Distinct Gene Expression Profiles in Egg and Synergid Cells of Rice as Revealed by Cell Type-Specific Microarrays

Takayuki Ohnishi, Hideki Takanashi, Mirai Mogi, Hirokazu Takahashi, Shunsuke Kikuchi, Kentaro Yano, Takashi Okamoto, Masahiro Fujita, Nori Kurata, and Nobuhiro Tsutsumi*

Graduate School of Agricultural and Life Sciences, University of Tokyo, Bunkyo-ku, Tokyo 113–8657, Japan (T. Ohnishi, H. Takanashi, M.M., H. Takahashi, N.T.); Department of Life Sciences, Faculty of Agriculture, Meiji University, Kawasaki, Kanagawa 214–8571, Japan (S.K., K.Y.); Department of Biological Sciences, Tokyo Metropolitan University, Hachioji, Tokyo 192–0397, Japan (T. Okamoto); and Genetic Strain Stock Center, National Institute of Genetics, Mishima, Shizuoka 411–8540, Japan (M.F., N.K.)

Double fertilization in flowering plants refers to a process in which two sperm cells, carried by the pollen tube, fertilize both the egg and the central cell after their release into a synergid cell of the female gametophyte. The molecular processes by which the female gametophytic cells express their unique functions during fertilization are not well understood. Genes expressed in egg and synergid cells might be important for multiple stages of the plant reproductive process. Here, we profiled genome-wide gene expression in egg and synergid cells in rice (Oryza sativa), a model monocot, using a nonenzymatic cell isolation technique. We found that the expression profiles of the egg and synergid cells were already specified at the micropylar end of the female gametophyte during the short developmental period that comprises the three consecutive mitotic nuclear divisions after megaspore generation. In addition, we identified a large number of genes expressed in the rice egg and synergid cells and characterized these genes using Gene Ontology analysis. The analysis suggested that epigenetic and posttranscriptional regulatory mechanisms are involved in the specification and/or maintenance of these cells. Comparisons between the rice profiles and reported Arabidopsis (Arabidopsis thaliana) profiles revealed that genes enriched in the egg/synergid cell of rice were distinct from those in Arabidopsis.

The life cycle of plants alternates between a haploid gametophyte stage and a diploid sporophyte stage. The gametophytes are embedded within the sexual organs of the flower. The male gametophytes develop within the anthers (McCormick, 2004), and the female gametophytes develop within the ovule (Yadegari and Drews, 2004; Fig. 1, A–C). In angiosperms, the female gametophyte is formed by meiotic division of a diploid sporophytic cell, followed by mitotic divisions of one or more haploid cells to generate a multicellular gametophyte. In the developing ovule, a megaspore mother cell undergoes meiosis to generate four spores, three of which undergo programmed cell death, leaving only the single spore as the functional megaspore in each ovule. In the most common form, the polygamous type, the megaspore then undergoes three sequential mitotic nuclear divisions to generate the eight nuclei of the mature embryo sac (Maheshwari and Johri, 1950). Subsequent cellularization results in the formation of just seven cells due to nuclear migration and eventual fusion of two nuclei in the large central cell. In the mature embryo sac, the micropylar end of the female gametophyte has an egg cell and two synergid cells, and the chalazal end of the female gametophyte has three antipodal cells (Fig. 1, B and C), which proliferate into a mass of cells before fertilization in rice (Oryza sativa; Dong and Yang, 1989).

Each female gametophyte cell type has specialized roles in reproduction. The egg cell has the potential to receive additional genetic material from the male and to stably transmit it to the next generation. Other cell types in the female gametophyte surrounding the egg cell support the passage of sperm cells into the egg cell. In particular, the synergid cells secrete attractants that guide the pollen tube to the female gametophyte (Higashiyama et al., 2001, 2003; Okuda et al., 2009) and also probably contain factors that control the cessation.
of pollen tube growth, pollen tube discharge, and gamete fusion (Weterings and Russell, 2004).

