Molecular analysis of NOZZLE, a gene involved in pattern formation and early sporogenesis during sex organ development in Arabidopsis thaliana

(anther/organogenesis/ovule)

URSULA SCHIEFTHALER†, SURESHKUMAR BALASUBRAMANIAN†, PATRICK SIEBER†, DAVID CHEVALIER†, ELLEN WISMAN‡, AND KAY SCHNEITZ†§

†Institute of Plant Biology, University of Zurich, Zollikerstr. 107, CH-8008 Zurich, Switzerland; and ‡Max-Plank-Institut für Züchtungsforschung, Carl-von-Linné-Weg 10, D-50829 Cologne, Germany

Communicated by Elliot M. Meyerowitz, California Institute of Technology, Pasadena, CA, July 26, 1999 (received for review June 3, 1999)

ABSTRACT Sexual reproduction is a salient aspect of plants, and elaborate structures, such as the flowers of angiosperms, have evolved to aid in this process. Within the flower the corresponding sex organs, the anther and the ovule, form the male and female sporangia, the pollen sac and the nucellus, respectively. However, despite their central role for sexual reproduction little is known about the mechanisms that control the establishment of these important structures. Here we present the identification and molecular characterization of the NOZZLE (NZZ) gene in the flowering plant Arabidopsis thaliana. In several nzz mutants the nucellus and the pollen sac fail to form. It indicates that NZZ plays an early and central role in the development of both types of sporangia and that the mechanisms controlling these processes share a crucial factor. In addition, NZZ may have an early function during male and female sporogenesis as well. The evolutionary aspects of these findings are discussed. NZZ encodes a putative protein of unknown function. However, based on sequence analysis we speculate that NZZ is a nuclear protein and possibly a transcription factor.

The elucidation of the molecular mechanisms underlying organogenesis continues to represent a formidable challenge for biologists. In this context the plant reproductive organs recently have attracted considerable attention. The life cycle of plants alternates between the diploid sporophytic and the haploid gametophytic phase. The sporophyte produces the unisexual spores that produce the gametophyte, which in turn encodes a homeodomain putative transcription factor. For example, how are the three elements, the nucellus, chalaza, and funiculus, set up? There is a lot of indirect evidence that supports a corresponding early subdivision of the ovule primordium (4, 5); however, only one candidate is known that may be directly involved in establishing one of these elements. Mutations in the BELL (BEL1) gene lead to an altered chalazal development and outgrowths of unknown identity develop in place of integuments (6–8). The BEL1 locus encodes a homeodomain putative transcription factor and its RNA expression pattern becomes restricted to the central region before the initiation of the integuments (9).

In this context a number of important questions surface. For example, what genes regulate the formation of the nucellus and are they related to BEL1? Does the formation of the nucellus and pollen sac, two quite different sporangia, involve common factors? As a first step toward the elucidation of these questions we describe in this paper the identification and molecular characterization of the NOZZLE (NZZ) gene in A. thaliana. The NZZ locus encodes a novel protein. The nzz mutant phenotype indicates that NZZ has a role in the establishment of the pollen sac and nucellus and possibly an early role in sporogenesis.

Abbreviations: MCC, megaspore mother cell; PMC, pollen mother cell; RT-PCR, reverse transcription–PCR.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF146794).

†To whom reprint requests should be addressed. E-mail: Kay.Schneitz@access.unizh.ch.
METHODS

Plant Work and Genetics. Plants were grown as described (8), and A. thaliana (L.) Heynh. var. Landsberg (erecta mutant) was used as the wild-type strain. The nzz-1 and nzz-3 mutants were found in an En-1/spm transposon-induced mutant population (Columbia ecotype) (10) whereas the nzz-2 mutant was isolated in a screen by using ethyl methane sulfonate as mutagen (Ler ecotype). The mutations were shown to be allelic (not shown). The nzz-1 and the nzz-2 alleles were used in two independent mapping experiments involving outcrosses to Ler and Columbia plants, respectively. The combined results yielded a map position between markers PG11 (2/130 recombinant chromosomes) and PHA (7/132) on chromosome 4.

