An Analysis of the Arabidopsis Pollen Transcriptome

BACKGROUND

In flowering plants, male gametogenesis occurs in the anthers. The division of a diploid sporogenic cell results in two cells with very different fates—the tapetal initial cell and the pollen mother cell. The tapetum, which supplies nutrients to the developing pollen, is formed from the tapetal initial cell while the pollen mother cell undergoes meiosis to produce microspores, a tetrad of haploid cells. Subsequently, microspores undergo an asymmetric mitotic division, creating a smaller generative cell enclosed in the larger vegetative cell. The vegetative cell then forms the pollen tube while the generative cell undergoes mitosis once again to form the two sperm cells (tricellular pollen grain). At what point this second division occurs is species specific. In the majority of flowering plants, it is during pollen tube growth. However, in the case of the crucifers and grasses, the division occurs while the pollen grain is still in the anther. In most species, pollen is released in a partially hydrated state and becomes fully hydrated upon contact with the stigma. The vegetative cell extends the pollen tube by tip growth, ultimately delivering the sperm cells to the embryo sac and completing the pollen development process. Pollen tube extension does not involve cell division, only cell elongation. Because of this, pollen grains have become a model system for studying cell growth.

Pollen must be able to rapidly produce the proteins necessary for germination and pollen tube growth. Due to its specialized function, pollen would be expected to have a different transcriptome than sporophytic tissues; thus, the identification of the uniquely expressed genes in pollen will aid in further studies of pollen germination.

WHAT WAS SHOWN

The pollen grain transcriptome from Arabidopsis (*Arabidopsis thaliana*) was compared with that of four vegetative tissues in a study by Pina et al. (Pina et al., 2005). This was an expansion of an earlier study by Becker et al. (Becker et al., 2003) that used ATH8 microarrays covering about one-third of the Arabidopsis genome. In both studies, viable, hydrated pollen grains were sorted from other pollen stages and cell debris with flow cytometry. Using ATH1 microarrays, representing 22,750 annotated genes, Pina et al. (2005) found 6,587 genes expressed in Arabidopsis (Columbia-0) pollen. A previous pollen transcriptome study by Honys and Twell (2004) found 7,235 genes expressed in Arabidopsis pollen. Honys and Twell (2004) examined the change in gene expression as the pollen grain matures from a uninucleate microspore to a mature pollen grain. Differences between the two studies could be due to the use of different ecotypes or the way in which the data was analyzed.

Pina et al. (2005) performed a comparative analysis of gene family and gene ontology representation in the transcriptome of pollen and vegetative tissues. Of the genes identified in this study, a relatively large number were selectively expressed in pollen (11%) compared to those found to be selectively expressed in vegetative tissues.

A functional skew toward signaling, vesicle transport, and cell wall metabolism was also found, a reflection of what the pollen tube does and similar to what was seen by Honys and Twell (2004). Interestingly, transcripts for proteins involved in transcription and translation were underrepresented. This is surprising in light of a need for protein synthesis upon pollen germination and during pollen tube growth. The authors hypothesize that the lack of representation in the pollen transcriptome could be due to presynthesis and storage of proteins involved in this process. Indeed, proteomic studies have confirmed that pollen does store synthesized proteins for later use (Holmes-Davis et al., 2005).

In general, transcription factors and RNA-processing protein transcripts were not highly represented in the pollen transcriptome with the exception of MADS-box genes. Nonclassical MADS-box genes, type I and MIKC*, were represented in the transcriptome. Although the type II family of MADS-box proteins, found not to be highly represented in the pollen transcriptome, is known to be involved in floral organ identity, the functions of the nonclassical groups is unknown.

An apparent absence of small RNA pathway components was also observed. Data from Honys and Twell (2004) was reanalyzed to assess which stages of pollen maturation genes involved in small RNA pathways are expressed. Pina et al. (2005) found expression of small RNA pathway genes in the early stages of pollen development (uninucleate microspores and bicellular pollen). However, in tricellular pollen only one gene was expressed, while most were detected in vegetative tissue tested. The authors hypothesize that this absence would allow the accumulation of small and micro RNA precursors in pollen tubes.

