparison of nucleotide sequences of sunflower genomic clone (SF15G), and four SF15 cDNA clones (SF15-49, SF15-2, SF15-3 and SF15-45). For the coding regions of the cDNAs, only the nucleotides that differ from the genomic clone are shown. Dots indicate identical nucleotides and a dash indicates a gap in the sequence introduced, in order to maintain an optimal alignment. Nucleotides in the 5' and 3' non coding regions are indicated in lower case letters. The translation start codon (ATG), translation termination codon (TAA), and the potential polyadenylation signals (AATTAT, AATTAT, aattat) are shown in bold letters. Three TATA boxes, two CAAT boxes, two cis-elements (GATA and GTGG) are also shown in bold letters. Inverted repeats present in the 5' region are underlined. Transcription initiation sites are indicated by asterisks.(The nucleotide sequence data reported here will appear in the EMBL data bank under the accession numbers X 82005, X 82006, X 82007, X 82008, and X 82009).



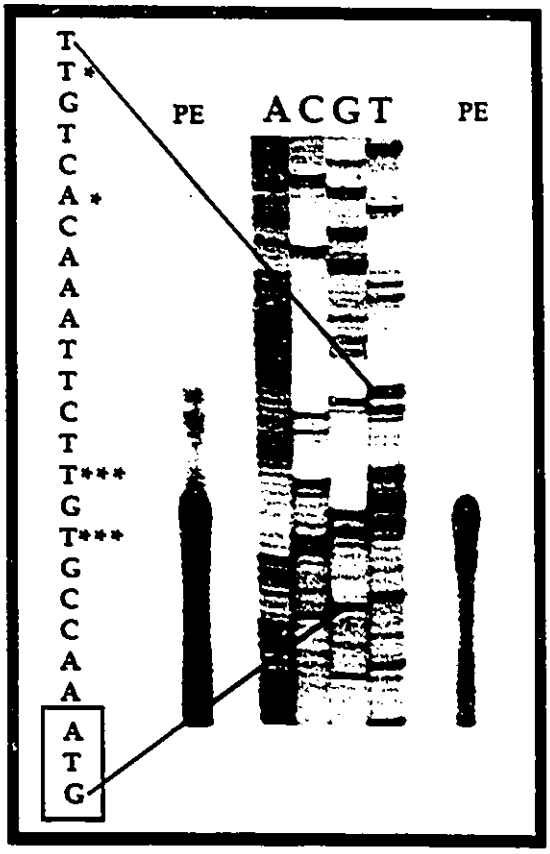


Figure 12. Mapping of transcription initiation sites for SF15. Primer extension was performed with a ^{32}P-labeled, 31 nucleotide oligomer (3'-C CTC AGT TAG AAC GGT TTC GAG GTT GTG GAG-5') complimentary to nucleotides +63 to +93 of the coding region of SF15G genomic clone (Fig. 11). The end labeled primer was annealed to 100 µg of total RNA (of stage 3 disc floret) and extended by reverse transcriptase. The primer extension products (PE) and the products of sequencing reactions (A,C,G,T) using the same primer, were size fractionated on a 6% polyacrylamide-urea (8M) gel, and analyzed by autoradiography. The nucleotide sequence arcund the transcription start site(s) are shown. Asterisks indicate the positions of major and minor initiation sites.

Table 2. Direct repeat sequences found in the 5' flanking region of SF15G gene. Position numbers are relative to the transcription initiation site.

No.	Sequence	Size	Positions
1.	5'-TA(T/C)GAATGAT(G/A)TGTTTGAA-3'	19 n	-1741-1723 and -1676 -1658
2.	5'-TAC(T/A)TAA(A/C)ATTAAA(T/A)T(C/T/G)-3'	17 n	-1433 -1417; -1235-1219 and -1047-1031
3.	5'-ATTGTTC(A/G)TACA(A/G)TA-3'	15 n	-1415 -1401 and -165 -151
4.	5'-TAGA(A/T)AAATGAAATAAAAA(C/T)AA-3'	22 n	-1205-1184 and -1026-1005
5.	5'-ATAACATACCTGCAT(G/A)GA-3'	18 n	-1182 -1165 and -1003-986
6.	5'-TTTCCTTTGTGTATT-3'	15 n	-537- 523 and -453 -439
7.	5'-TT(A/T)G(A/T)GATT(A/C)TATTTT-3'	16 n	-367 - 352 and - 276 -261

Table 3. Nucleotide sequence and deduced amino acid sequence identity among the members of the SF15 family.

	SF15-2	SF15-3	SF15-45	SF15-49
SF15-G	90.6% 84.0%	90.9% 84.1%	92.8% 85.8%	93.6% 87.6%
SF15-2		99.7% 99.6%	89.8% 81.9%	90.5% 83.2%
SF15-3			89.8% 82.4%	90.8% 83.4%
SF15-45				95.6% 94.0%

top - nucleotide sequence identity
bottom - amino acid sequence identity

9. Protein structure of SF15 products

The amino acid sequence of the putative polypeptides encoded by the SF15 gene family is shown in Fig. 12. The molecular mass of these proteins range from 29.1 to 31.7 kDa, if the missing sequences of incomplete cDNAs are taken into consideration . The complete SF15G open reading frame encodes a putative polypeptide of 280 amino acids (Fig. 13) with a molecular masss of 31.7 kDa. The predicted iso-electric point (pI) for mature protein is 9.6. The hydropathy plot according to Kyte and Doolittle indicates that SF15G protein is relatively hydrophilic (Fig. 14). The N-terminal region of the SF15G protein shows the characteristic features of a signal peptide (Fig. 13) (von Heijne, 1986, 1990). A possible cleavage site is located between position 25 and 26 (Fig. 13) conforming to the (-1, -3) rule (von Heijne, 1986). Three potential N-linked glycosylation sites are predicted in SF15 protein at amino acids 41, 63 and 64 from the initial methionine (underlined in Fig. 13). The repeated amino acid sequence, Val-Asn-Leu-Ala-Lys, is also found at positions 20 and 161 (overline in Fig.13).

The predicted amino acid sequences of the SF15 encoded by each clone are 81.9% to 99.6% identical (Table 3). Most of differences among the sequences are located in the N-terminal region (first 35-42 codons) and in the last 16 to 28 codons at the carboxyl terminus. Comparison of the nucleotide and deduced amino acid sequences for SF15 G with sequences in the GenBank databases revealed no significant homology to known proteins.

```
MQMKSIILFL IIIQSLINGV NLAKAPTPLI RNWHACRGGE LKSNTIQKNR 50 SF15G
 E-------- -----I----. -------TIV GD--VS---D FR-------- 49 SF15-49
   ------- F-L-AI---- --VI-QS-KT DR.-V----- FR-------- 46 SF15-2
       --- F-L-AI---- --VI-QS-KT DR.-V----- FR-------- 42 SF15-3
       --- -----I---- -------TIV GD--VS---D FR-------- 43 SF15-45

DSALDKLMVL IKNNSSYNGF YHTQSAGKPE EQVSATFFCA LNVVKGLCEC 100 SF15G
--V----LL- ---------- ---------- ---------- ----------  99 SF15-49
---------- ---.------ ---------- -----S---- --I-------  95 SF15-2
---------- ---.------ ---------- -----S---- ----------  91 SF15-3
--V----LL- ---------- ---------- ---------- ----------  93 SF15-45

CLRNVVQIIR KYCPKQQEGV AWDFYPYLQC MVRYSTGRKI FSVLDDWANC 150 SF15G
-------YI- -N-.Q---C- ---------- ---------- ---------K 148 SF15-49
--K----Y-- -N-------- ---------- ---------- ---------Y 145 SF15-2
--K----Y-- -N-------- ---------- ---------- ---------Y 141 SF15-3
------RY-- -N------A- ---------- ---------- L--------Y 143 SF15-45

RSGDDLSVST VNLAKTMDSM INKLKVKAAG GDALRKYASD TIHYDGDEHA 200 SF15G
-L-SG----- A--------- ---------- ---------- ---------- 198 SF15-49
---S------ ---------- L--------- ---------- S----E---- 195 SF15-2
---S------ ---------- L--------- ---------- S----E---- 191 SF15-3
-L----Y-K- ---------- ---------- ---------- S--------- 193 SF15-45

LYADVQCTPD LTKENCLKCL TKGSNEIRSF TRKPLFSGRV ISTNCYVRYA 250 SF15G
---------- ----D----- --------N- ----R----- ---------D 248 SF15-49
--V------- ----D----- -----K---- -K-------- ---------K 245 SF15-2
--V------- ----D----- -----K---- -K-------- ---------K 241 SF15-3
--V------- ----D----- --------N- ---------- ---------R 243 SF15-45

HTSLFNPPTV YDTKLLRLKL MFEFKVTMW                        279 SF15G
---I-----E                                             258 SF15-49
---I-KA-ID F-PEACG                                     262 SF15-2
---I-KA-ID F-PEACG                                     258 SF15-3
--N------E -SGGDECG                                    261 SF15-45
```

Figure 13. Comparison of the deduced amino acid sequences of SF15 proteins. The deduced amino acid sequence of the genomic clone is shown on the upper line in single letter code. For the cDNA clones only amino acids differing from the corresponding amino acid sequence of SF15G are shown. Dashes indicate sequence identity and gaps (indicated by dots) are inserted in the amino acid sequence to get the maximum alignment. The arrow indicates the cleavage site of the signal peptide. Putative glycosylation sites are underlined, identical amino acid regions are overlined. Bold faced amino acids in the C-terminus represents the putative vacuolar targeting region.

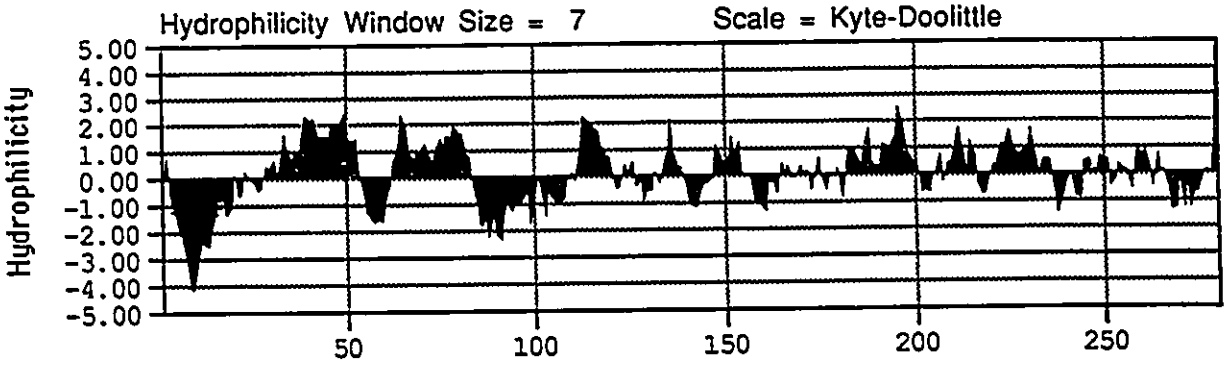


Figure 14. Hydropathy plot of SF15G protein. The hydropathy plot of deduced SF15G protein is according to Kyte and Doolittle (1982).

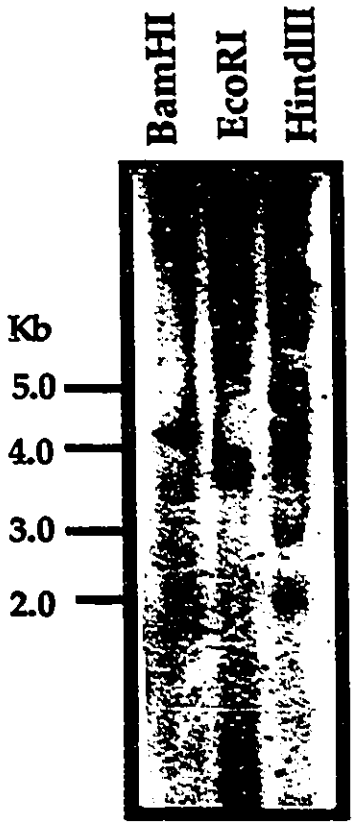


Figure 15. Southern blot analysis of sunflower nuclear genomic DNA. Nuclear genomic DNA (20 μg) was restricted with indicated enzymes, electrophoresed on agarose gel, transferred to Hybond-N^+ membrane and probed with ^{32}P-labeled SF15 probe. Hybridization and washings were performed at 65 ^0C

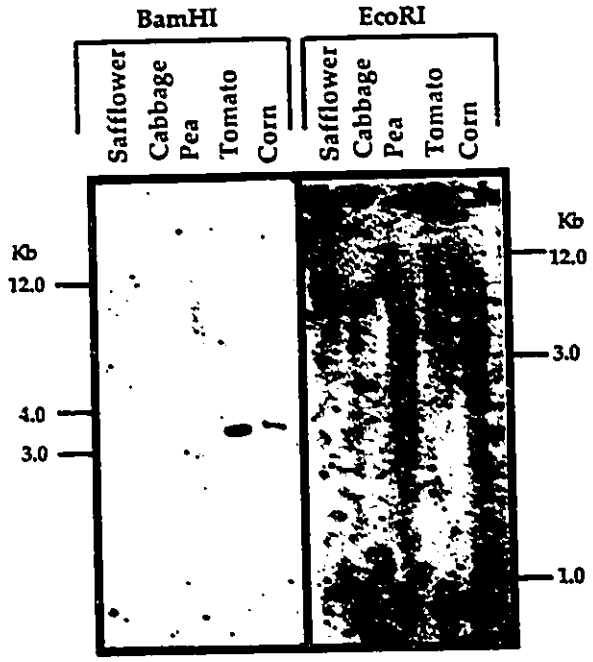


Figure 16. Southern blot analysis of genomic DNAs of different plants. Nuclear genomic DNAs of safflower, cabbage, pea, tomato and corn (20 μg) were restricted with indicated enzymes, electrophoresed on agarose gel, transferred to Hybond-N^+ membrane and probed with ^{32}P-labeled SF15 probe. Hybridization and washings were performed at 60 ^0C.

10. SF15 gene is a member of a multigene family

To know the copy number of the SF15 gene, Southern hybridization analysis was performed on sunflower nuclear genomic DNA. Nuclear genomic DNA was digested to completion with three restriction enzymes (BamHI, EcoRI and HindIII). The DNA gel-blot was hybridized with 526 bp SF15 cDNA probe. Results shown in Figure 15 reveal that the probe hybridized to several restriction fragments of each enzyme (from three to eight). This confirms that the SF15 gene belongs to a multigene family in the sunflower nuclear genome.