Molecular Identification of nzz Mutations. Genomic DNA flanking the left side of the En-1 insertions in nzz-1 (after outcrossing to reduce the En-1 copy number) was isolated by PCR using nested En-1-specific and splinkerette-based primers (11) and subcloned into the vector pCR 2.1 (Invitrogen). The En-1-specific primers were En-225 (CAGCACGCTCTG- TAGAATAGGATTGGGGAATTT) and En-93 (GGTGCAAGCAACACACTTTAATCTGGATTAAAG). Genomic clones from the right side of the En-1 insertion were obtained by using a second set of primers based on the sequence of clone ATM422.140. They were; for nzz-1, En-1-specific CGGAA-TTCGATGCTGGCTGTCTTGAAG and NZZ-specific CCAAAAGGCTTATCACAAGGC; for nzz-3, En-1-specific TCAGCGTACATCAGTCTGAC (Th-763) and NZZ-specific GAATCGACAGGATCATGG. Genomic clones from the left side of nzz-3 were isolated by using a third set of primers: En-1-specific, AGCACGCGCTGTAGAATAGGATTG (En-203); NZZ-specific, CTCGAGTTCTACAACAACGCT. The footprint allele nzz-1* was confirmed by subcloning three independent copies of PCR fragment subclones by using nzz-1*-specific genomic DNA and two primers flanking the En-1 insertion site in the nzz-1 mutant (AGGACTGCAATCACCTACGAG, GTGACACTTGCTTGAGCCC). Genomic DNA from the right side of the nzz-2 insertion was amplified by PCR and subcloned, and three different clones were sequenced. An identical approach was used to generate the Ler wild-type sequence. The ATM422 sequence (AL030978) and the Ler genomic sequence of NZZ (AI467974) can be found in the GenBank database.

DNA Work and cDNA Isolation. Regular protocols were used for routine DNA work (12). Several cDNA clones were obtained through a 5’ and 3’ rapid amplification of cDNA ends (RACE) reverse transcription–PCR (RT-PCR) (13) by using the Marathon kit (CLON-TECH) and poly(A)+ RNA from flowers of stages 1–12 (14) (Columbia ecotype). The following nested NZZ-specific primers were used: CTGATCGTGGCCCTTCATCGGAG and GTGTATTTTGGAAATCCACACCGTTGCG (5’ RACE) as well as GATCTCTACACAACAGCTACCGCTTT- ACCC and GTGTTGTGCTACAAGGCTTCCACAA (3’ RACE). The longest PCR fragments were sequenced on both strands. Sequences were obtained by standard cycle sequencing using an ABI 373 sequencer (Applied Biosystems).

RNA Work and Reverse Transcription–PCR (RT-PCR). Total RNA was isolated from inflorescences (12) carrying flowers from stage 1 to 12. The poly(A)+ RNA was isolated by using oligo(dT)25-Dynabeads according to the manufacturer’s protocol (Dynal). For Northern analysis 10 μg of poly(A)+ RNA was loaded on a formaldehyde-agarose gel. After electrophoresis the poly(A)+ RNA was blotted onto a Hybond-N nylon membrane (Amersham Pharmacia) and hybridized according to the manufacturer’s recommendations. The RNA size marker G319 was included for size calibration (Promega). A genomic DNA PCR fragment spanning the 3’ two-thirds of the gene (nucleotides 58168–59191, see Fig. 2) was labeled with [α-32P]dCTP (3,000 Ci/mmol) and used as probe. Exposure was for 4 days at −80°C by using a Kodak Biomax MS film with the help of an intensifying screen. For the RT-PCR experiments 2 μg of poly(A)+ RNA, isolated from the corresponding tissue, was reverse-transcribed with Moloney murine leukemia virus reverse transcriptase (NEB, Beverly, MA). The following NZZ-specific primers, flanking the first intron, were used for PCR: AGTACTAGAAGCAGAATCTGAAGCTG and CATCGGATCTTACAAACGCT. The PCR products were loaded on an agarose gel and a Southern blot analysis was performed. The same probe as in the Northern experiment was used. As a control a RT-PCR using primers corresponding to the Arabidopsis GAC gene (15) (CCTGTGTGCGCAAGTGC, CACTTGAAGG- GTGGTGCCAAAG) was performed.