The molecular machineries underlying plant reproduction are becoming an area of active research. However, the small number and inaccessibility of these cells have hampered molecular and genome-wide studies; thus, we know little about the molecular basis of cell specification, differentiation, and function in the female gametophyte. The identification of genes expressed in the female gametophyte is essential to understanding how female gametophyte cells become specified and acquire their unique features and functions. Previous research in this field has identified genes that are expressed in specific cells of the female gametophyte in various plant species (Vrinten et al., 1999; Kasahara et al., 2005; Márton et al., 2005; Sprunck et al., 2005; Yang et al., 2006; Steffen et al., 2007; Okuda et al., 2009; Amien et al., 2010; Wang et al., 2010; Wuest et al., 2010). With the exception of Arabidopsis (Arabidopsis thaliana), microarray-based, comprehensive screens for genes exhibiting female gametophyte expression have never been done for plants.

In rice, several methods for isolating egg cells have been reported. Most require a step to degrade the cell wall with catalytic enzymes, such as cellulase (Han et al., 1998; Zhao et al., 2000; Khalequzzaman and Haq, 2005). Egg and synergid cells can also be isolated from unfertilized ovaries by manual manipulation without using enzymes (Zhang et al., 1999; Uchiumi et al., 2006). We chose this nonenzymatic method with some modifications (Takanashi et al., 2010) for the isolation of egg and synergid cells. The sizes of rice egg and synergid cells range between 30 and 50 μm in diameter, which exceeds the size of other cells released from dissected ovaries. Since many vacuoles, ranging in size from 2 to 25 μm, were present in the peripheral region of the egg cells, egg cells can be distinguished from synergid cells by their internal contrast (Fig. 1, D–G). Because an intact egg apparatus composed of a single egg cell and two synergid cells was obtained only rarely, we picked up one or two targeted cell(s) from every ovule. As the synergid cells are more delicate than egg cells, a large number of synergid cells were disrupted during the isolation procedure. We obtained 3,000 egg cells and 1,000 synergid cells from the basal portions of dissected ovaries several days before flowering (Fig. 1; Supplemental Fig. S1) and then extracted the RNA between the data obtained and the transcriptomes of a variety of diploid tissues showed that the egg and synergid cell transcriptome was distinct. We also identified and characterized a large number of the genes expressed in rice egg and synergid cells. This study provides insights into the roles of genes expressed in the female gametes before, during, and possibly after double fertilization.

RESULTS

Cell Type-Specific Microarrays Using RNA from Isolated Rice Egg and Synergid Cells

Egg and synergid cells can be isolated from unfertilized ovaries by manual manipulation without using enzymes (Zhang et al., 1999; Uchiumi et al., 2006). We chose this nonenzymatic method with some modifications (Takanashi et al., 2010) for the isolation of egg and synergid cells. The sizes of rice egg and synergid cells range between 30 and 50 μm in diameter, which exceeds the size of other cells released from dissected ovaries. Since many vacuoles, ranging in size from 2 to 25 μm, were present in the peripheral region of the egg cells, egg cells can be distinguished from synergid cells by their internal contrast (Fig. 1, D–G). Because an intact egg apparatus composed of a single egg cell and two synergid cells was obtained only rarely, we picked up one or two targeted cell(s) from every ovule. As the synergid cells are more delicate than egg cells, a large number of synergid cells were disrupted during the isolation procedure. We obtained 3,000 egg cells and 1,000 synergid cells from the basal portions of dissected ovaries several days before flowering (Fig. 1; Supplemental Fig. S1) and then extracted the RNA...
from the isolated cells for microarray analysis. RNA samples derived from egg and synergid cells were of sufficient quality for use in microarray analysis (Supplemental Fig. S2; Supplemental Table S2). We conducted a 44K microarray analysis of egg cell, synergid cell, ovary, and whole plant of *japonica* rice (cv Nipponbare). In the cell isolation procedure, isolated cells were first washed in mannitol solution and then stored transiently in it (Supplemental Fig. S1). Therefore, we needed to account for any changes in gene expression in the isolated cells caused by the transient storage in mannitol solution. To pay attention to fluctuations in gene expression caused by these washing steps, we prepared two types of ovary RNA: one was derived from ovaries just after cutting the middle portion, and the other was derived from ovaries treated with mannitol after cutting.