11. Homologous sequences of SF15 in other higher plants

To identify related sequences in other higher plants, nuclear genomic DNAs have been isolated from monocot (maize) and dicots (cabbage, pea, safflower and tomato). The genomic DNAs were digested to completion with BamHI and EcoRI. DNA gel-blots were hybridized with SF15 cDNA probe, under reduced stringency. Figure 16 shows that the probe strongly hybridized with only one EcoRI or BamHI fragment in maize and tomato genomes. Two weak bands have been found in cabbage genome indicating the presence of less closely related sequences. These data demonstrate that SF15 is a single-copy gene in maize and tomato genomes whereas in sunflower it exists as a family of six to eight members.

DISCUSSION

1. Spatial and temporal expression of SF15

The initial objective of this study was to isolate and characterize new flower-specific genes in the sunflower. Differential screening of the cDNA library constructed in λgt10, with poly(A)$^+$ mRNA from sunflower (phaseIII) inflorescence has resulted in the isolation of a number of the flower-specific cDNA clones (Herdenberger et al., 1990). Some of these clones have already been characterized on the basis of their expression specificity in the floral organs (Baltz et al., 1992 a,b; Domon et al.,1990, 1991,1994; Dudareva et al.,1994 a,b; Evrard et al., 1991).

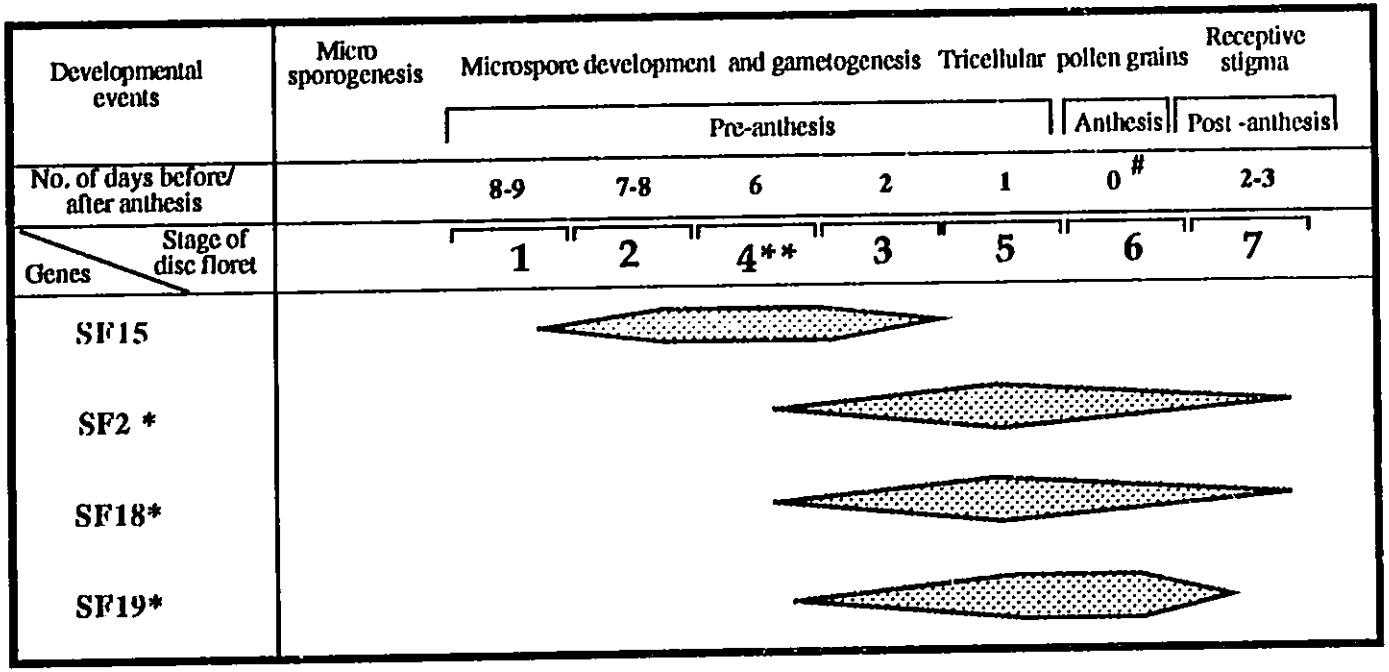
The SF15 gene family is expressed in young green closed pre-anthesis disc florets (Fig.1). The temporal expression of the SF15 gene family was examined by Northern blot analysis of total RNA extracted from florets of different stages of development. Disc floret development in sunflower was divided into seven stages (Fig. 2 and Table 1) and the first five stages were microscopically characterized (Fig. 3 and Table 1). The earliest stage at which SF15 mRNA could be detected was in closed stage 1 green florets (microspore release). Thereafter, the steady state level increased and reached its highest level in closed green flowers at stage 4 (pollen grain development). The amount of SF15 mRNA slightly decreased at stage 3 (closed yellow flowers, which developmentally follow the stage 4; pollen maturation stage before anthesis). The expression of SF15 transcripts was undetectable in the later stages of floret development i.e., stages 5-7 (Fig. 4A). Transcripts corresponding to the SF15 gene were not

found in male-sterile disc florets (Fig. 4B). *In situ* hybridization experiments demonstrated anther specificity of SF15 transcripts (Fig. 6). These results clearly indicate that expression of the SF15 gene family is spatially limited to the anther epidermal cell layer and observed only during a short period of floret development in male-fertile disc florets.

2. Anther-specific SF2, SF18, SF19 and SF15 are expressed in the epidermis

The characterization of three anther-specific genes (SF2, SF18 and SF19), expressed exclusively in the single cell layer of the anther epidermis in sunflower, has been reported previously (Evrard et al., 1991). These genes are expressed at the late stage of anther development, before and during pigmentation and elongation, and also in mature anthers (after anthesis). But the temporal pattern of expression of SF15 transcripts is different from those of SF2, SF18 and SF19. Table 4 presents a schematic representation of sunflower anther-specific gene expression programs, that have been identified in this study (and reported previously by Evrard et al., 1991). In this expression program, SF15 belongs to the earliest sunflower anther-specific transcripts identified to date. Its expression begins during microspore release (stage1). These results and also studies from various groups (Goldberg et al.,1993; Koltunov et al., 1990; McCormick, 1991, 1993; Scott 1993), demonstrate that the expression of anther-specific genes is regulated by many different (temporal and spatial) control programs and these programs correlate with differentiation and degeneration of specific tissues and cell types during anther development.

Table 4. Temporal patterns of expression of anther-specific epidermal genes of sunflower.



* The temporal expression of SF 2, 18 and 19 genes reported by Evrard et al., 1991 are presented according to the staging system described in this report.

** Developmentally the florets are midway between stage 2 and 3.

Florets are harvested 1-2 hours after anthesis

3. SF15 multigene family and SF15 proteins

Southern blot analysis shows that SF15 gene family in sunflower contains six to eight members (Fig.15). The four cDNA clones (SF15-2; -3; 45; and -49) and one genomic clone (SF15G) represents transcripts of at least five different cognate cistrons (Fig. 11). The simplest interpretation is that the different cDNA clones represent transcripts from different genes. In many cases it has been shown that individual members of multigene families are selectively expressed at different developmental stages in a tissue- or organ-specific manner or differentially regulated in response to distinct external stimuli (Coruzzi et al., 1984; Kuhlemeier et al., 1987; van Tunen et al., 1988). All of the genes in the SF15 family, however, are expressed coordinately in anthers during a short period of time. Since the nucleotide sequence of the genes are quite similar (89.8% to 99.7% identity, Table 3 and Fig. 11) it was impossible to discriminate different transcripts in Northern blot or *in situ* hybridization experiments. Nevertheless, the mRNAs produced from various gene family members might be present at different levels of abundance in anthers.

The SF15G gene encodes a basic polypeptide of 31.7 kDa. In common with the majority of anther-specific proteins described to date (Koltunov et al., 1990; Scott et al., 1991b; Scott 1993; Staiger and Apel, 1993) the SF15 polypeptide has a recognizable N-terminal signal peptide. This suggests that newly synthesized SF15 proteins enter the biosynthetic-secretory pathway, by crossing the ER membrane (Chrispeels, 1991). From the ER the proteins are transported to the Golgi apparatus, where they are modified by a series of glycosyltransferase enzymes. The genomic protein (SF15G) contains three potential glycosylation sites near its N-terminus, whereas SF15-49 and SF15-45 proteins have two, and SF15-2 and SF15-3 proteins have only one glycosylation site, respectively

(Fig. 13). It is known that the proteins traversing the secretory pathway are sorted to their respective compartments by selective retention information (or targeting information) contained in their molecular structures (Rothman, 1987). Proteins lacking specific retaining determinants follow a default pathway and are consequently secreted to the cell surface (Rothman, 1987; Denecke et al., 1990).

A comparison of the deduced amino acid sequences of a number of the proteins that have both vacuolar and cell wall forms indicates the presence of C-terminal extensions on the vacuolar forms. The C-terminal extensions are usually rich in hydrophobic amino acids, and this characteristic may be recognized by the sorting machinery (Chrispeels and Raikhel, 1992). To date, no common consensus sequences or structural elements that function as vacuole localization signals in the amino- and carboxyl-terminal propeptides have been identified. This means that there may be multiple independent mechanisms for vacuolar protein sorting (Chrispeels and Raikhel, 1992).

A comparison of the coding region of the SF15G gene and the homologous cDNA clones indicates that different members of the SF15 gene family code for basic proteins with similar amino acid sequences but with different sizes (Table 3 and Fig.13). The longest putative polypeptide corresponds to (SF15G) genomic protein of 31.7 kDa. The higher molecular weight of the genomic protein (Fig. 13) is caused by the presence of a short hydrophobic C-terminal domain (10 hydrophobic amino acids out of 16). It is likely that this terminus contains the vacuolar sorting information, but it remains to be confirmed. So, the genomic SF15 protein is an example of a vacuolar protein that is processed at its C-terminus. All the other SF15 proteins (corresponding to isolated cDNA clones), lack the putative

vacuolar sorting determinant, and could be retained in the ER or in the Golgi or secreted by the default mechanism via secretory vesicles. These results indicate that the SF15 proteins have both extracellular and a vacuolar location and are the products of different genes (Fig. 17). In the case of barley lectin (a vacuolar protein), the vacuolar targeting sequence is removed before or concomitant with deposition of the mature protein in vacuoles (Bednarek et al., 1990).

4. Evolutionary conservation

If SF15 gene is important during earlier stages of flower development it probably will be evolutionarily conserved and consequently present in the genomes of a variety of plant species. To investigate the presence of the SF15 genes in other plant species, a Southern blot analysis of nuclear genomic DNAs from safflower, cabbage, pea, tomato and corn was performed (Fig. 16). The results indicate that homologous SF15 genes are present in a single copy in corn and tomato genomes, whereas in sunflower it exists as a family of six to eight members (Fig. 15). Two weak hybridization bands in cabbage DNA indicate the presence of less closely related genes in the cabbage genome.

Several anther-specific genes (expressed in the pollen grains) have been reported and these genes have sequence similarities to various hydrolytic enzymes. These sequence similarities include two tomato anther-specific cDNA clones, LAT56 and LAT59, (Wing et al., 1989) and two maize Zm58.1 and Zm58.2 clones to pectate lyases (Turcich et al., 1993); P2 gene family from *Oenothera organensis* (Brown and Crouch, 1990) and tobacco Npg1 gene (Tebbutt et al., 1994) to polygalacturonase and *Brassica napus* Bp19 gene (Albani et al., 1991) and *Petunia inflata* PPE1 gene (Mu et al., 1994) to pectin esterases. A search of existing data bases, using both the nucleotide

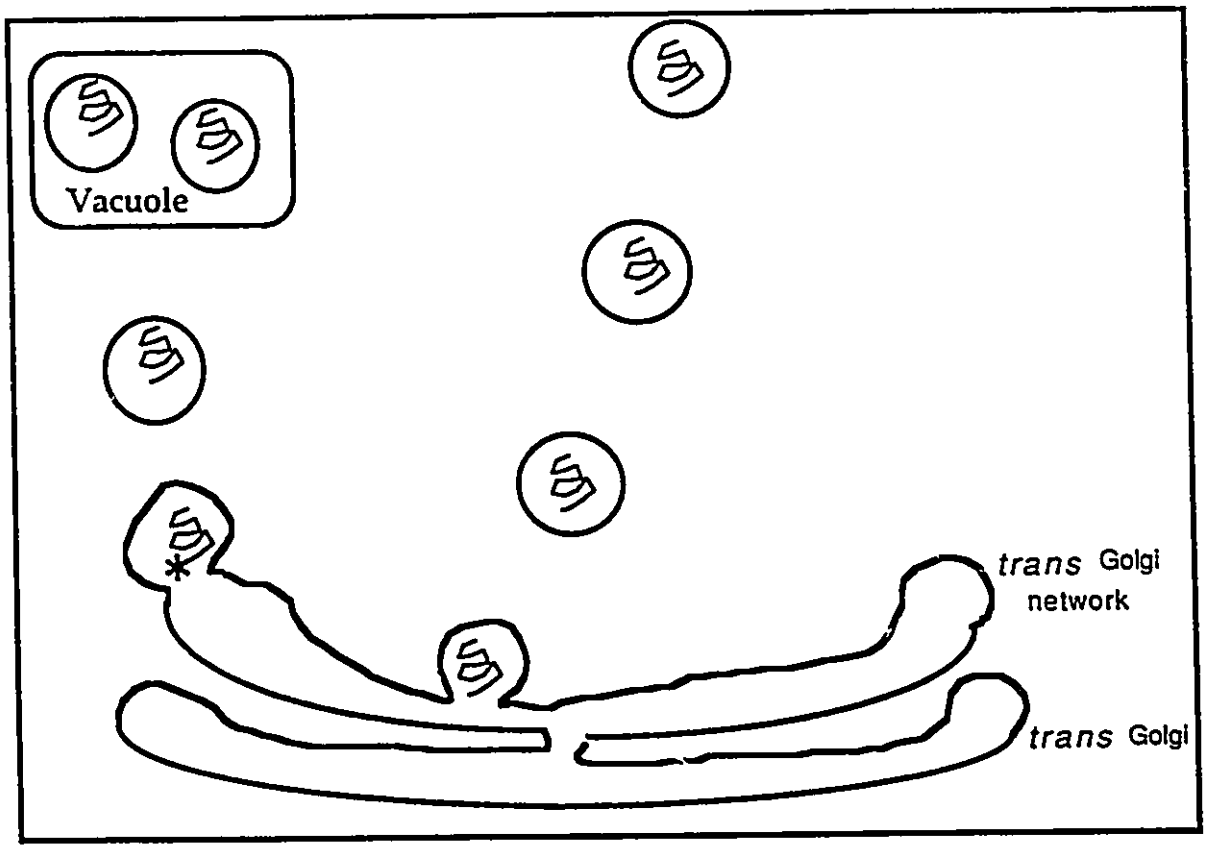


Figure 17. Vacuolar and secretory pathways of SF15G protein and SF15 cDNA proteins. Proteins translated from SF15G transcripts contain a vacuolar targeting sequence (*) and are targeted to the vacuole. Proteins translated from SF15 cDNA transcripts lack a vacuolar targeting sequence and follow the secretory pathway (Figure is drawn based on Bednarek et al., 1990).

and deduced amino acid sequences of the SF15 cDNA and genomic clones, revealed no significant homology with any known proteins. The biological function of the SF15-encoded protein is yet to be determined.