RESULTS

The nzz Mutant Phenotype. A summary of the visible defects that occur in nzz mutants is given in Fig. 1. A detailed genetic analysis of NZZ will be presented elsewhere (S.B. and K.S., unpublished work). The wild-type development of ovules and anthers in Arabidopsis has been described by several authors (5–7, 18–20). In nzz mutants alterations apparently are restricted to the anthers and ovules. Very early defects in the development of both organs can be observed. In the ovule the distal tip tissue (normally the nucellus) usually is very much reduced, sometimes even appears absent, and a MMC usually is not detected. As a result no nucellus and no embryo sac can be seen at later stages. Furthermore, the inner integument initiation is delayed and the integuments often are reduced, in some instances they are missing (not shown), and the funiculus is longer because of a higher cell number (not shown). In the young anther the initial periclinal division of the archespore cells that leads to the primary parietal layer and the PMCs cannot be observed. Only a mass of subepidermal uniform cells is apparent throughout development and at late stages no pollen sacs are seen. As a result of these defects nzz mutants are male and female sterile.

Molecular Identification of the NZZ Gene. We have isolated three independent mutant alleles of NZZ, two of which were obtained by using an En-1 transposon-tagging approach (see Methods). The frequent presence of fertile revertant sectors in nzz-1 mutants (not shown) indicated that the NZZ locus was tagged. Genomic DNA flanking the left side of the En-1 insertion sites in nzz-1 mutants was isolated and sequenced. Sequence database searches showed one of the sequences obtained to be identical to a sequence corresponding to clone ATM422.140. This clone stems from the Arabidopsis genome project, corresponds to a gene-coding region, and maps to the same interval on chromosome 4 as NZZ. The sequence information from this clone was used to design PCR primers that were successfully used to clone the right flanking region of nzz-1, to identify and clone the genomic DNA of the En-1 insertion site of nzz-3, and to identify a point mutation in this
region in nzz-2 (Fig. 2 A and B) (see Methods). During the process of identifying the En-1 insertion that causes the nzz mutant phenotype in nzz-1 mutants on a Southern blot, using genomic DNA derived from members of a segregating popu-

lation, a mutant plant carrying a putative footprint mutation (nzz-1*) was identified (not shown). The footprint mutation was expected to have arisen because of the imprecise excision of the En-1 transposon. This possibility was verified by sequence analysis of the nzz-1* DNA spanning the corresponding En-1 insertion site. The result showed that a 2-bp insertion is present in the nzz-1* mutant (Fig. 2B).

Molecular Characterization of NZZ. The longest cDNA spans a region of 1.285 kb (Fig. 2A), a value in good agreement with the 1.4-kb transcript length as shown by Northern analysis (see Fig. 4A). There is an at least 88-bp 5' and a 252-bp 3' untranslated region and there are two small introns of 99 bp and 89 bp, respectively. The longest ORF, after conceptual translation, encodes a 314-aa protein with a calculated molecular mass of 31.1 kDa. The putative NZZ protein sequence does not show any significant homology to entries in the European Molecular Biology Laboratory/GenBank sequence databases. However, close examination of the protein sequence by computer analysis revealed several features (Figs. 2 and 3).