An accumulation of G2/M-associated cell cycle factors was also observed. The authors suggest that these factors are involved in the first mitotic division of the zygote. Cell cycle protein transcripts that were called "absent" include D3-type cyclin and subunits of the heterodimeric adenovirus E2 promotor-binding protein-dimerization partner, both of which are required for entry into the S phase.

www.plantphysiol.org/cgi/doi/10.1104/pp.104.900237

THE IMPACT

The study by Pina et al. (2005) and earlier studies (Becker et al., 2003; Honys and Twell, 2003, 2004) identified RNA present in the pollen transcriptome. However, RNA presence alone does not always ensure that the RNA is translated. Thus, despite all of the information that can be gleaned from transcriptome studies, one still cannot ascertain exactly which proteins are expressed. A logical next step would be to look at the protein content of the pollen grains. To address this, proteome studies have been done on mature pollen from Arabidopsis (Holmes-Davis et al., 2005; Noir et al., 2005) and, more recently, from rice (*Oryza sativa*) germinating pollen (Dai et al., 2007). These studies found a functional skew toward cell wall metabolism, carbohydrate/energy metabolism, and cell structure, similar to what was seen in the transcriptome studies. Protein was also found in the proteome that was not represented in the transcriptome, suggesting that protein is stored in pollen. Most notably, protein involved in protein synthesis was represented in proteome but not transcriptome studies.

In the pollen transciptome, an increase in nonclassical MADS-domain transcription factors was found in the late stages of pollen development (Honys and Twell, 2004; Pina et al., 2005). Five members of the Arabidopsis MIKC* (AtMIKC*) subgroup were characterized in a study by Verelst et al. (2007). Binding studies—both in vitro using yeast two-hybrid system and in planta—demonstrated that these proteins are able to form multiple heterodimeric complexes with high DNA-binding specificity. In vitro DNA binding as well as in vitro pollen germination experiments suggest functional redundancy between some of the complexes. Using an in silico method, Verelst et al. (2007) identified putative targets of AtMIKC* complexes in pollen, several of which have reported or proposed functions in pollen germination.

CONCLUSIONS

Analysis of the pollen transcriptome has proven to be useful starting point for many other studies, both in terms of specific genes found to be expressed as well as in allowing a glimpse into which genes are necessary for pollen growth and germination. The apparent absence of small RNA pathway genes and preferential expression of nonclassical MADS-box genes suggest that transcriptional regulation in pollen involves different players than vegetative tissues. Proteomic studies of mature and germinated pollen in both Arabidopsis (Holmes-Davis et al., 2005; Noir et al., 2005) and rice (Dai et al., 2007) confirm the functional skew observed in transcriptome studies (Becker et al., 2003; Honys and Twell, 2003, 2004; Pina et al., 2005).

LITERATURE CITED

Becker JD, Boavida LC, Carneiro J, Haury M, Feijo JA (2003) Transcriptional profiling of Arabidopsis tissues reveals the unique characteristics of the pollen transcriptome. Plant Physiol **133:** 713–725

Dai S, Chen T, Chong K, Xue Y, Liu S, Wang T (2007) Proteomics identification of differentially expressed proteins associated with pollen germination and tube growth reveals characteristics of germinated Oryza sativa pollen. Mol Cell Proteomics **6:** 207–230

Holmes-Davis R, Tanaka CK, Vensel WH, Hurkman WJ, McCormick S (2005) Proteome mapping of mature pollen of *Arabidopsis thaliana.* Proteomics **5:** 4864–4884

Honys D, Twell D (2003) Comparative analysis of the Arabidopsis pollen transcriptome. Plant Physiol **132:** 640–652

Honys D, Twell D (2004) Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biol **5:** R85

Noir S, Bräutigam A, Colby T, Schmidt J, Panstruga R (2005) A reference map of the Arabidopsis thaliana mature pollen proteome. Biochem Biophys Res Commun **337:** 1257–1266

Pina C, Pinto F, Feijo JA, Becker JD (2005) Gene family analysis of the Arabidopsis pollen transcriptome reveals biological implications for cell growth, division control, and gene expression regulation. Plant Physiol **138:** 744–756

Verelst W, Saedler H, Munster T (2007) MIKC* MADS-protein complexes bind motifs enriched in the proximal region of late pollen-specific Arabidopsis promoters. Plant Physiol **143:** 447–460

Aleel K. Grennan
University of Illinois
Urbana, IL 61801