5. Anther dehiscence

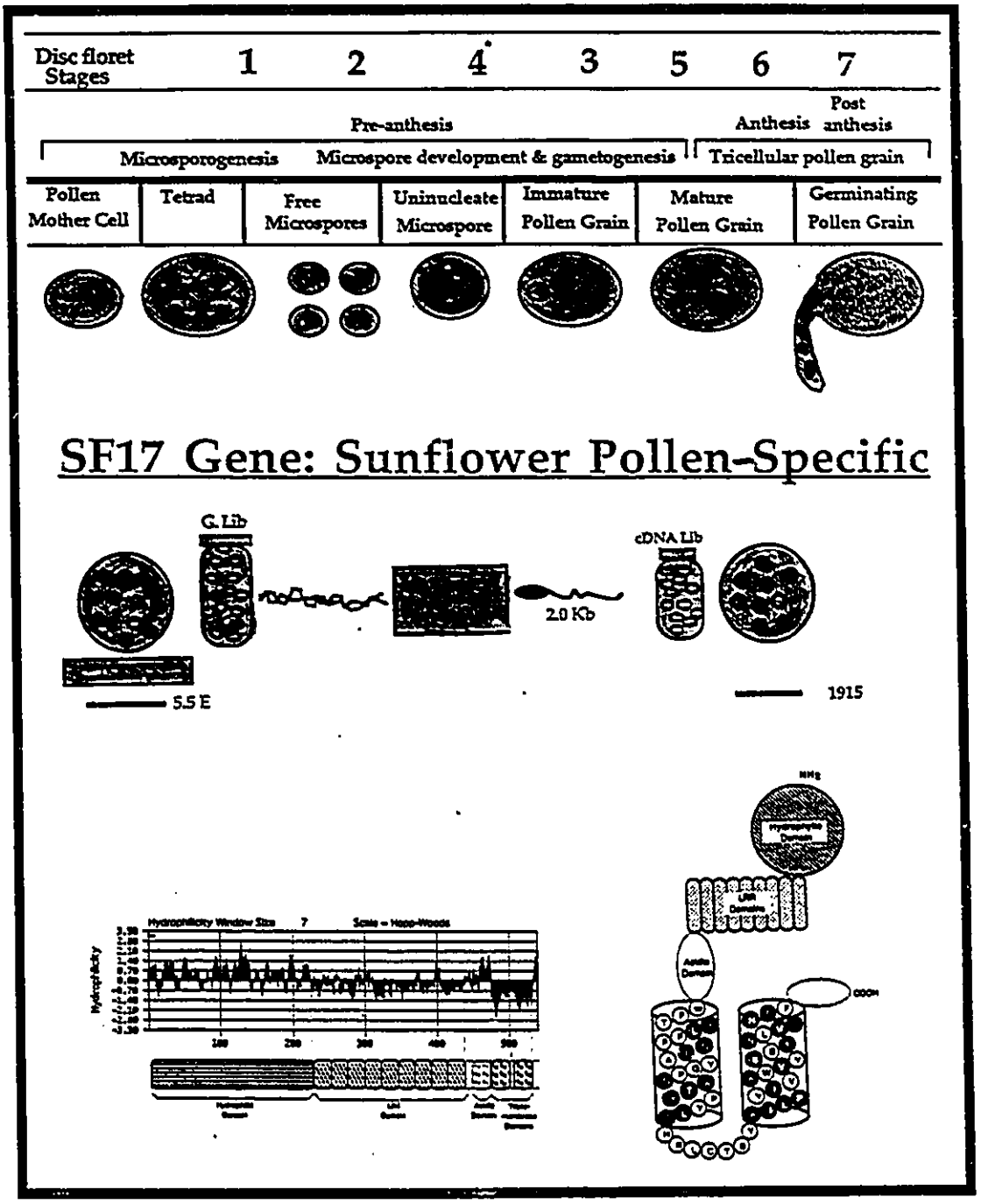
Anther is the male reproductive part of the flower, and contains reproductive as well as nonreproductive tissues. The major function of the anther is the production and release of pollen grains. Two major aspects of anther development are histospecification, cell degeneration and dehiscence (Goldberg et al., 1993). Dehiscence of anther occurs at the time of anthesis and is essential for the liberation of pollen grains. The dehiscence program in sunflower anther starts immediately after stage 1 (microspore release). Disc floret anthers from stage 2 onwards have a simple structure, sustained by the filament, consisting of a single layer of epidermis enclosing pollen grains. In the case of tomato (Bonner et al., 1989) differential growth of epidermal cells has been implicated in the rupture of anther. The epidermis defines the point of rupture (stomium) and the enlarged epidermal cells also generate force needed for the dehiscence. The process of dehiscence has not been characterized in sunflower. But it is reasonable to speculate the involvement of complex mechanical processes.

In male-sterile disc florets, the degeneration of the tapetum and microspore tetrads (the degeneration occurs immediately after meiosis) leads to the male sterility (Horner,1977). The anthers of CMS male-sterile florets generally do not project out of corolla, and are empty of pollen and enclosed by the epidermis. SF15 is the fourth epidermal gene reported. Unlike the previous three epidermal genes (SF2, SF18 and SF19) which are transcribed just

before and during anthesis, SF 15 begins to accumulate from very early stages after microspore release. Interestingly, for the above four epidermal genes reported so far, the expression patterns are either undetectable or very weak in case of male-sterile disc florets (in Northern blot hybridization experiments). It also needs to be pointed out that there is no dehiscence in male-sterile florets.

It is tempting to speculate that the role of SF15 proteins or its down stream products may be vital to the events that prepare the anther epidermis to dehisce. SF15 proteins are both vacuolar and secretory in their destination. It is known that the vacuole is a multifunctional organelle important in the regulation and maintenance of plant cell growth and development (Boller and Wiemken, 1986). The presence of homologous genes of SF15 in corn, tomato and cabbage also suggests that the encoded proteins might perform an important function(s) in the anther development of higher plants, explaining their evolutionary conservation. However, the role of SF15 in anther dehiscence has to be determined.

SF17


Disc floret Stages
1
2
4
3
5
6
7
Pre-anthesis
Anthesis
Post anthesis
Microsporogenesis
Microspore development & gametogenesis
Tricellular pollen grain
Pollen Mother Cell
Tetrad
Free Microspores
Uninucleate Microspore
Immature Pollen Grain
Mature Pollen Grain
Germinating Pollen Grain
SF17 Gene: Sunflower Pollen-Specific
G. Lib
cDNA Lib
2.0 Kb
5.5 E
1915

RESULTS

1. Organ specificity of SF17

A floral cDNA library constructed in λgt10 vector (with poly(A)$^+$ RNA from sunflower inflorescence at anthesis), was differentially screened. This led to the isolation of about 31 flower-specific cDNA clones (Herdenberger et al., 1990). SF17 was one such cDNA clone. Size fractionation on agarose gel indicated that the insert of the SF17 cDNA clone is about 300 bp. This insert was subcloned into pUC19 vector and sequenced completely. The size of insert was 329 bp and the sequence analysis revealed the presence of an open reading frame. In order to determine the organ specificity of SF17 gene, RNA gel-blot analysis was performed. Total RNA was isolated from various vegetative (roots and leaves) and reproductive organ systems (ovaries, corolla, pollen depleted anthers, styles+stigmas, mature pollen, immature seeds) of sunflower. The reproductive organs i.e. ovaries, corolla, anthers, styles+stigmas, mature pollen and immature seeds were harvested from male-fertile disc florets of 1-3 days after anthesis. Closed disc florets were harvested before anthesis. The total RNA was size fractionated on urea-agarose vertical gel and blotted to Hybond N+ membrane. The gel-blot was probed with labelled SF17 cDNA insert (329 bp). The probe hybridized to transcripts of mature pollen grains only (Fig. 1). The size of mRNA transcripts was about 1.9 Kb. The faint hybridization signal in the lanes of anthers and stigmas+styles, is due the contamination of pollen grains with these organ systems. Longer exposure of the autoradiogram did not reveal any hybridization signal in the lanes of other organs. The results revealed that SF17 gene is expressed in the mature pollen grains of sunflower. No hybridization

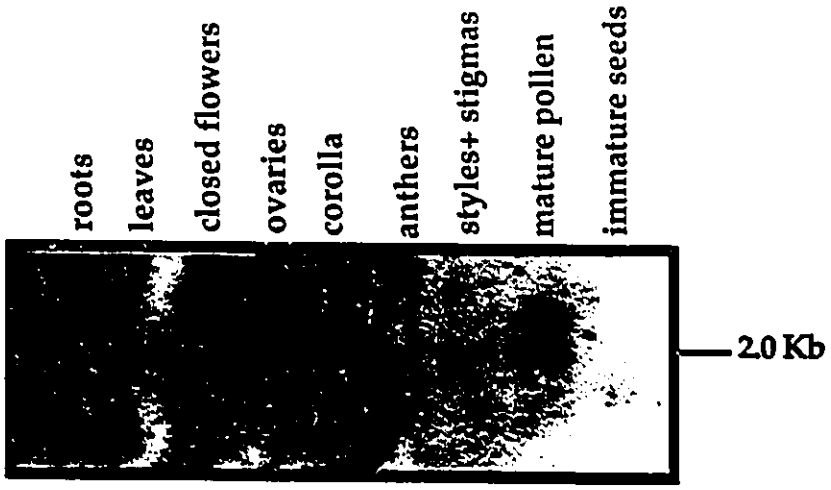


Figure1. Organ specificity of SF17 gene .
Total RNA from various organ systems was isolated, and RNA gel-blot analysis was performed (20 µg of total RNA from various organ systems and 7.0 µg from pollen grains). The RNA samples were separated in a vertical urea-agarose gel and blotted to Hybond-N+ membrane and probed with ^{32}P-labeled SF17 cDNA probe (329 bp). The probe hybridized to the transcripts of the pollen grains only. Faint hybridization signals, in other lanes (anthers and styles + stigmas) are due to the contamination of pollen grains in those organ systems.

signal was detected in Northern hybridization experiments with total RNA of corn pollen grains and anthers of tomato at reduced stringency (60 ^{0}C) (data not shown).

2. Isolation of full-length SF17 cDNA clone

In order to isolate a full-length SF17 cDNA clone, the sunflower floral cDNA library was screened with 329 bp insert of SF17 as a probe. About 0.5x10^{6} λgt10 recombinant cDNA clones were screened. The initial round of screening led to the isolation of 7 clones. Two of these clones (clones-21 and -23) were plaque purified to homogeneity. The DNA inserts of these two clones were size fractionated on 0.7% agarose gel, and the results revealed the presence of similar sized inserts of about 1.9 kb (Fig. 2A). One of these clones (clone-23) was selected for further investigation. The insert (of clone-23) was subcloned into EcoRI site of pUC19 vector (Fig. 2B) and sequenced completely. Sequence analysis revealed that the cDNA insert is 1915 bp long . It has a complete open reading frame, starting with the first ATG codon at base 175 and ending with TGA termination codon at base 1795 (Fig. 5). The size of the cDNA clone (1915 bp) closely approximates the size of mRNA transcripts of Northern hybridization experiments (1.9 kb in Fig. 1). The full-length SF17 cDNA clone has a small poly(A)$^{+}$ tail in the 121-nucleotide long untranslated region of the 3' end. The putative polyadenylation signal (AATAAA) is 18 bp upstream of the poly(A)$^{+}$ tail (Fig. 5).

3. Isolation and characterization of SF17 genomic clone (SF17G)

In order to isolate a SF17 genomic DNA clone, sunflower nuclear genomic DNA library was screened with 1915 bp insert (of full-length SF17 cDNA clone) as a probe. About 2.0x10^{6} λcharon 40

recombinant clones were screened. The initial round of screening led to the isolation of 7 clones. One genomic clone (SF17G) of 15.5 kb was isolated and purified to homogeneity. SF17G was restricted with EcoRI and EcoRI+BamHI and the restriction products were size fractionated on a 0.7% agarose gel and blotted to a Hybond N+ membrane (Fig. 3A). The blot was assayed with labeled SF17 cDNA insert as a probe (1915 bp). The probe hybridized to a 5.5 kb EcoRI fragment (Fig. 3B). The 5.5 kb EcoRI fragment was subcloned into pUC19 vector (Fig. 3C). Restriction mapping of 5.5 kb of SF17G was performed (Fig. 4A). The DNA sequence analysis revealed that the coding region is located in the 3.8 kb EcoRI and SalI fragment (Fig. 4A). The coding region of of the (SF17G) genomic clone is identical to the cDNA, but it is split into three exons of 741, 167, 715 bp and two introns of 371 and 246 bp (Fig. 4A). The introns are AT rich and their boundaries have well conserved GT/AG splice sites (similar to typical eukaryotic introns). Both introns contain a number of inverted repeats, which may play a role in the stabilization of their secondary structures. The genomic clone contains about 1.0 kb of 5'-regulatory sequence (Fig. 5).

The 5' upstream region contains a number of classical sequence motifs, CAAT box and GATA box (shown in bold letters in Fig. 5), but lacks the typical TATA box. In addition there are two pairs of inverted repeats (underlined in Fig. 5), and one pair of direct repeats of 18 bp in size (boxed in Fig. 5). These inverted repeats can form a hairpin structure and may be involved in the regulation of SF17 gene expression.

Fifteen C-boxes (XCCCCX) have been identified in the 5'-upstream region of SF17G (shown in bold letters in Fig. 5). At present their significance is unknown, but six similar C-boxes are located in the 5' upstream region of a pollen-specific maize polygalacturonase gene (W22476) (Allen and Lonsdale, 1993).

4. Mapping of transcription start sites of SF17 gene

Transcriptional initiation sites of SF17G were determined by primer extension analysis. A 25-mer oligonucleotide (3'-G CGG GAC TAG TTG AGT ATA TAT TTG-5') and total RNA isolated from sunflower pollen grains, were used. The primer is complimentary to nucleotides +74 to +98 of the open reading frame of SF17 genomic clone (Fig. 5). The end labeled primer was annealed to 35 µg of total RNA (of sunflower pollen grains) and extended by AMV reverse transcriptase. The primer extension products (PE) and the products of DNA sequencing reactions (A,C,G,T) with the same primer, were size fractionated on a 6% polyacrylamide-urea (8M) gel, and analyzed by autoradiography. The results show that there are two major and few minor transcription start sites (indicated by three and one asterisks respectively in Fig. 6). The two major transcription initiation sites (G and C, indicated by arrows in Fig. 5) are 38 nucleotides apart. The SF17 cDNA clone contained the second major initiation site (nucleotide C) but lacked the first major initiation site (nucleotide G) (Fig. 5). These results indicate that SF17 cDNA clone is incomplete: 16 bp are missing in the 5'-untranslated region (with reference to first major transcription initiation site).