**Gametic, Nongametic, and Nonreproductive Cell Identities**

To examine for correlations in gene expression between samples, we created scatterplot diagrams and correlation plots for the microarray data sets. When replicates from the same cell type were compared, the points fell on the 45° identity line (Fig. 2, A and B), showing that the data was highly reproducible for microarray replicates of samples with the same biological origin (egg or synergid cell). In contrast, the points were spread widely when we compared the averaged microarray data set from egg cell samples with that from synergid cell samples (Fig. 2C). This means that the gene expression profiles of egg and synergid cells are very different. Therefore, the cell types are already differentiated with genome-wide transcriptional responses within the egg apparatus at the micropylar end of the female gametophyte during the period comprising the three sequential mitotic nuclear divisions after the creation of the megaspore. These conclusions were supported by correlation plot analysis (Fig. 2D). The expression profiles of the synergid cell were less correlated with the profiles of other samples than were the expression profiles of the egg cell.

To clearly demonstrate that the egg cell and the synergid cell had already differentiated from the diploid cells that conferred ovule identity, we created a heat map based on the expression patterns of MADS box genes that control floral organ identity (Fig. 3). We confirmed that the levels of the transcripts derived from the three classes of genes (C, D, and E) were higher in both types of ovary samples than in whole plant samples (Fig. 3). Although both egg and synergid cells develop within the ovule, our results show that transcripts from the three classes of genes do not accumulate in these haploid cell samples. In particular, the expression of *OsMADS13* was clearly suppressed. *OsMADS13* is specifically expressed in the ovule, and its expression is first detected in the ovule primordium, where it persists during further development of the ovule. Inside the ovule, *OsMADS13* is expressed in integuments and nucellus tissues (Lopez-Dee et al., 1999). This heat map is consistent with the fact that the egg and synergid cells had already differentiated from...
the diploid cells that form the ovule and that they had acquired unique gametic and nongametic cell identities, respectively.

Interestingly, the heat map of the expression patterns of MADS-box genes showed abundant transcripts of the MADS box genes controlling floral organ identity are shown. The color scale (representing the average of normalized values) is shown at the bottom. Each column of the matrix represents average expression levels of five samples: egg cell (EC), synergid cell (SC), ovary with mannitol treatment (Ova M+), ovary without mannitol treatment (Ova M−), and whole plant (WP). Class A genes are OsMADS14 (Os03g0752800, OsMADS15 (Os07g0108900, OsMADS16 (Os07g00752800, OsMADS18 (Os07g065200, and OsMADS20 (Os12g0517000); class B genes are OsMADS2 (Os01g0883100), OsMADS4 (Os05g0423400), and OsMADS16 (Os06g07127000; class C genes are OsMADS J (Os10g0273400), OsMADS3 (Os01g0217000) and OsMADS5 (Os05g0302800, class D genes are OsMADS13 (Os12g0207000) and OsMADS21 (Os10g0886200), and class E genes are OsMADS1 (Os03g0215400), OsMADS55 (Os06g0162800), and OsMADS24 (Os09g0572000), OsMADS34 (Os03g0753100), and OsMADS45 (Os08g0531700).

Figure 3. Expression patterns of MADS box genes controlling floral organ identity. Representative gene expression profiles of the MADS box genes controlling floral organ identity are shown. The color scale (representing the average of normalized values) is shown at the bottom. Each column of the matrix represents average expression levels of five samples: egg cell (EC), synergid cell (SC), ovary with mannitol treatment (Ova M+), ovary without mannitol treatment (Ova M−), and whole plant (WP). Class A genes are OsMADS14 (Os03g0752800, OsMADS15 (Os07g0108900, OsMADS16 (Os07g00752800, OsMADS18 (Os07g065200, and OsMADS20 (Os12g0517000); class B genes are OsMADS2 (Os01g0883100), OsMADS4 (Os05g0423400), and OsMADS16 (Os06g07127000; class C genes are OsMADS J (Os10g0273400), OsMADS3 (Os01g0217000) and OsMADS5 (Os05g0302800, class D genes are OsMADS13 (Os12g0207000) and OsMADS21 (Os10g0886200), and class E genes are OsMADS1 (Os03g0215400), OsMADS55 (Os06g0162800), and OsMADS24 (Os09g0572000), OsMADS34 (Os03g0753100), and OsMADS45 (Os08g0531700).