Overall the protein appears to be hydrophilic (Fig. 3). In particular there is an about 40-residue domain, which we call PC domain for polar-charged, spanning residues 41–80 with 22.5% polar residues and 45% charged residues (35% basic and 10% acidic aa). A hydrophobicity plot (21) revealed three somewhat less hydrophilic regions. There is a larger, so-called middle region, which includes residues 81–146 and contains 39.4% hydrophobic residues. In addition, there may exist two smaller regions, separated by 40 aa, which span residues 232–247 (D1) and 288–302 (D2). Scanning the PROSITE database (22) for protein domain signatures revealed imperfect matches of the D1 and D2 sequences with the signature for a myc-type helix-loop-helix (HLH) protein dimerization domain (but see Discussion). In addition, one can find two candidate nuclear localization signals (23) (Fig. 2A). A number of potential amidation, glycosylation, myristylation, and phosphorylation signals are present in the putative NZZ protein sequence (not shown), which could indicate that the NZZ protein undergoes posttranslational modifications.

All three nzz mutations map within the transcribed region of NZZ (Fig. 2). In the case of nzz-1 the En-1 insertion is just after nucleotide 58638. In this mutant a partially truncated protein of 254 residues still could be formed (with the last 10 residues being encoded by the left end of the En-1 element; not shown). Such a protein would carry an aberrant D1 part and lack the remainder, including the D2 region, of the wild-type NZZ protein. In nzz-2 mutants a G to A transition at nucleotide 59040 results in a stop codon. Thus, a protein of 142 residues could be produced that lacks a putative nuclear localization signal and the D1 and D2 region. The nzz-3 mutation is caused by an En-1 insertion just after nucleotide 58315 in the 3' untranslated region. Thus, it raises the possibility for a role of this region in for example mRNA stability or translational control of NZZ (24).

The Temporal and Spatial Expression Pattern of NZZ Transcripts. As expected from the spatial extent of the nzz mutant phenotype (Fig. 1) NZZ transcripts are found in flowers (Fig. 4A). Interestingly, transcripts also are detected in rosette leaves, stem tissue, and seedlings (Fig. 4B). To investigate the spatial and temporal localization of NZZ transcripts during floral development a series of in situ hybridization experiments were performed (Fig. 5). Here we will focus on the expression of NZZ during early anther and ovule development. However, NZZ transcripts can be detected in petals, carpels, and nectaries, and weak staining also is seen in the vascular tissue of the inflorescence (not shown).

With respect to ovule development NZZ expression is seen during all of stage I and throughout the ovule primordium (Fig. 5 A and B). At stage 2-1 the MMC also is staining. During stages 2-II/III, when the two integuments initiate, NZZ is

Fig. 1. Summary of the nzz mutant phenotypes. (A and B) Scanning electron micrographs. (C–H) Optical sections through whole-mount tissue. (A, C, E, and G) Wild type. (B, F, and H) nzz-2. (D) nzz-1. Stages: (A–D) ovule stage 2-III, (E and F) ovule stage 3-VI, (G and H) floral stage 7. (A and C) The nucellus is prominent and the two integuments are initiated. (B and D) Note the reduced distal tip region (arrow) and compare with A and C. (E) A fully developed embryo sac and fully developed integuments are seen. (F) There is a prominent absence of a nucellus and an embryo sac at this stage (arrow). The integuments are reduced; however, an endothelium still can be recognized. (G) An anther at a premeiotic stage. The PMC nuclei and the descendants of the primary parietal layer are clearly recognizable. (H) Only a mass of uniform cells is detectable in this anther (arrow), cc, central cell; e, epidermis; ec, egg cell; en, endothecium; et, endothecium, fu, funiculus; ii, inner integument; m, middle layer; mmc, megaspore mother cell; nu, nucellus; oi, outer integument; pmc, pollen mother cell; syn, synergid; tp, tapetum; vs, vascular strand. (Scale bars, 20 μm.)
somewhat stronger expressed in both developing integuments (Fig. 5C) whereas weak expression throughout the ovule is still apparent. During stage 3 prominent staining is visible in the proximal part of the endothelium (not shown).