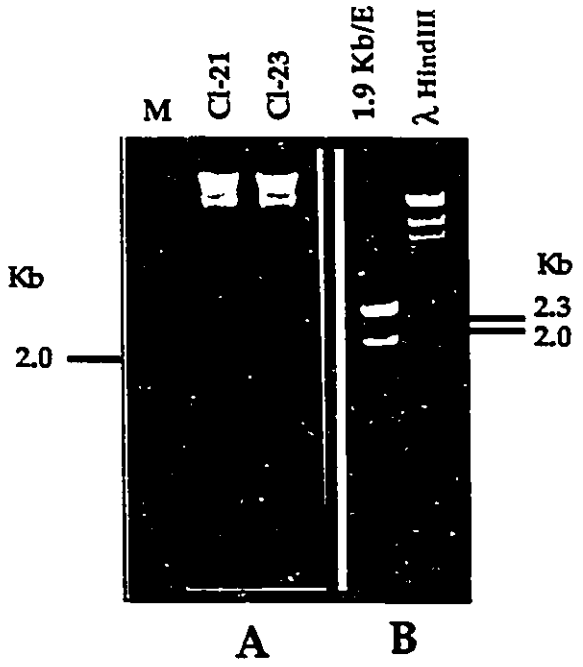


Figure 2. Isolation of full-length SF17 cDNA clones and subcloning. Sunflower floral cDNA library was screened with ^{32}P-labeled SF17 cDNA probe (329 bp). Two cDNA clones (Cl-21 and Cl-23 containing similar sizes of inserts ca. 1.9 Kb) were isolated and the insert from Cl-23 was subcloned into the EcoRI site of pUC19 vector. (A) Restriction analysis of λgt10 recombinant clones (Cl-21 and Cl-23) with EcoRI. The insert of Cl-23 was subcloned into pUC19. (B) Restriction analysis of the 1.9 Kb subclone (of Cl-23) with enzyme EcoRI. (M-molecular size marker; λ HindIII-restriction product of λ DNA with HindIII, another molecular size marker)

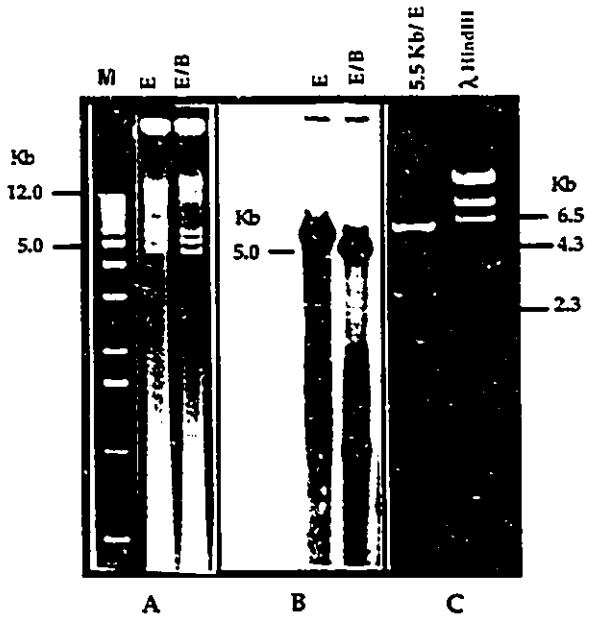


Figure 3. Isolation and restriction analysis of the genomic clone SF17G. About 2x 10^6 recombinant genomic DNA clones were screened, with ^{32}P-labeled SF17 cDNA (1915 bp) as a probe, and a genomic clone (SF17G) of ca. 15 Kb was isolated and purified to homogeneity. (A) Restriction analysis of SF17G with EcoRI (E) and EcoRI/BamHI (E/B). (B) DNA gel-blot analysis of SF17G restriction products of (A) with SF17 cDNA probe(1915 bp). The 5.5 Kb EcoRI fragment hybridizing with the probe was subcloned into pUC19 vector. (C) Restriction analysis of 5.5 Kb subclone with EcoRI enzyme (E). (M-molecular size marker; λ HindIII- λ DNA digested with HindIII, another molecular size marker).

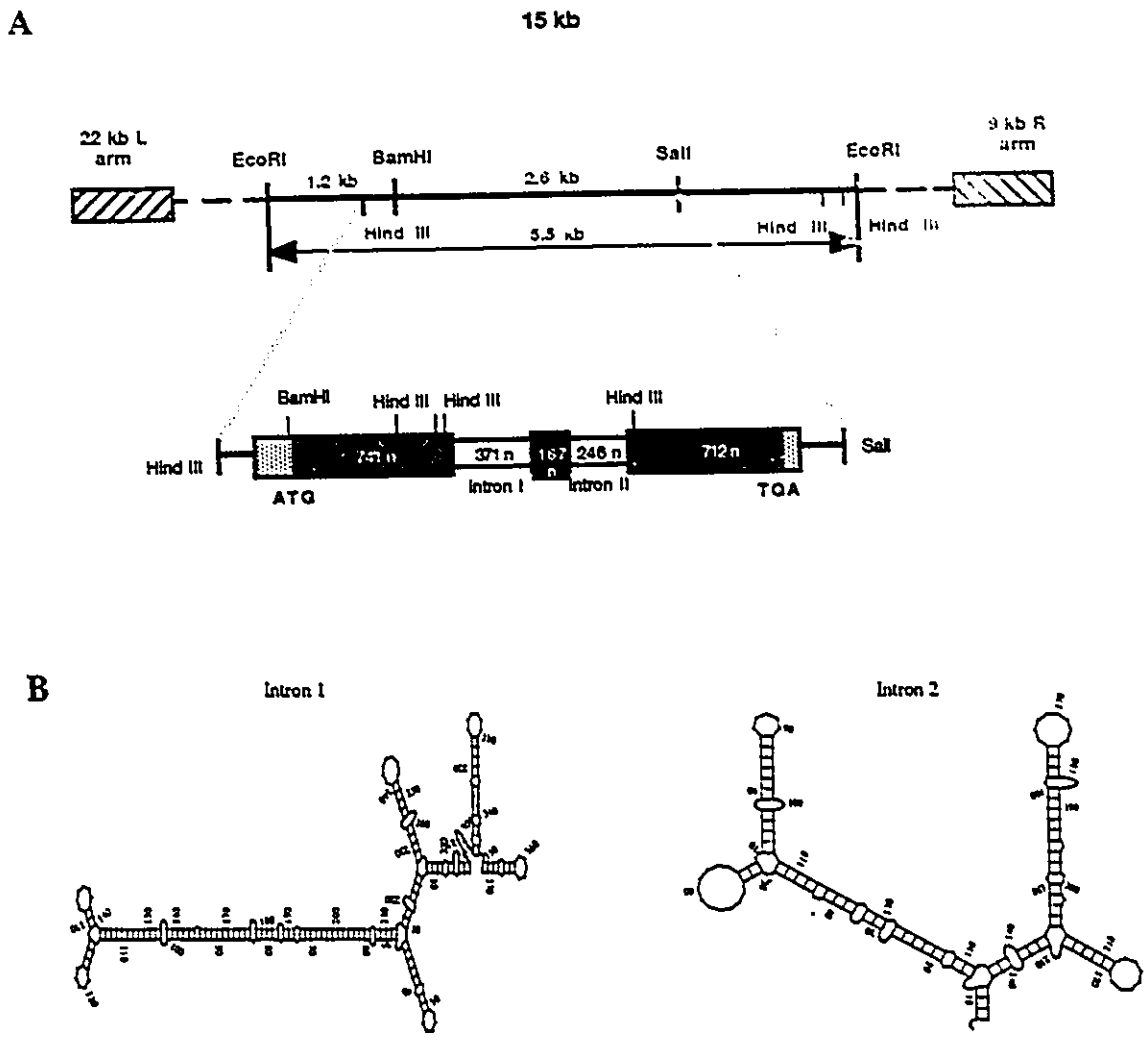


Figure 4. Structure of 5.5 Kb EcoRI subclone of SF17G.
(A) Restriction map of 5.5 Kb EcoRI subclone of SF17G. The open box indicates the introns and closed box indicates the exons and stippled boxes at the ends represent the λ Charon 40 vector arms. (B). Secondary structure of introns. Both the introns (intron1 and 2) contain a number of inverted repeats, which may play a role in stabilizing their secondary structure.