Genes Enriched in Female Gametophyte Cell Types

We characterized the gene expression profiles of the egg and synergid cells by focusing on the genes enriched in these cells. Using one-way ANOVA and multiple t test analysis, we identified enriched genes in the egg or synergid cell (see “Materials and Methods”). Based on our criteria, 44 genes (48 probes) and 56 genes (71 probes) were considered as egg cell- and synergid cell-enriched genes, respectively (Supplemental Tables S3 and S4). Among the egg cell-enriched genes (44 genes), we compared the average values of signal intensities between the egg cell and the ovary treated with mannitol and made a list of the top 10 probes by sorting the probes with the EC/Ova M+ value (Table I; these terms are defined in the tables). In a similar way, among the synergid cell-enriched genes (56 genes), we compared the average values of signal intensities between the synergid cell and the ovary treated with mannitol and made a list of the top 10 probes by sorting the probes with the SC/Ova M+ value (Table II). There was a sharp contrast in the expression profiles between the egg cell and the synergid cell, and few genes listed in Tables I and II showed a significant difference in signal intensity between the two types of ovary sample. This suggests that the cell type-dependent accumulation of the transcripts of these listed genes was not caused by artificial mannitol treatment or ovary cutting. We confirmed the expression patterns of several of the listed genes in Tables I and II by reverse transcription (RT)-PCR (Fig. 4). Each surveyed gene showed a cell type-dependent expression pattern that was consistent with the microarray data.

In Table I, two of nine genes expressed in egg cells (Os03g0296600, which we named ECA1-like1 gene, and Os11g0168000, which we named ECA1-like2 gene) share significant sequence similarity with egg cell transcripts belonging to the EARLY CULTURE ABUNDANT (ECA1) family in barley (Hordeum vulgare; Vrinten et al., 1999), wheat (Triticum aestivum; Sprunk et al., 2005), and Arabidopsis (Steffen et al., 2007; Supplemental Fig. S3). No probes for the third gene of the rice ECA1 gene family (Os12g0167601, which we named ECA1-like3 gene) were present on the microarray chip, but an RT-PCR analysis of the ECA1-like3 gene showed the same expression pattern as the other members (Fig. 4). Two genes (Os11g0187600 and Os03g0276800) encode heat shock protein 70 (HSP70). HSP70 is highly expressed in the egg cells of mouse (Mus musculus; Curci et al., 1987), newt (Pleurodeles waltl; Billoud et al., 1993), maize (Zea mays; Yang et al., 2006), and Arabidopsis (Wuest et al., 2010), and HSP70 is essential for fertilization in bovine (Bos taurus; Matwee et al., 2001) and boar (Sus scrofa domesticus; Spinaci et al., 2005). These results suggest that HSP70 abundance is a common characteristic of animal and plant eggs and that it is required for the fertilization process in both animals and plants. Os06g0285200 encodes a protein that is similar to LEAFY COTYLEDON1-like (LIL). In Arabidopsis, LIL is expressed primarily during seed development and is an essential regulator of embryo development (Kwong et al., 2003). Therefore, it seems likely that LIL is expressed in the rice egg cell before fertilization for swift embryogenesis.

In Table II, we found a gene (Os06g0162700) that had an MYB domain that was highly similar to that of Arabidopsis MYB98. In Arabidopsis, MYB98 controls the development of specific features within the synergid cell during female gametophyte development (Kasahara et al., 2005; Punwani et al., 2007, 2008). Os04g0289600 encodes allergen V5/Tpx-1 related protein, which belongs to a family of Cys-rich secretary proteins called the CAP superfamily (Gibbs et al., 2008).
Specific rice ovule autofluorescence was observed at the periphery of the egg cell-expressed ECA1-like1 gene (Os04g0289600) (Table II. Ten probes showing genes enriched in the egg cell). Values show averages of binary log values of three biological replicates, and values in parentheses are mean P values of the t test with egg cell. EC, Egg cell; Ova M+/M−, ovary with or without mannitol treatment; SC, synergid cell; WP, whole plant.