The NZZ expression is visible throughout the anther primordium at early floral stage 7 and before the initial archesporial periclinal cell divisions (Fig. 5D). During further anther development NZZ continues to be broadly expressed with somewhat less stain visible in the epidermis (Fig. 5E). By the time of early meiosis NZZ transcripts apparently are restricted to the PMCs and the tapetum (Fig. 5F). The staining continues to be present in these two cell types until the pollen grains are released from the tetrads. After this stage NZZ expression can be detected only in individual pollen grains but quickly vanishes there as well (not shown).

The transcript patterns in other floral organs such as petals, carpels, and nectaries, as well as the expression found in rosette leaves, stem tissue, and seedlings, was unexpected as we could not detect a visible phenotype in these tissues. This finding can be explained in several ways. There is the possibility that NZZ may have no function in these parts of the plant or that the defects cannot easily be detected. One also could imagine that in these tissues the absence of NZZ activity in nzz mutants is compensated for by additional redundant factors.

**DISCUSSION**

Here we report our results on the molecular characterization of NOZZLE. The cloning of the NZZ locus was achieved by a transposon-tagging approach. We have molecularly characterized three independent nzz alleles. All the mutations map within the same gene, within exons, and two are located in the protein-coding part of the gene. Taken together, these data show that we have cloned and identified the NZZ locus.

**Fig. 2.** Sequence analysis of the NZZ gene. (A) The genomic DNA sequence of the NZZ gene [between 113 bp upstream of the longest cDNA (NZZc7) and 94 bp downstream of the poly(A) addition site] and the corresponding amino acid sequence of the putative NZZ protein are shown. Uppercase letters represent exon sequences. The 5’ end of NZZ points to the southern telomer. The numbering of nucleotides corresponds to the numbering given in the sequence of clone ATM4I22. The numbering of the amino acids is indicated above the sequence. The former alteration leads to an amino acid change. The positions of the three nzz mutations and the 5’ and 3’ ends of NZZ-cDNAs also are indicated. At the amino acid level the charged and hydrophobic domains as well as the D1 and D2 domains are flanked by brackets and braces, respectively. Double lines above the sequence highlight the two putative nuclear localization motifs. (B) Identification of NZZ. Genomic sequences found in the nzz-1, nzz-1*, nzz-2, and nzz-3 mutants and flanking the individual mutations are given as are the corresponding wild-type sequences. nzz-1* represents a mutant revertant of nzz-1. Bold letters refer to the target site duplications generated by the En-1 transposons. In the case of the nzz-1 and nzz-3 mutations the orientation of the En-1 elements is the same (left end to the left). Note the presence of a 2-bp footprint in nzz-1*. The position of the G to A transition in nzz-2 is underlined.
The Role of \textit{NZZ} During Anther and Ovule Development.

The mutant phenotype and the expression pattern indicate that \textit{NZZ} plays a wider role during ovule development as compared with anther ontogenesis. Here we focus on the requirements of \textit{NZZ} for sporangium formation. The observed expression of \textit{NZZ} in the chalaza, the integuments, and the funiculus is compatible with a function of \textit{NZZ} in aspects of ovule ontogenesis other than nucellus development. They will be discussed in great detail in a separate paper (S.B. and K.S., unpublished work).

It seems reasonable to infer from these data that \textit{NZZ} has equivalent functions during the early formation of the two types of sporangia, the nucellus and the pollen sac (see also below). What role could that be? The data are compatible with essentially two possibilities. One primary function of \textit{NZZ} could reside in the establishment of the distal region of the ovule primordium and the equivalent regions in the young anther. In \textit{nzz} mutants the reduction of the nucellus and pollen sacs are compatible with this explanation. Thus, the failure to form a MMC in ovules and the PMCs and parietal cell layers in anthers would be a secondary effect. The early broad transcript expression of \textit{NZZ} in both types of primordia supports such a view as well. Alternatively, \textit{NZZ} could have a primary role in establishing the sporogeneous lineage. In the absence of a functional MMC and the male archesporia the nucellus and pollen sacs would fail to proceed in their proper development. If true this would indicate that cells of the sporogeneous lineage somehow influence the development of the sporangium through cell-cell communication. The expression of \textit{NZZ} in the MMC and the PMCs corroborates this possibility. To us the former view appears as the most parsimonious explanation of our findings and thus we favor the view of \textit{NZZ} being an important factor in establishing the distal, i.e., the prospective spore-producing region of the ovule and anther primordia. However, at present we cannot rule out that \textit{NZZ} could reside in the establishment of the distal region of the ovule primordium and the equivalent regions in the young anther.