Figure 5. Comparison of nucleotide sequence of SF17G and SF17 cDNA clone. Both the clones are colinear (except for the presence of two introns in the genomic clone). The colinear sequence of both the clones is represented in upper case letters. The nucleotide sequence of introns, upstream and down stream region of genomic clone is represented in lower case letters. The translation start codon (ATG), translation termination codon (TGA), and the potential polyadenylation signals (AATAAA) are indicated by bold letters. The transcription initiation sites are indicated by arrows and asterisks. Two pairs of inverted repeats are underlined, and one pair of 18 bp direct repeat is boxed. Fifteen C-boxes (XCCCCX) and classic sequence motifs, i.e., CAAT box and GATA box, are shown in bold letters. (The nucleotide sequence data reported here will appear in the EMBL data bank under the accession numbers X 81997 and X 82089).

```
gaattcattt tttttaatgg atattattag cagtgtagaa actttgagaa cattttaaat aggttttgaa tggaaggcgt agtgtactag -922
aagatttaag atttaactga aatgtgtgga gttttttttta attaatattg tagtttaaat accttttcta aaacgcgtga tctttttttgg -832
aaccaacgat ttttcgtaaa cggatttttt aaacaggtct gtttttaacg agtgtagcga ttcccatgca cccaacgatg atgattcact -742
tgcctaatgc ttgtgagctc acccaagccc tgccccccag cccctcatgt gccaccagtg acagtgtggc cacgagcccg ccctacatgt -652
caacttccat gcccttatcg aagcctatta gggtcacccc ccactcccct aagaggggag tcccctctct tacgcacaac ccaataggtt -562
atgccacgtc aactcccctg cttaaactcc cctcacaccc ccatttgatg gcggcactcc cccttagat gacttgtttt ttatttaaaa -472
aaaaaaattc attggttgat tgaatggggg gccccacctc tctctcctcc cctcttcggt gccccctccc cgatcgttca ccccgatttg -382
gcctcccgcc tgatggcggc accgcccgga tcggtaaaaa tggtccccgc acgggtcacc gagggtgccc ctcaccctta gatttgtagc -292
tctaaaaggg tgatatgagt caaataatgt ttatatttgt acaagtttaa tagatacatg taattccatg tgttaaatat ttatatacta -202
cttaaatagt taaattaatt aagatgtatt aatttccgtt aaggaggggc ttaaaatgaa gcttaaaaat aattaaaaag ccatagccct -112
atttttggca tgaatgaatg aatgatatcc tttgaaagtt gtccagaaaa catggaagcc acggttgcaa aatttgtctt caaattgatg  -22

caacaatggc cattatttct ttGTCATCGT CCTTCAATTC TTATTGTTTG TTGCCTTCTT CGTTCTTTCT CCCGTTCGGA TGAGACATAG   68
                                           TC TTATTGTTTG TTGCCTTCTT CGTTCTTTCT CCCGTTCGGA TGAGACATAG   52

CGTGACGCCC TGATCAACTC ATATATAAAC AACCCCCAAA GCGTTACGTT CATTGATCTC TCTGAACACC ATCACCACCA TCATCCTCCT  158
CGTGACGCCC TGATCAACTC ATATATAAAC AACCCCCAAA GCGTTACGTT CATTGATCTC TCTGAACACC ATCACCACCA TCATCCTCCT  142 cDNA

CCTCATTCCT CATCTAAATC ATTATCATCG CCATGGATCC CAACCCCAGC CCAAGGAAGA TGCGAATCCT CAAATATGTC ATGACCAAAA  248
CCTCATTCCT CATCTAAATC ATTATCATCG CCATGGATCC CAACCCCAGC CCAAGGAAGA TGCGAATCCT CAAATATGTC ATGACCAAAA  232 cDNA

TCCCCAGCTT CAAACGCCGG CAATTGCAGG AGCTTGAGAA CGAAAGACTA GCTCGCCCAT ACGCCTTGCC AGAAACCTAC ACCGAGCTAT  338
TCCCCAGCTT CAAACGCCGG CAATTGCAGG AGCTTGAGAA CGAAAGACTA GCTCGCCCAT ACGCCTTGCC AGAAACCTAC ACCGAGCTAT  322 cDNA

CTGAACGTGA ATCATATACC GAGTTAGCCG AACGTGTGTC TCATTTAACA GAAGATGATA TTCTAGCCAA CATCCCTGCG GTTGTTGTTG  428
CTGAACGTGA ATCATATACC GAGTTAGCCG AACGTGTGTC TCATTTAACA GAAGATGATA TTCTAGCCAA CATCCCTGCG GTTGTTGTTG  412 cDNA

AGGTTAAACA GATCCGTTCG GTTATTAAAT CTCTAGGTGA CCGTCCCGAC CCAGAAACCG TTGATCTGGC TAGGTTGCGG TACCGAGAAG  518
AGGTTAAACA GATCCGTTCG GTTATTAAAT CTCTAGGTGA CCGTCCCGAC CCAGAAACCG TTGATCTGGC TAGGTTGCGG TACCGAGAAG  502 cDNA

CCGAATCTCC AGTCGCGGGT CAATTCGATG AGAATGCTGA ATACGATATG GAGATGGAGA AGAGGAAGAG ATTGGTGAAA AGAGAGAGCC  608
CCGAATCTCC AGTCGCGGGT CAATTCGATG AGAATGCTGA ATACGATATG GAGATGGAGA AGAGGAAGAG ATTGGTGAAA AGAGAGAGCC  592 cDNA

AGATGTATAA GGCATTGATT TCTTTAGATG AGATGCATGA AACGTATAGC GATCTGCTTG TTGTCGCCGA GAGGAGGCTT CAGAAGCTTT  698
AGATGTATAA GGCATTGATT TCTTTAGATG AGATGCATGA AACGTATAGC GATCTGCTTG TTGTCGCCGA GAGGAGGCTT CAGAAGCTTT  682 cDNA

ATGATACGGC TAAATCAGCT GGGAAGTTAT CCGCGCTTGA TAAGAGGGTG TCTTCTATGC TTCCGACGAT AGCTGAGGAG GTGAAGGAGG  788
ATGATACGGC TAAATCAGCT GGGAAGTTAT CCGCGCTTGA TAAGAGGGTG TCTTCTATGC TTCCGACGAT AGCTGAGGAG GTGAAGGAGG  772 cDNA

AGATGGCGGA TATTCTTCAG GATGCGTTGA TGAACGGTGT TGAACGGATT GATTTGTCGC GCAGGCCGTT GCCGTTTGTT CCTGAAGCTT  878
AGATGGCGGA TATTCTTCAG GATGCGTTGA TGAACGGTGT TGAACGGATT GATTTGTCGC GCAGGCCGTT GCCGTTTGTT CCTGAAGCTT  862 cDNA

TTGGGAAGCT TCATACATTG GTTTCTCTTG ACCTCTCCTC CAATAAACTT ACGgttagtt gttttgcacc atcaacgttc tcaaaagttg  968
TTGGGAAGCT TCATACATTG GTTTCTCTTG ACCTCTCCTC CAATAAACTT ACG                                         915 cDNA

aaaacgtttt ttagctgagg tagcaggtga catgggaaaa tgaccatctg ctgtttaagt agcacgtctc aaattagtta atgtaattaa 1058
aaagttcaat cttctaattt gagacatcca ctacttcaag cagcggatgt cgtggtcact ttttgtcatg tcacctgcta tcggtcagat 1148
tacacctgca tggtctcata aaccaaacat gtttagaaac ttgaccataa accatataaa aaaacctaaa aaataaaata aaaaaaggat 1238

tccatcttca tatatatgtt cttatatttg aattattgac gtttgcttac aatcgttgtt gcagGCTATA CCAGAATCAC TTGCGGGACT 1328
                                                                     GCTATA CCAGAATCAC TTGCGGGACT  941 cDNA

TACGAGTCTC GAGGAGCTTA ATCTTTCTGC TAATCTCTTT GAATCACTAC CAGATACCAT CGGCTCGTTG CAACATTTGC AGTTTCTAAA 1418
TACGAGTCTC GAGGAGCTTA ATCTTTCTGC TAATCTCTTT GAATCACTAC CAGATACCAT CGGCTCGTTG CAACATTTGC AGTTTCTAAA 1031 cDNA

TGTCTCAAGG AACAAACTCA CTTCCTTGCC CGATGGCATT TGTAAATGCA Cgtaattgct taatccgtag gaaaacttgt gtcagaaatg 1508
TGTCTCAAGG AACAAACTCA CTTCCTTGCC CGATGGCATT TGTAAATGCA C                                           1082 cDNA

cgtaggaaaa gttacaaaaa aaaaataaaa aaaataaaaa gggggcaaaa tgccttttca ctaatttttt tcccgcaatg gtttttccga 1598
cgatacttct tttgtgagaa aagggtgttt cctaggcctt atgacctatt ttagcgtaga aacgtcccag ttatctatta gtgtcaagaa 1688

ctaattaaat atcacatgtt ggtgcagATC ATTATTAGAA TTGCATGCTA GTTTCAATCA GATTACATAT CTTCCTGCCA ACATTGGGTA 1778
                             ATC ATTATTAGAA TTGCATGCTA GTTTCAATCA GATTACATAT CTTCCTGCCA ACATTGGGTA 1145

TGGACTAATC AATCTAAAGA AGCTTATAAT GCCTTTGAAC AACGTGCGCT CTCTTCCTAC CTCCATCGGC GAAATGATAT CTCTTCAAGT 1868
TGGACTAATC AATCTAAAGA AGCTTATAAT GCCTTTGAAC AACGTGCGCT CTCTTCCTAC CTCCATCGGC GAAATGATAT CTCTTCAAGT 1235 cDNA

CCTTGATGTG CACTTCAATA CGCTCAGAGG CCTGCCACCC TCAATCGGAA TGCTTAAAAA ACTCGAGGTT CTTAACTTGG GCAGTAACTT 1958
CCTTGATGTG CACTTCAATA CGCTCAGAGG CCTGCCACCC TCAATCGGAA TGCTTAAAAA ACTCGAGGTT CTTAACTTGG GCAGTAACTT 1325 cDNA

TAACGACTTT ACAGCACTCC CCGAAACAAT CGGGAGTCTA ACTCGCCTAA GGGAACTCGA TATATGCAAC AATCAGATTC AACAGCTGCC 2048
TAACGACTTT ACAGCACTCC CCGAAACAAT CGGGAGTCTA ACTCGCCTAA GGGAACTCGA TATATGCAAC AATCAGATTC AACAGCTGCC 1415 cDNA

CATTACATTT GGGCGGTTGG TGAGCCTGAC CAGACTGGTG GTGGACCATA ACCCGTTGAC GGTGTCACCA CCAGAGGTGG TGGCAGAAGG 2138
CATTACATTT GGGCGGTTGG TGAGCCTGAC CAGACTGGTG GTGGACCATA ACCCGTTGAC GGTGTCACCA CCAGAGGTGG TGGCAGAAGG 1505 cDNA

TGTTGAAGCT GTGAAGGTGT ATATGAGCAA GAGGTTATAT GATATGATTG TAGAAGAAGA GAGAAGGGTG ATGTGGGAAC GGGAAGAACA 2228
TGTTGAAGCT GTGAAGGTGT ATATGAGCAA GAGGTTATAT GATATGATTG TAGAAGAAGA GAGAAGGGTG ATGTGGGAAC GGGAAGAACA 1595 cDNA

GGCACAACAA GCTGGGTGGT TTACGTTTTT GTTTTTCGGG CCGGCTTATG GCCCCGGCGC AACGTATCCT TACCTCACTC ATCGACTTTG 2318
GGCACAACAA GCTGGGTGGT TTACGTTTTT GTTTTTCGGG CCGGCTTATG GCCCCGGCGC AACGTATCCT TACCTCACTC ATCGACTTTG 1685 cDNA

TACAAGTTAT AAAATCCTAC CATATATAGT AGCTTGGGTT TATTCTTCGT TCAACCTTGT TATTCATTGT TTTTTAAAAT GTTATGTTTT 2408
TACAAGTTAT AAAATCCTAC CATATATAGT AGCTTGGGTT TATTCTTCGT TCAACCTTGT TATTCATTGT TTTTTAAAAT GTTATGTTTT 1775 cDNA

TTCGGACCGG AAACAAACCT GAAAACATGT ATGTAAAGGT TGCACGCAGG ATGGATACAT GTAACATCCT AAATTTTATT CCATTCCATG 2498
TTCGGACCGG AAACAAACCT GAAAACATGT ATGTAAAGGT TGCACGCAGG ATGGATACAT GTAACATCCT AAATTTTATT CCATTCCATG 1865 cDNA

TGTTAAGGAC GATTTATTGA AATAAATGGC CTCCTTTATT TGTTgatgtg tttgatggag aaattctaat tgttaatata ttttatatgt 2588
TGTTAAGGAC GATTTATTGA AATAAATGGC CTCCTTTATT TGTTAAAAAA                                             1915 cDNA

tattttatac ataaattata aaaccacgta tctaggattt tgtccaggga tagttagcaa aacaatcaag gttgaccata acttttcaaa 2678
aaggggaatt tacgaaaatc tgcgttgacc aaccatgttt tcgaatttgt cacgctaatg cgttcgtgga aggacctcta agtatgggat 2768
gcgtcgac                                                                                           2776
```

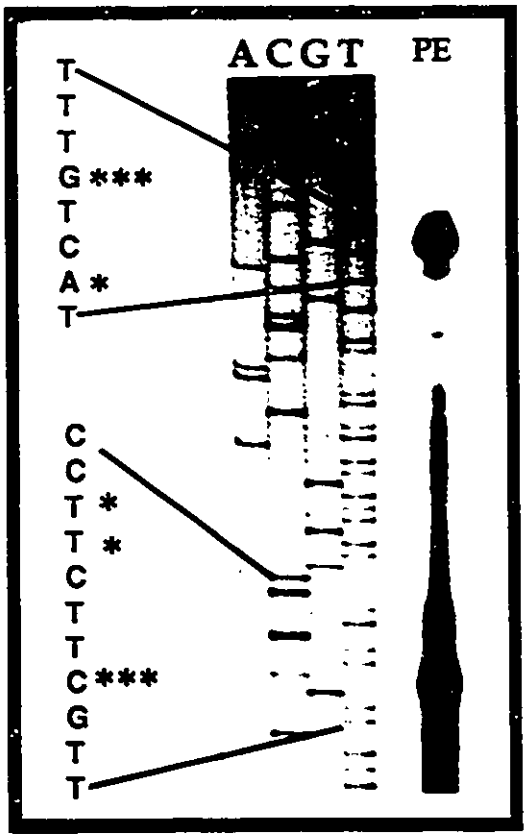


Figure 6. Mapping of transcription initiation sites of SF17. Primer extension was performed with a ^{32}P-labeled 25-mer oligonucleotide (3'-G CGG GAC TAG TTG AGT ATA TAT TTG-5'). The primer is complimentary to nucleotides +74 to +98 of the ORF of SF17 genomic clone (Fig. 5). The end labeled primer was annealed to 35 µg of total RNA (of pollen grains) and extended by reverse transcriptase. The primer extension products (PE) and the products of sequencing reactions (A,C,G,T) with the same primer, were size fractionated on a 6% polyacrylamide-urea (8M) gel, and analyzed by autoradiography. The nucleotide sequence around the transcription start site is shown. Asterisks indicate the positions of major and minor initiation sites

5. Structural features of SF17 protein

Nucleotide sequence of the open reading frame (ORF) of SF17 encodes a 540 amino acids long polypeptide (Fig. 7A). The putative polypeptide of SF17 is leucine rich (14%) with a molecular mass of 61.7 kDa and an isoelectric point (pI) 6.2. Hydropathy analysis of deduced SF17 protein (Hopp-Woods) reveals the presence of five major structural features (Fig. 7B and Fig.10) i.e., a hydrophilic domain, a leucine-rich-region (LRR) domain, an acid domain and two trans-membrane domains.

Hydrophilic domain: The hydrophilic domain is present in the N-terminal part of the protein (a.a. 1-226). Analysis revealed no evidence of signal peptide at the N-terminal. Of the three potential N-linked glycosylation sites (NXS) (underlined in Fig. 7A), one is located close to N-terminus of the polypeptide (a.a residues 4-6).

LRR domain: The second structural region of SF17 protein (a.a residues 227-438) contains nine leucine-rich imperfect tandem repeats (LRR) (Fig. 7B; Fig. 8A). The basic repeat unit of LRR, consists of 24 amino acids and is defined by the conservative positions of proline (at position 1), isoleucine (at position 4), leucine (at position 8,11,14 and 24), serine (at position 17), asparagine (at position 19 and some times at 15) and aliphatic amino acids (at positions 16 and 21) (Fig. 8A). The LRR domains have been reported in various organisms. In case of humans PR1, the placental glycoprotein, mediates the high affinity interaction with RNAse A (Lee and Valle, 1990). In humans, the leucine rich *rsp-1* (RSU-1) is capable of suppressing v-RAS transformation (Cutler et al., 1992). The α and β chains of human platelet glycoprotein GPIb, mediate the interaction with von Willebrand factor (Lopez et al., 1988). The rat p34 protein, characteristic of rough microsomes is believed to play a role in ribosome-membrane association. In the case of *Drosophila*, chaoptin

mediates the adhesive interaction between cells (Reinke et al., 1988), and the*Toll* protein mediates the dorsal-ventral polarity in the embryo (Hashimoto et al., 1988). In *Trypanosoma brucei*, the VSG expression site encodes a leucine -rich- repeat protein ESAG8 (Revelard et al., 1990). The catalytic domains of adenylate cyclases of both *Schizosaccharomyces pombe* (Young et al., 1989) and *Saccharomyces cerevisiae* (Kataoka et al., 1985) contain tandemly repeated motifs, rich in leucine. In plants many proteins implicated in a variety of functions, have also been shown to contain leucine rich tandem repeats (Chang et al., 1992; Valon et al., 1993; Mindrinos et al., 1994; Mu et al., 1994; Stotz et al., 1994 and Whitham et al., 1994). The LRR domain of SF17 protein contains two potential N-linked glycosylation sites (NXS) (shown in red in Fig. 7A).

Acid domain and two trans-membrane domains: The C-terminal region of SF17 protein contains two hydrophobic domains (a.a. residues 483-500 and a.a. residues 508-537). These hydrophobic segments are preceded by a cluster (a.a. residues 458-479) of mainly negatively charged residues (acid domain) (Fig. 7B)

6. Sequence similarity of SF17 polypeptide to other proteins

The deduced amino acid sequence of the full-length SF17 protein was compared with the data bases using the BLAST program (Altschul et al., 1990). No striking sequence similarity was found in the hydrophilic domains of SF17 protein with any other protein present in the data bases. However a striking resemblance of LRR domain was found with a series of proteins, including human *rsp-I* protein (42% identity; 58% similarity), *D. melanogaster* flightless-I gene (31% identity; 49% similarity), rat-ribosome binding protein p34 (37% identity; 57% similarity), and yeast adenylate cyclase (33% identity; 52% similarity) (Fig. 9a).

a)

```
MDPNPSPRKM RILKYVMTKI PSFKRRQLQE LENERLARPY ALPETYTELS ERESYTELAE  60
RVSHLTEDDI LANIRAVVVE VKQIRSVIKS LGDRPDPETV DLARLRYREA ESPVAGQFDE 120
NAEYDMEMEK RKRLVKRERQ MYKALISLDE MHETYSDLLV VAERRLQKLY DTAKSAGKLS 180
ALDKRVSSML PTIAEEVKEE MADILQDALM NGVERIDLSR RRLPFVPEAF GKLHTLVSLD 240
LSSNKLTAIP ESLAGLTSLE ELNLSANLFE SLPDTIGSLQ HLQFLNVSRN KLTSLPDGIC 300
KCRSLLELDA SFNQITYLPA NIGYGLINLK KLIMPLNNVR SLPTSIGEMI SLQVLDVHFN 360
TLRGLPPSIG MLKKLEVLNL GSNFNDFTAL PETIGSLTRL RELDICNNQI QQLPITFGRL 420
VSLTRLVVDH NPLTVSPPEV VAEGVEAVKV YMSKRLYDMI VEEERRVMWE REEQAQQAGW 480
FTFLFFGPAY GPGATYPYLT HRLCTSYKIL PYIVAWVYSS FNLVIHCFLK CYVFSDRKQT 540
```

b)

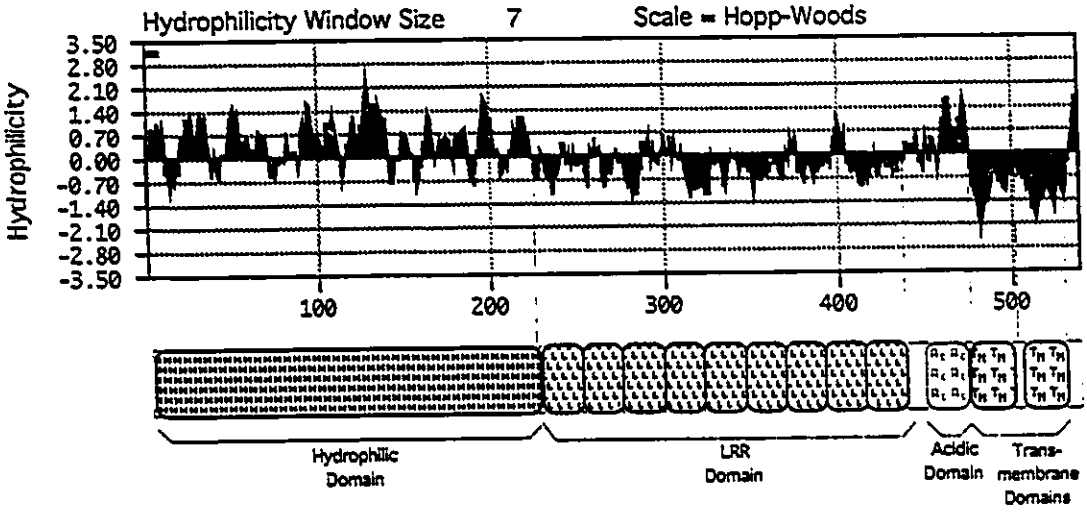


Figure 7. Amino acid sequence and hydropathy plot of SF17 protein. (a) The predicted amino acid sequence of SF17 protein. The polypeptide is 540 a.a. long. Three putative N-linked glycosylation sites (NXS) are shown in red. (b) Hydropathy plot (Hopp-Woods) and linear model of deduced SF17 protein. SF17 protein consists of five major structural features, a hydrophilic domain, a Leucine-Rich-Repeat (LRR) domain, an acid domain, and two trans-membrane domains.