Table I. Ten probes showing genes enriched in the egg cell

<table>
<thead>
<tr>
<th>Probe Name</th>
<th>Annotation</th>
<th>EC</th>
<th>SC</th>
<th>Ova</th>
<th>WP</th>
<th>Log2 (EC/M+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Os11g0168000</td>
<td>mRNA</td>
<td>Protein of unknown function, DUFI278 family protein (ECA1-like2)</td>
<td>9.09</td>
<td>4.54</td>
<td>−6.64</td>
<td>−4.83</td>
</tr>
<tr>
<td>Os06g0296600</td>
<td>mRNA</td>
<td>Similar to ECA1 protein (ECA1-like1)</td>
<td>11.21</td>
<td>7.42</td>
<td>−3.10</td>
<td>−0.61</td>
</tr>
<tr>
<td>Os07g0598400</td>
<td>mRNA</td>
<td>Similar to heat shock protein 70</td>
<td>6.91</td>
<td>0.82</td>
<td>−6.17</td>
<td>−5.36</td>
</tr>
<tr>
<td>Os11g0187600</td>
<td>mRNA</td>
<td>Protein of unknown function, Hsp70 family protein</td>
<td>10.99</td>
<td>5.69</td>
<td>−1.69</td>
<td>−0.42</td>
</tr>
<tr>
<td>Os05g0115600</td>
<td>mRNA</td>
<td>Hypothetical conserved gene</td>
<td>10.71</td>
<td>5.51</td>
<td>−1.96</td>
<td>−1.70</td>
</tr>
<tr>
<td>Os06g0228900</td>
<td>COMBINER_EST</td>
<td>Hypothetical conserved gene</td>
<td>7.45</td>
<td>1.56</td>
<td>−3.75</td>
<td>−1.73</td>
</tr>
<tr>
<td>Os01g0299700</td>
<td>mRNA</td>
<td>Similar to Leafy cotyledon1-like L1L protein</td>
<td>4.40</td>
<td>−0.18</td>
<td>−6.64</td>
<td>−6.64</td>
</tr>
<tr>
<td>Os10g0508900</td>
<td>mRNA</td>
<td>Heat shock protein, Hsp70 family protein</td>
<td>9.64</td>
<td>4.59</td>
<td>−1.38</td>
<td>−2.03</td>
</tr>
<tr>
<td>Os06g0157000</td>
<td>COMBINER_EST</td>
<td>Hypothetical conserved gene</td>
<td>4.14</td>
<td>−6.64</td>
<td>−6.20</td>
<td>−6.37</td>
</tr>
<tr>
<td>Os07g0598400</td>
<td>mRNA</td>
<td>Glycine hydrolase family 79, N-terminal protein</td>
<td>7.02</td>
<td>5.07</td>
<td>−3.89</td>
<td>−2.80</td>
</tr>
</tbody>
</table>

To validate the microarray data, we generated promoter::GFP transgenic rice lines containing promoter regions about 3 kb upstream of the translation start site of the egg cell-expressed ECA1-like1 gene (Os11g0187600) and the synergid cell-expressed allergen V5/Tpx1-related gene (Os04g0289600). Some site-specific rice ovule autofluorescence was observed at the periphery of the embryo sac, especially at the micropylar end (Fig. 5A, white arrows). In all transgenic rice lines, we observed consistent cell type-dependent GFP expression (Fig. 5; B and C). The chalazal end of the egg cell contains developed vacuoles; therefore, GFP fluorescence driven by the egg cell-dependent promoter was localized to the micro-