The Biochemical Function of the Putative \textit{NZZ} Protein.

Because of the novel structure of \textit{NZZ} the biochemical function of the \textit{NZZ} protein at present remains obscure. The significance of the match of the D1 and D2 regions to the PROSITE myc-type helix–loop–helix (HLH) protein dimerization signature is unclear as well because no HLH protein sequences were pulled out in corresponding database searches. The myc-related HLH proteins represent a group of transcription factors that act as dimers (25). Nevertheless, a few hints prompt us to speculate that \textit{NZZ} could be a transcription factor. There are two putative nucleolar localization signals and the basic polar-charged region might bind DNA. The middle region stretch and the D1/D2 region could be involved in protein–protein interactions. The \textit{nzz}-1 and \textit{nzz}-2 (a strong allele) mutations indicate that the region is important for \textit{NZZ} function. Obviously more data are required to address these questions.

\textbf{The Role of \textit{NZZ} in the Establishment of the Distal Region of the Ovule.}

In this section the posterior (abaxial) side of the inner integument is not visible. Note the stronger epidermal signal that can be detected in the developing integuments. (\textit{D}) An anther primordium at early floral stage 7 before the archesporial periclinal divisions. The \textit{NZZ} RNA is expressed throughout the primordium. (\textit{E}) An anther slightly older than in D. Note the reduced staining in the epidermis (arrow). (\textit{F}) An anther at an early meiotic stage. The \textit{NZZ} expression is restricted to the PMCs and the surrounding tapetum. Control experiments using a sense probe gave no signals (not shown). ii, inner integument; fu, funiculus; mmc, megaspore mother cell; oi, outer integument; pmc, pollen mother cell; tp, tapetum. (Scale bars, 20 \textmu m.)
also is required for a very early step of male and female sporogenesis.

The Evolution of Heterospory. The likelihood that \( NZZ \) has equivalent functions during the early establishment of the two types of sporangia possibly has implications with respect to questions that address the evolution of heterospory and heterosporangy. Higher plants exhibit heterospory, which is always coupled with heterosporangy (1). It is accepted that heterospory evolved from homospory, and that this process may have occurred several times independently during vascular plant evolution, including in the ancestor of the seed plants (26). However, little is known about the evolutionary transition between homospory and heterospory and the ontogenetic development of the latter. The analysis of \( NZZ \) provides genetic and molecular support for the view that the heterosporous seed plants evolved from a homosporous ancestor, and that \( NZZ \) orthologues may exist in homosporous basal vascular plants. Thus, the \( NZZ \) gene should provide an excellent tool to address these important questions in vascular plant evolution.

Note Added in Proof. While this work was in press, Yang et al. (27) reported the cloning and identification of the \( \textit{NOZZLE} \) gene under the name of \( \textit{SPOROCYTELESS} \).

We thank Heinz Saedler and ZIGIA, Max-Plank-Institut für Züchtungsforschung, Cologne, for making available the \( \textit{En-1} \)-induced mutant lines. We thank Peter Endress, Yin-Long Qiu, and Ulrike Utans Schneitz for helpful discussions and Röbi Dudler for comments on the manuscript. We also thank Antje Redweik for expert technical assistance and Jean-Jacques Pittet for the artwork. U.S. is the recipient of a postdoctoral fellowship from the Austrian Schrödingerstiftung. This work was funded by grants from the Swiss National Science Foundation (31-42175.94, 31-53032.97) to K.S. and by the Kanton of Zurich.