Figure 8. The SF17 Leucine-Rich-Region (LRR).
(A) The alignment of 9 leucine rich tandem repeats. The second structural region of SF17 protein (residues 227-438) contains nine leucine-rich imperfect tandem repeats (LRR). The basic repeat unit consists of 24 amino acids and is defined by the conservative positions of proline (at position 1), isoleucine (at position 4), leucine (at position 8,11,14 and 24), serine (at position 17), asparagine (at position 19 and some times at 15) and aliphatic amino acids (at positions 16 and 21). (B) Comparison of LRR motifs of SF17 protein and other members of LRR super family. The members of the superfamily include: rat ribosome-binding protein p34 (Ohsumi et al., 1993); yeast adenylate cyclase (Kataoka et al., 1985); human carboxy peptidase N (Tan et al., 1990); *Arabidopsis* receptor like protein kinase RLK5 (Walker, 1993); *Petunia* pollen receptor like kinase PRK1 (Mu et al., 1994); *Arabidopsis* transmembrane kinase TMK1(Chang et al., 1992); *Arabidopsis* transmembrane kinase like protein TMKL1 (Valon et a l., 1993); *Arabidopsis* disease resistance protein RPS2 (Mindrinos et al., 1994); Tobacco resistance gene N (Whitham et al., 1994); *Drosophila* dorso-ventral polarity mediating protein Toll (Hashimoto et al., 1988).

A

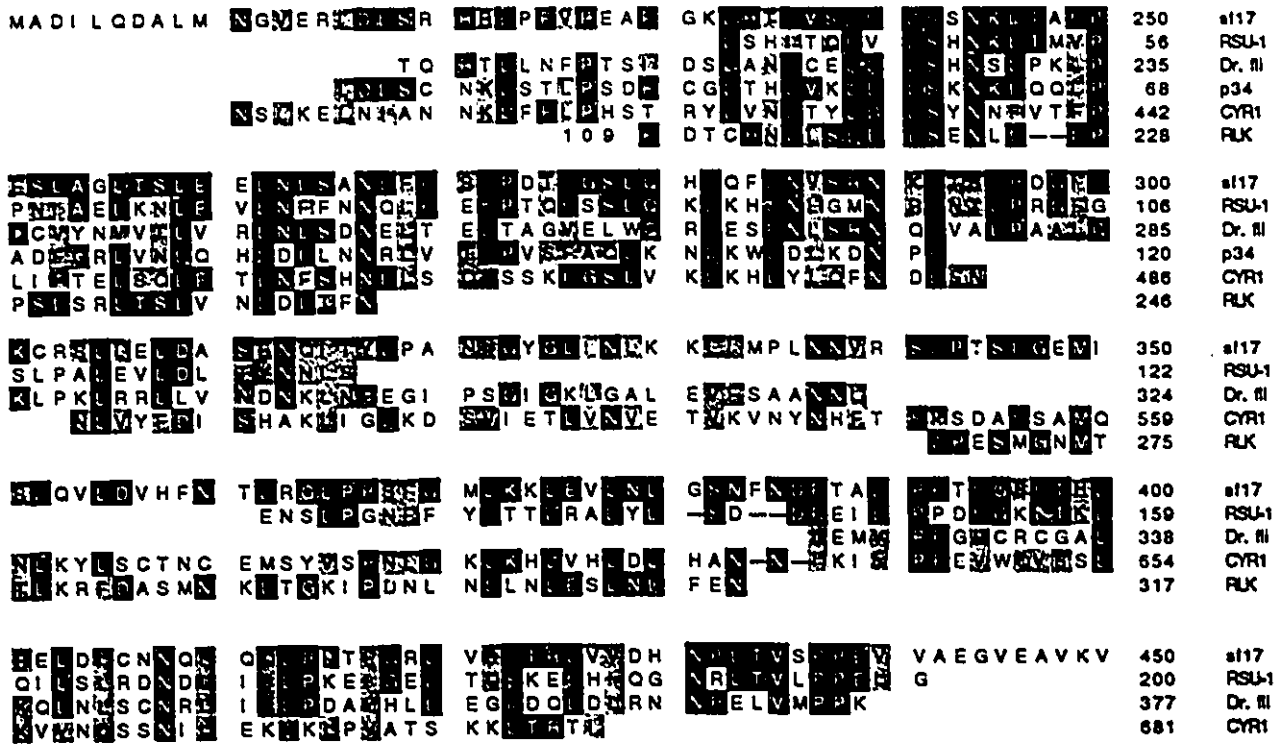
	227	P E A E G K - L H T L V S L D L S S N K L T A I
	250	P E S L A G - L T S L E E L N L S A N L E E S L
	273	P D T I G S - L Q H L Q F L N V S R N K L T S L
	296	P D G I C K - C R S L L E L D A S F N Q I T Y L
	319	P A N I G Y G L I N L K K L I M P L N N V R S L
	343	P T S I G E - M I S L Q V L D V H F N T L R G L
	366	P P S I G M - L K K L E V L N L C S N F N D F T A L
	391	P E T I G S - L T R L R E L D I C N N Q I Q Q L
	414	P I T E G R - L V S L T R L V V D H N P L T V S
Consensus		P X X I X X - L X X L X X L N/D a S X N X a X X L

B

sunflower SF17	P--I-- L--L--L N/D aS-N-a--L
rat p 34	P---F--L--L--LDLS-N-L--L
adenylate cyclase	P--a --L--L--L aS-N-a--a
c-peptidase N	P--aF--L--L--L-L--N-L--L
RLK5 Arabidopsis	P--L --L--L--L-L--N-LSG-I
PRK1 Petunia	L --L--L--L-L-NN---G-IP
TMK1 Arabidopsis	L --L--L--L-L--N-a-G-aP
TMKL1 Arabidopsis	--I -----L-SL-L--N-LSG-LP
RPS2 Arabidopsis	P--a-- L--L--L-a----a--a
Tobacco N gene	P--a --L--L--L-L-----L--L
Toll	P--LF-H--NL--L-L--N-L--L

Figure 9. (a) Amino acid sequence comparison of SF17 protein with human RSU1 (Cutler et al., 1992); *Drosophila* flightless-I protein Campbell et al., 1993); rat ribosome-binding protein p34 (Ohsumi et al., 1993); yeast adenylate cyclase CYR1 (Kataoka et al., 1985); and *Arabidopsis* RLK (Walker, 1993). Black boxes indicate conserved amino acids between two or more proteins and grey boxes indicate similar amino acids. (b) Comparison of amino acid sequence of the C-terminal region of SF17 protein with the transmembrane domains of rat olfactory receptor protein, OLF17 (Buck and Axel, 1991). Vertical lines indicate identical amino acids and asterisks indicate similar amino acids.

a



b

```
VAEGVEAVKV YMSKRLYDMI VEEERRVMWE REEQAQQAGW FTFLFFGPAY 490 SF17
   |*  ||| **  ||                                  |*  ||
   GISMVKV FLISRL 171                              FILLGPLS 217 OLF17

GPGATYPYLT HRLCTSYKIL PYIVAWVYSS FNLVIHCFLK CYVFSDRKQT 540 SF17
  ||*|  *|        |** * * | * *  || *|*|
VTGASYMAIT        KLV SVLYAVIVPL FNPIIYC               296 OLF17
```

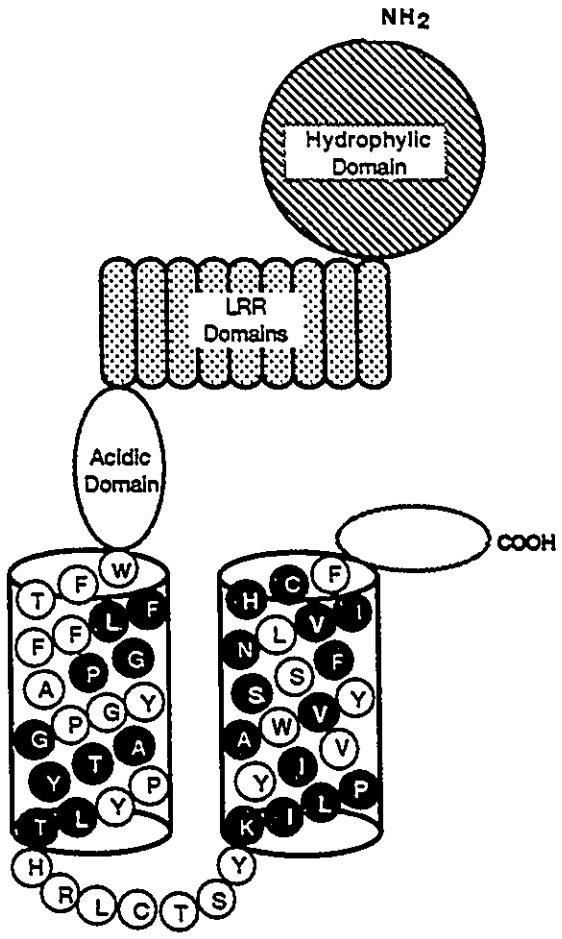


Figure 10. Structural model of SF17 protein.
The amino-terminal region contains the hydrophilic part of the protein as well as the nine leucine-rich repeats (LRR). The carboxyl-terminal contains an acid domain and two hydrophobic domains. The SF17 protein resides in a membrane with two membrane-spanning domains (shown in vertical cylinders) and its amino-terminal region is disposed in the cytoplasm. The conserved amino acids are shown in black circles and similar amino acids are shown in red circles (Comparison with the rat olfactory protein OLF17)

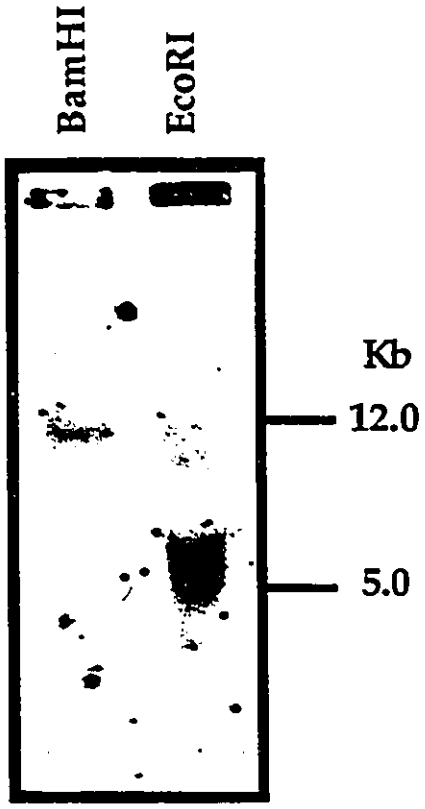


Figure 11. Southern blot analysis of sunflower nuclear genomic DNA. Nuclear genomic DNA (20 μg) was restricted with BamHI and EcoRI, fractionated on agarose gel, transferred to Hybond-N$^+$ membrane and probed with ^{32}P-labeled SF17 cDNA probe (1915 bp). Hybridization and washings were performed at 65 ^{0}C .

Alignment of SF17 LRR domain and receptor-like protein kinase gene of *A. thaliana* revealed three stretches of conserved sequences with 58-42% identity (64-50% similarity) (Fig. 9a).

In addition, the data base searches also revealed significant sequence homology between C-terminal domain of the SF17 protein and the rat olfactory receptor protein OLF17 (46-30% identity; 69-55% similarity) (Fig. 9b). The homology was found in the trans-membrane domain of OLF17.

7. Genomic copy number of SF17 gene

In order to determine the copy number of SF17 gene, sunflower nuclear genomic DNA was isolated and restricted with BamHI and EcoRI. The products of restriction digestion were size fractionated on a 0.7% agarose gel, and transferred to a Hybond N^+ membrane. The DNA gel-blot was assayed with labeled SF17 cDNA insert (1915 bp). The probe hybridized strongly to two restriction fragments of EcoRI and to a single fragment of BamHI (Fig. 11). These results reveal that SF17 gene exists in few copies (one or two) in the sunflower nuclear genome. Southern blot analysis, performed with nuclear genomic DNAs of safflower, cabbage, pea, tomato and corn at low stringency, did not show any hybridization signal (data not shown).

DISCUSSION

1. SF17 gene and protein

Northern blot analysis revealed (Fig. 1) that SF17 gene is pollen-specific. Previously two sunflower pollen-specific cDNA clones (SF3; SF16) were reported (Baltz et al., 1992a,b; Natalia et al., 1994a,b). SF17 is the third sunflower pollen-specific gene reported to date. The expression pattern of SF17 is similar to those of SF3 and SF16. Southern blot analysis demonstrates that, like other pollen specific genes, SF17 is present in few copies in the nuclear genome.

SF17 gene encodes a protein of 540 amino acids. The SF17 protein consists of five domains: an N-terminal hydrophilic domain, a Leucine-Rich Repeat (LRR) region, an acid domain and two trans-membrane domains (Fig. 7). Data base search using the amino acid sequence of entire SF17 protein revealed no striking similarity to other proteins. However the LRR region of SF17 protein showed similarity to LRR domains of a wide spectrum of eukaryotic proteins. A consensus sequence for this leucine-rich repeat is found in a number of diverse proteins from mammals (Takahashi et al., 1985 Lopez et al., 1988; Schneider and Schweiger,1991), *Drosophila* (Hashimoto et al., 1988; Reinke et al., 1988), to yeast (Kataoka et al., 1985). The LRR region contains 9 imperfect tandem repeats of 24 amino acids (Figure. 8A). These repeats are present in many different classes of proteins and may be extracellular or cytoplasmic.

Recently several proteins containing LRRs are characterized in higher plants: *Arabidopsis thaliana RPS2* gene that confers disease resistance to the bacterial pathogen *Pseudomonas syringae*

(Mindrinos et al., 1994), resistance gene *N* of tobacco that mediates resistance to the tobacco mosaic virus (Witham et al., 1992) and *TMK1* gene of *Arabidopsis thaliana* encoding a putative receptor-like transmembrane kinases (Chang et al., 1992; Valon et al., 1993; Walker, 1993). A receptor-like kinase (PRK1) specific to the male reproductive tissue with highest level in mature pollen was isolated and characterized in *Petunia inflata* (Mu et al., 1994). These proteins found in plants contain similar but not identical LRR motifs that lack one or more of the defining elements present in the SF17 protein. Figure 8B, shows the alignment of the consensus sequence of the LRR motif of SF17 with ten other LRR-containing proteins. The consensus sequence derived for the aligned LRRs of SF17 is identical to that found in the regulatory region of the yeast adenylate cyclase (CYR1) and the rat ribosome-binding protein (p34).

2. Functional implications

The LRR domains are thought to be involved in specific protein-protein interactions (Nose et al., 1992). The LRRs of human glycoprotein GPIbα bind to von Willebrand factor (Wicki and Clemetson, 1985; Handa et al., 1986). Binding of the von Willebrand factor to the LRR region of human glycoprotein Ibα heterodimer, induces adhesion of human platelets to the subendothelium (Titani et al., 1987) and the LRR of human placental ribonuclease inhibitor binds pancreatic ribonuclease (Lee and Vallee, 1990). In yeast LRRs of adenylate cyclase are involved in the interaction with *ras* protein and are necessary for adenylate cyclase activation (Suzuki et al., 1990). In addition to protein-protein interactions, LRRs could mediate interactions between the protein and cellular membranes (Thakahasi et al., 1985; Kataoka et al., 1985). The presence of LRRs in the SF17 protein suggests that it might act through a protein-protein contact.

The different LRR-containing proteins have one common feature: most are membrane-associated proteins. SF17 protein may be no exception from the rule. The C-terminus of SF17 protein contains two hydrophobic segments long enough to span the membrane bi-layer. These trans-membrane domains have significant sequence homology with transmembrane domain (5 and 7) of the rat olfactory receptor protein OLF17 (Fig. 9a) (Buck and Axel, 1991). The membrane-spanning sequences are preceded by a cluster of mainly negatively charged residues, called acid domain (Fig. 7B), which can provide a barrier to translocation. The amino acid sequence of rat ribosome binding-protein p34 (a member of LRR protein super family, Ohsumi et al., 1993) contains a single hydrophobic region and lacks an N-terminal signal sequence. The SF17 protein contains two hydrophobic domains and lacks an N-terminal signal peptide. Protein translocation across the membrane bilayer also occurs post-translationally, in addition to the co-tanslational translocation mediated by N-terminal signal peptides (Sanders and Schekman, 1992). SF17 protein shares conserved features of two transmembrane domains of rat odorant receptor OLF17 (Fig. 9B). The OLF17, like other members of the receptor superfamily, displays no evidence of an N-terminal signal peptide (Buck and Axel, 1991). On the basis of the topology of rat p34, rat OLF17, and other newly found topological features (Ohsumi et al., 1991), the proposed structural model of SF17 protein is presented in Figure 10. It is postulated that SF17 protein resides in a membrane with two membrane-spanning segments between amino acid residues 483-500 and 508-537. Its N-terminal side and LRR region being disposed in the cytoplasm like rat p34 polypeptide.

The pollen grain (microgametophyte) represents the extremely reduced haploid generation, in the diplo-haplontic life cycle of flowering plants. The major fuction of a pollen grain is the delivery of two sperm nuclei to the female embryo sac. The process of pollen

tube penetration through the stylar canal requires the degradation of the cell wall components. Investigations from various laboratories, reveal that a large number of cloned pollen-specific genes code for cell wall degrading enzymes (Waterings,1992; Mu et al., 1994). Another biochemical aspect of a rapidly germinating pollen grain, is a large incrase in protein synthesis on the polyribosomes and a decrease in single ribosomes (Mascarenhass and Bell, 1969). Similar initiation of protein synthesis has also been reported in the pollen of several other plants (Linskens et al., 1970; Doekstra and Bruinsma, 1979).

Based on the timing of gene expression during pollen development (microsporogenesis/gametogenesis), various pollen-specific genes have been categorized as 'early pollen genes' or 'late pollen genes' (Mascarenhas, 1990). In Northern blot analysis (Fig.1) the expression of SF17 was not detected in closed disc florets (which consists of pollen containing green and yellow florets harvested before anthesis). The closed yellow florets contain well developed tri-cellular pollen grains. Based on this temporal expression pattern, the SF17 gene (similar to SF3 and SF16) was placed in the category of post-anthesis "late" pollen genes. Late pollen genes are implicated to play an important role in pollen maturation and/or pollen tube growth (Mascarenhas, 1990).

Rat p34 is a membrane protein characteristic of rough ER and is believed to play a vital role in the ribosome and ER membrane association (Ohsumi et al., 1993). SF17 protein shows 37% identity (57% similarity) with the rat ribosome-binding protein p34 (Fig. 9a). The homology was found in an array of leucine-rich-repeats which may provide a site of interaction with the ribosome or may associate with proteins to form translocation sites through which nascent polypeptides are transferred across the membrane bi-layer of the rough microsomes (Ohsumi et al., 1993). By analogy, it will be

tempting to hypothesize, that the LRRs of SF17 might be involved in the binding of ribosomes to the rough ER, leading to the accumulation of polysomes, and as a corollary to the rapid initiation of protein synthesis during pollen germination or pollen tube growth. Subcellular localization of the SF17 protein, and identification of interacting cellular component(s), will shed more light on its function.

REFERENCES

Albani, D., Altosaar, I., Arnison, P.G., and Fabijanski, S.F. (1991). A gene showing sequence similarity to pectin esterase is specifically expressed in developing pollen of *Brassica napus*. Sequences in its 5' flanking region are conserved in other pollen-specific promoters. Plant Mol. Biol. 16, 501-513.

Allen, R.L. and Lonsdale, D.M. (1993). Molecular characterization of one of the maize polygalacturonase gene family members, which are expressed during pollen development. Plant J. 3, 261-271.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Liopman, D.J. (1991). Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Arrand J.E. (1985). preparation of nucleic acid probes, In Nucleic acid hybridization: A practical approach, B.D. Hames., and S.J. Higgins, eds (Oxford: IRL Press) pp. 17-45.

Baltz, R., Domon, C., Pillay, D.T.N., and Steinmetz, A. (1992a). Characterization of a pollen-specific cDNA from sunflower encoding a zinc finger protein. Plant J. 2, 713-721.

Baltz, R., Evrard, J.L., Domon, C., and Steinmetz, A. (1992b). A LIM motif is present in a pollen-specific protein. Plant Cell 4, 1465-66.

Bedinger, P. (1992). The remarkable biology of pollen. Plant Cell 4, 879-887.

Bednarek. S., Wilkins, T., Dombrowski, J.E., and Raikhel, N. (1990). A carboxyl-terminal propeptide is necessary for proper sorting of barley lectin to vacuoles of tobacco. Plant Cell 2, 1145-1155.

Benton, W. D., and Davis R.W. (1977). Screening λgt recombinant clones by hybridization to single plaques in situ. Science 196, 180-182.

Bernier. G., Havelange, A., Houssa, C., Petitjean, A., and Lejeune, P. (1993). Physiological signals that induce flowering. Plant Cell 5, 1147-1155.

Birnboim, H.C., and Doly, J. (1979). A rapid extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7, 1513-1523.

Boller, T., and Wiemken, A. (1986). Dynamics of vacuole compartmentation. Annu. Rev. Plant Physiol. 37, 137-164.

Bonner, L.J., and Dickinson, H.G. (1989) Anther dehiscence in *Lycopersicon esculentum* Mill. I. Structural aspects. New Phytol. 113, 97-115.

Brown, S.M., and Crouch, M.L. (1990). Characterization of a gene family abundantly expressed in *Oenothera organensis* pollen that shows sequence similarity to polygalacturonase. Plant Cell 2, 263-274.

Buck, L., and Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. Cell 65, 175-187

Chang C, G.E., Schaller, S.E., Patterson, S.F., Kwok, E.M., Meyerowitz and Bleecker, A.B. (1992). The TMK1 gene from Arabidopsis codes for a protein with structural and biochemical characteristics of a receptor protein kinase. Plant Cell 4, 1263-1271.

Chasan. R., and Walbot, V. (1993). Mechanisms of plant reproduction: Questions and approaches. Plant Cell 5, 1139-1146.

Chrispeels, M.J. (1991). Sorting of proteins in the secretory system. Annu. Rev. Plant Physiol. Plant Mol. Biol. 42, 21-53.

Chrispeels, M.J., and Raikhel, N.V. (1992). Short peptide domains target proteins to plant vacuoles. Cell 68, 613-616.

Coen S.C., and Carpenter, R. (1993). The metamorphosis of flowers. Plant Cell 5, 1175-1181.

Coruzzi, G., Broglie, C., Edwards, C., and Chua, N.H. (1984). Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5- bisphosphate carboxylase. EMBO J 3: 1671-1679

Cox, K.H., and Goldberg, R.B.(1988). Analysis of plant gene expression. In Plant molecular biology, C. H. Shaw ed (Oxford: IRL press) pp. 1-36.

Cutler, M.L., Bassin, R.H., Zanoni, L., and Talbot, N. (1992). Isolation of *rsp-1*, a novel cDNA capable of suppressing v-RAS transformation. Mol. Cell. Biol. 12, 3750-3756.

Davis, R.W., Botstein, D., and Roth, J.R. (1980). Advanced bacterial genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Dellaporta, S.L ., Wood, J., and Hicks J.B. (1983). A plant DNA minipreparation: Version 2. Plant Mol. Biol. Rep. 1, 19-21.

Denecke, J., Botterman, J., and Deblaere, R. (1990). Protein secretion in plant cells can occur via a default pathway. Plant Cell 2, 51-59.

Dedio ,W., and Putt, E.D. (1980). Sunflower. In Hybridization of Crop Plants, W.R. Fehr., H.H. Hadley, eds (Madison, WI: American Society of Agronomy-Crop Science Society of America), pp. 631-644.

Domon, C., Evrard, J.L., Herdenberger, F., Pillay, D.T.N., and Steinmetz, A. (1990). Nucleotide sequence of two anther-specific cDNAs from sunflower. Plant Mol. Biol. 15, 643-646.

Domon, C., Evrard, J.L., Pillay, D.T.N.., and Steinmeitz, A. (1991). A 2.6 Kbp intron separates the signal peptide coding sequence of an anther-specific protein from the rest of the gene in sunflower. Mol. Gen. Genet . 229, 238-244.

Domon, C., and Steinmeitz, A. (1994). Exon shuffling in anther-specific genes from sunflower. Mol. Gen. Genet. 244, 312-317.

Dudareva, N., Evrard, J.L., Pillay, D.T.N., and Steinmetz, A. (1994a) Nucleotide sequence of a pollen-specific cDNA from *Helianthus annuus* L. encoding a highly basic protein. Plant Physiol. 106, 403-404

Dudareva, N., Evrard, J.L., Pillay, D.T.N., and Steinmetz, A. (1994b) Molecular characterization of a pollen-specific cDNA from *Helianthus annuus* L. and its expression. Mol. Biol. (Life Sci. Adv.) 13, 261-268

Dunn, I.S., and Blattner, F. (1987). Charons 36 to 40: multienzyme , high capacity, recombination deficient vectors with polylinkers and polystuffers. Nucleic Acids Res. 15, 2677-2698.

Esau, K. (1993). Anatomy of seed plants. (New York: John Wiley)

Everett, N..P., Robinson, K.E.P., and Masccarenhas, D. (1987). Genetic engineering of sunflower (*Helianthus annuus* L.) Biotechnology 5, 1201-1204.

Evrard, J.L., Jako, C., Saint-Guily, A., Weil, J.H., and Kuntz, M. (1991). Anther-specific developmentally regulated expression of genes coding a new class of proline-rich proteins in sunflower. Plant Mol. Biol. 16, 271-281.

Feinberg, K., and Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 132, 6-13.

Goldberg, R.B. (1988.) Plants: Novel developmental processes. Science, 240: 1460-1467.

Goldberg, R.B., Beals, T. P., and P.M. Sanders (1993). Anther development: Basic principles and practical applications. Plant Cell 5, 1217-1229.

Gubler ,U., and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. Gene 25, 263-269.

Hanahan, D. (1983). Studies on transformation of *Eschrichia coli* with plasmids. J. Mol. Biol. 166, 557-580.

Handa, M., Titani, K., Holland, L.J., Roberts, J.R., and Ruggeri, Z.M. (1986). The von Willebrand factor-binding domain of platelet membrane glycoprotein Ib. J. Biol. Chem. 261, 12579-12585.

Hashimoto, C., Hudson, K.L., and Anderson, K.V. (1988). The *Toll* gene of Drosophila, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52, 269-279.

Herdenberger, F., Evrard, J.L., Kuntz, M., Tessier, L.H., Klein, A.L., Steinmetz, A., and Pillay, D.T.N. (1990). Isolation of flower specific cDNA clones from sunflower (*Helianthus annuus* L.) Plant Sci. 69, 111-122.

Horner, H.T. (1977). A comparative light- and electron-microscopic study of microsporogenesis in male-fertile and cytoplasmic male-sterile sunflower (*Helianthus annuus*). Am. J. Bot. 64, 745-759.

Huala, E ., and Sussex., I.M. (1993). Determination and cell interactions in reproductive meristems. Plant Cell 5, 1157-1165.

Joshi, C.P. (1987a) An inspection of domain between TATA box and transcription start site of 79 plant genes. Nucleic Acids Res. 15, 6643-6653.

Joshi, C.P. (1987b). Putative polyadenylation signals in nuclear genes of higher plants: a compilation and analysis. Nucleic Acids Res. 15, 9627-9640

Kamalay, J.C., and Goldberg, R.B. (1980). Regulation of structural gene expression in tobacco. Cell 19, 935-946.

Kamalay, J.C., and Goldberg, R.B. (1984). Organ-specific nuclear RNAs in tobacco. Proc. Natl. Acad. Sci. USA 81, 2801-2805.

Kataoka, T., Broek, D., and Weigler, M. (1985). DNA sequence and characterization of the *S. cerevisiae* gene encoding adenylate cyclase. Cell 43, 493-505.

Knowles, P.F. (1978). Morphology and anatomy. In Sunflower science and technology, Agronomy Ser. No. 19, J.F. Carter, ed (Madison, WI: American Society of Agronomy), pp.55-87.

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

Kuhlemeier, C., Green, P.J., and Chua, N.H. (1987). Regulation of gene expression in higher plants. Annu. Rev. Plant Phyiol. 38, 221-257.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Lam, E., and Chua, N.H. (1989). ASF-2: A Factor that binds to the cauliflower mosaic virus 35S promoter and conserved GATA motif in Cab promoters. Plant Cell 1, 1147-1156.

Lee, F.S., and Valle, B.L. (1990). Modular mutagenesis of the human placental ribonuclease inhibitor, a protein with leucine-rich repeats. Proc. Natl. Acad. Sci. USA 87, 1879-1883.

Locker, J .(1979). Analytical and preparative electrophoresis of RNA in agarose-urea. Anal. Biochem. 98, 358-367.

Lopez, J.A., Chung, D.W., Fujikawa, K., Hagen, F.S., and Roth, G.J. (1988). The α and β chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence. Proc. Natl. Acad. Sci. USA 85, 2135-2139.

Mariani, C., De Beuckeleer M., Truttner, J., Leemans J., and Goldberg, R.B. (1990). Induction of male sterility in plants by a chimaeric ribonuclease gene. Nature 347, 737-741.

Mariani, C., Gossele, V., De Beuckeleer, M., De Block, M., Goldberg, R.B., De Greef, W., and Leemans, J. (1992). A chimaeric ribonucleasee-inhibitor gene restores fertility to male sterile plants. Nature 357, 384-387.

Mascarenhas, J.P. (1989). The male gametophyte of flowering plants. Plant Cell 1, 657-664.

Mascarenhas, J.P. (1990). Gene activity during pollen development. Annu. Rev. Plant Physiol. Plant Mol. Biol. 41, 317-338.

Mascarenhas, J.P. (1993). Molecular mechanisms of pollen tube growth and differentiation. Plant Cell 5, 1303-1314.

Mascarenhas, J.P., and Bell, E. (1969). Protein synthesis during germination of pollen: Studies on polysome formation. Biochim. Biophys. Acta. 179, 199-203.

McCormick, S. (1991). Molecular analysis of male gametogenesis in plants. Trends Genet. 7, 298-303.

McCormick, S. (1993). Male gametophyte development. plant Cell 5, 1268-75.

Meyerowitz, E.M. (1987). In situ hybridization to RNA in Plant tissue. Plant Mol. Biol. Rep. 5, 242-250.

Mindrinos, M. , Katagiri, F., Yu, G-L., and Ausubel, F.M. (1994). The A. thaliana disease resistance gene RPS2 encodes a protein containing a nucleotide-binding site and leucine- rich repeats. Cell 78, 1089-1099.

Mu, J.H., Stains, J.P., Kao, T.H. (1994). Characterization of a pollen-expressed gene encoding a putative pectin esterase of *Petunia inflata*. Plant Mol. Biol. 25, 539-544.

Nacken, W.K.F., Huijser, P., Beltran, J.P., Saedler, H., and Sommer. H. (1991). Molecular characterization of two stamen-specific genes tap1 and fil1 , that are expressed in wild type, but not the deficiens mutant of *Antirrhinum majus* . Mol. Gen. Genet. 229, 129-136.

Nakashima, H., and Hosokawa, S. (1974). Studies on histological features of male-sterility in sunflower (*Helianthus annuus* L.). Proc. Crop Sci. Soc. Japan 43, 475-481.

Nanda, K., and Gupta, S.C. (1975). Syngenesious anthers of *Helianthus annuus* - a histochemical study. Bot. Not. 128, 450-454

Niogret, M.F., Dubaid, M., Mandaron, P., and Mache, R. (1991). Characterization of pollen galacturonase encoded by several cDNA clones in maize. Plant Mol. Biol. 17, 1155-1164.

Nose A, Mahajan, V.B., and Goodman, C.S. (1992) Connectin: A homophilic cell adhesion molecule expressed on a subset of muscle and the motoneurons that innervate them in Drosophila. Cell 70, 553-567.

Ohsumi, T., Ichimura, T., Sugano, H., Omata, S., Isobe, T., and Kuwano, R. (1993). Ribosome-binding protein p34 is a member of the leucine-rich-repeat-protein superfamily. Biochem. J. 294, 465-472.

Pressy, R., and Reger, B.J. (1989). Polygalacturonase in pollen from corn and other grasses. Plant Sci. 59, 57-62.

Raven, P.H., Evert , R.F., and S.E. Eichhorn, S.E. (1991). Biology of Plants. (New York: Worth Publishers).

Reinke, R., Krantz, D.E., Yen, D., and Zipursky, S.L. (1988). Chaoptin, a cell surface glycoprotein required for Drosophila photoreceptor cell morphogenesis, contains a repeat motif found in yeast and human. Cell 52, 291-301.

Revelard, P., Lips, S., and Pays, E. (1990). A gene from the VSG expression site of *Trypanosoma brucei* encodes a protein with both leucine-rich repeats and a zinc finger. Nucleic Acids Res. 18, 7299-8303.

Rogers, H.J., Harvey, A., Lonsdale, D.M. (1992). isolation and characterization of a tobacco gene with homology to pectate lyase which is specifically expressed during microsporogenesis. Plant Mol. Biol. 20, 493-502.

Rothman, J.E. (1987). Protein sorting by selective retention in the endoplasmic reticulum and Golgi stack. Cell 50, 521-522.

Saiger, D., and Apel, A. (1993). Molecular characterization of two cDNAs from *Sinapis alba* L. expressed specifically at an early stage of tapetum development.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd edition. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Sanders, S.L, and Schekman, R. (1992). Polypeptide translocation across the endoplasmic reticulum membrane. J. Biol. Chem. 267, 13791-13794.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Schneider, R., and Schweiger, M. (1991). A novel modular mosaic of cell adhesion motifs in the extracellular domains of the neurogenic trk and trk B tyrosine kinase receptors. Oncogene 6, 1807-1811.

Scott, R., Dagless, E., Hodge, R., Paul, W., Soufleri, I., and Draper, J. (1991a). Patterns of gene expression in developing anthers of *Brassica napus*. Plant Mol. Biol. 17, 195-207.

Scott, R., Hodge, R., Paul, W., and Draper, J. (1991b).The molecular biology of anther differentiation. Plant Sci. 80, 167-191.

Scott, R. (1993). Anther Development: A Molecular Perspective. In Molecular Biology of Flowering, B.R. Jordan ed (Oxon, UK: CAB), pp. 141-164.

Shen, J.B. and Hsu, F.C. (1992). Brassica anther-specific genes: Characterization and in situ localization of expression. Mol. Gen. Genet. 234, 379-389.

Smith, A.G., Gasser, C.S., Budelier, K., andFraley R.T.(1990). Identification and characterization of stamen-and tapetum-specific genes from tomato. Mol. Gen. Genet. 222, 9-16.

Staiger, D., Apel, K. (1993) Molecular characterization of two cDNAs from *Sinapis alba* L. expressed specifically at an early stage of tapetum development. Plant J. 4, 697-703

Steeves, T.A and Sussex. I.M. (1991). Patterns in plant development. 2nd edition. (N.Y: Cambridge Univ. Press).

Steinmetz, A., Baltz, R., Domon, C., Dudareva, N., Evrard, J.L., and Krauter, R.(1992). Gametophytic and sporophytic gene expression in *Helianthus annuus* L. In Reproductive biology and plant breeding, Y. Datte, C. Dumas, and A. Gallais, eds (Berlin: Springer-verlag), pp. 59-68.

Steinmetz, A., Baltz, R., Domon, C., Dudareva, N., Evrard, J.L., Krauter, R., Mundel, C., and V.Tabidze. (1993). Flower-specific genes in sunflower. Biotechnol. & Biotechnol. Eq. 4, 16-20.

Stotz, H.U., Contos, J.J.A., Powell, A.L.T., Bennet, A.B., and Labavitch, J.M. (1994). Structure and expression of an inhibitor of fungal polygalacturonase form tomato. Plant Mol. Biol. 25, 607-617.

Takahasi, N., Takahasi, Y., and Putnam, F.W. (1985). Periodicity of leucine tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich α2-glycoprotein of human serum. Proc. Natl. Acad. Sci. USA 82, 1906-1910.

Tebbutt, S.J., Rogers, H.J., Lonsdale, D.M.. (1994). Characterization of a tobacco gene coding a pollen specific polygalacturonase. Plant. Mol. Biol. 25, 283-297.

Titani, K,, Takio, K., Handa, M., and Ruggeri, Z.M. (1987). Amino acid sequence of the von Willebrand factor-binding domain of platelet membrane glycoprotein lb. Proc. Natl. Acad. Sci. USA 84, 5610-5614.

Turcich, M.P., Hamilton, D.A., Mascarenhas, J.P. (1993). Isolation and characterization of pollen-specific maize genes with sequence homology to ragweed allergens and pectate lyases. Plant Mol. Biol. 23, 1061-1065.

Tsuchiya, T., Toriyama, K., Nasrallah, M.E., Ejiri, S. (1992). Isolation of genes abundantly expressed in rice anthers at the microspore stage. Plant Mol. Biol. 20, 1189-1193.

Twell, D., Yamaguchi, J., Wing, R.A., Ushiba, J., and McCormick, S. (1991). Promoter analysis of genes that are coordinately expressed during pollen development reveals pollen-specific enhancer sequences and shared regulatory elements. Genes and Dev. 5, 496-507.

Valon, C., Smalle ,J., Goodman, H.M., and Giraudat, J. (1993). Characterization of an *Arabidopsis thaliana* gene (TMKL1) encoding putaative transmembrane protein with an unusual kinase-like domain. Plant Mol. Biol. 23, 415-421.

Van Tunen, A.J., Koes, R.E., Spelt, C.E., Van der Krol, R., Stuitje, A.R., and Mol, J.N.M. (1988). Cloning of two chalcone isomerase genes from *Petunia hybrida* : coordinated, light regulated and differential expression of flavonoid genes. EMBO Journal 7, 1257-1263.

Von Heijne, G. (1986). A new method for predicting signal cleavage sites. Nucleic Acids Res. 14, 4683-4690.

Von Heijne G (1990) The signal peptide. J Membrane Biol. 115, 195-201

Walker, J.C. (1993). Receptor-like protein kinase genes of *Arabidopsis thaliana*. Plant J. 3, 451-456.

Weigel, D. and Meyerowitz, E.M. (1994). The ABCs of floral homeotic genes. Cell 78, 203-209.

Whitham, S., Dinesh-Kumar, S.p., Choi, D., Hehl, R., Corr, C., and Baker, B. (1994). The product of tobacco mosaic virus resistance gene N: Similarity to Toll and interleukin-1 receptor. Cell 78, 1101-1115.

Wicki, A.N., and Clemetson, K.J. (1985). structure and function of platelet membrane glycoproteins Ib and V. Eur. J. Biochem. 153, 1-11.

Wing R.A., Yamaguchi., J., Larabell, S.K., Ursin, V.M., McCormick, S. (1989). Molecular and genetic characterization of two pollen-expressed genes that have sequence similarity to pectate lyases of the plant pathogen Erwinia. Plant Mol. Biol. 14, 17-28.

Wodehouse, R.P. (1935). Pollen Grains. (N.Y: McGraaw-Hill Book Co.).

Yamawaki-Kataoka, Y., Tamaoki, T. , Choe, T.R., and Tanaka, H. (1989). Adenylate cyclases in yeast: A comparison of the genes from *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae.*

Young, D., Riggs, M., Field, J., Vojtek, A., Broek, D., and Wigler, M. (1989). The adenylyl cyclase gene from *Schizosaccharomyces pombe*. Proc. Natl. Acad. Sci. USA 86, 7989-7993.

Zhang, H., Scholl, R., Browse, R., and Somerville, C. (1988). Double-stranded DNA sequencing as a choice for DNA sequencing. Nucleic Acids Res. 16, 1220.

CURRICULUM VITAE

J. Thirupathi Reddy

EDUCATION:

B.Sc.Ag	1980	Andhra Pradesh Agricultural University Hyderabad, India
M.Sc.Ag	1982	Andhra Pradesh Agricultural University Hyderabad, India
M.Sc.	1990	University of Western Ontario London, Ontario, Canada
Ph.D.	1995	University of Windsor, Windsor, Ontario, Canada

AWARDS:

Andhra Pradesh Agricultural University Graduate Scholarship	1980 (Declined)
Indian Council of Agricultural Research Fellowship	1982
University of Western Ontario Scholarship	1987
University of Western Ontario Visa Bursary	1989
University of Western Ontario Research and Teaching Assistantship	1990
University of Western Ontario Special University Scholarship	1989
University of Windsor Visa Bursary	1993
University of Windsor Research and Teaching Assistantship	1995

EMPLOYMENT:

Research Assistant (Rice Breeding): 1983-85 A.P Agricultural Univ., India
Worked in the Development of Disease Resistant Rice cultivars
Production of Rice Breeder's Seed & Certified Seed

Instructor & Lecturer (Genetics & Plt. Breeding) 1985-87 A.P Agril. Univ., India.
Taught: Principles of Plant Breeding to Senior Undergraduate students
Taught: Seed Production and Certification to Senior Undergraduates

Teaching & Research Assistant 1987-90 U.W.O., London, Canada

Teaching & Research Assistant 1990-95 Univ. of Windsor, Canada

INTENSIVE LAB & LECTURE COURSES ATTENDED:

cDNA Library Construction	Jan, 1992	LTI,GIBCO,BRL Washington D.C
Site Directed Mutagenesis	May, 1992	Catholic Univ. Washington D.C
Polymearse Chain Reaction	May, 1992	Catholic Univ. Washington D.C
DNA Sequencing	May, 1992	Catholic Univ. Washington D.C

OTHER RELEVANT TRAINING:

Institut de Biologie Moleculaire des Plantes,
C.N.R.S. Strasbourg, France May-July 1991

MEMBERSHIP IN SOCIETIES:

American Society of Plant Physiologists
Canadian Society of Plant Molecular Biology
International Society of Plant Molecular Biology
American Association for the Advancement of Science

CONFERENCES:

Third International Congress of Plant Molecular Biology
Tuscon, Arizona, October, 1991
Annual meeting of the Canadian Society of Plant Molecular
Biology, Toronto, August, 1993

PUBLICATIONS:

Reddy, J.T. and M. Locke (1990). The size limited penetration of gold particles through insect basal laminae. J. Insect Physiol. 36:397-407