Table II. Ten probes showing genes enriched in the synergid cell

<table>
<thead>
<tr>
<th>Probe Name</th>
<th>Annotation</th>
<th>EC</th>
<th>SC</th>
<th>Ova</th>
<th>WP</th>
<th>Log2 (SC/M+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Os04g0289600</td>
<td>mRNA</td>
<td>Allergen V5/Tpx1-related family protein</td>
<td>6.86</td>
<td>14.56</td>
<td>−3.84</td>
<td>−4.19</td>
</tr>
<tr>
<td>Os05g0415700</td>
<td>mRNA</td>
<td>Glycine hydrolase, family protein</td>
<td>3.57</td>
<td>12.56</td>
<td>−5.21</td>
<td>2.54</td>
</tr>
<tr>
<td>Os06g0162700</td>
<td>COMBINER</td>
<td>Similar to MYB98 (myb domain protein 98); DNA-binding/trancription factor</td>
<td>4.65</td>
<td>12.10</td>
<td>−3.26</td>
<td>−3.27</td>
</tr>
<tr>
<td>Os06g0162700</td>
<td>COMBINER</td>
<td>Similar to MYB98 (myb domain protein 98); DNA-binding/trancription factor</td>
<td>2.63</td>
<td>9.86</td>
<td>−4.72</td>
<td>−4.66</td>
</tr>
<tr>
<td>Os04g0589400</td>
<td>COMBINER</td>
<td>Conserved hypothetical protein</td>
<td>4.20</td>
<td>15.41</td>
<td>1.73</td>
<td>1.70</td>
</tr>
<tr>
<td>Os07g0222000</td>
<td>mRNA</td>
<td>Similar to α-amylase/trypsin inhibitor (RBI, RATI)</td>
<td>2.69</td>
<td>7.78</td>
<td>−5.69</td>
<td>−5.99</td>
</tr>
<tr>
<td>Os11g0629100</td>
<td>COMBINER</td>
<td>Heavy metal transport/detoxification protein domain-containing protein</td>
<td>2.57</td>
<td>6.22</td>
<td>−5.84</td>
<td>−1.35</td>
</tr>
<tr>
<td>Os04g0289500</td>
<td>COMBINER_EST</td>
<td>Allergen V5/Tpx1-related family protein</td>
<td>−0.64</td>
<td>10.56</td>
<td>−1.25</td>
<td>−1.54</td>
</tr>
<tr>
<td>Os10g0508900</td>
<td>mRNA</td>
<td>Conserved hypothetical protein</td>
<td>1.20</td>
<td>11.94</td>
<td>0.56</td>
<td>0.09</td>
</tr>
<tr>
<td>Os11g0529500</td>
<td>COMBINER_EST</td>
<td>Polyketide synthase, type III domain-containing protein</td>
<td>2.67</td>
<td>11.79</td>
<td>0.65</td>
<td>1.67</td>
</tr>
</tbody>
</table>

*A description of the build1 data set. There was no description in the latest data set.
Figure 4. Confirmatory RT-PCR analysis for genes highly expressed in egg cell and synergid cell. Expression of several genes presented in Tables I and II was tested by RT-PCR of RNA from egg cell, synergid cell, ovary with and without mannitol solution treatment, and whole plant. Os03g0296600, ECA1-like1; Os11g0168000, ECA1-like2; Os12g0167601, ECA1-like3; Os04g0289600, a gene coding allergen V5/Tpx-1-related family protein; Os06g0162700, a gene coding a protein similar to MYB98; Os04g0589400, a gene coding an unknown protein. OsMADS13 (Os12g0207000), OsMADS24 (Os09g0507200), and PSII 10-kd polypeptide (Os07g0147500) were used as negative controls; OsActin1 (Os03g0718100) was used as a positive control. EC, egg cell; Ova M+/M-, ovary with or without mannitol treatment; SC, synergid cell; WP, whole plant.

Characterization of the Gene Expression Profiles of the Rice Egg and Synergid Cells

To determine which gene functions are enriched for in the egg or synergid cell, we performed a Gene Ontology (GO) analysis of the 2,000 most highly expressed genes in the egg cell, synergid cell, and ovary after mannitol treatment (Fig. 6). The results showed significant differences in the number of genes categorized by each GO term among these samples. In the egg cell, a high number of genes, whose translated products were estimated to localize to mitochondria and to participate in the respiratory electron transport chain, were expressed. In the synergid cell, a large number of genes involved in secretion were expressed. The fact that genes coding extracellular signaling molecules are expressed preferentially in the synergid cell is a common characteristic of dicots, such as Arabidopsis (Jones-Rhoades et al., 2007), which have a high number of ovules per ovary, and the monocot rice, which has a single ovule per ovary. Furthermore, the results showed that genes involved in transcriptional regulation, such as those involved in DNA binding and transcriptional factor activity, were expressed at a significantly lower level in both the egg and synergid cells than in the ovary. On the other hand, genes involved in gene silencing, regulation of gene expression (epigenetic), and chromatin modification were expressed at significantly higher levels in haploid cell types than in diploid tissues. These results suggest that posttranscriptional regulation or epigenetic control contributes to the establishment of the specification of these haploid cells in the last stages of egg apparatus development and fertilization.

By focusing on egg cell- and synergid cell-enriched genes, we compared the rice results with those obtained previously (Wuest et al., 2010) from laser-assisted microdissection (LAM)-derived egg and synergid cells from Arabidopsis. We checked whether Arabidopsis homologs of the rice egg/synergid cell-enriched genes existed by searching for candidates with high protein sequence similarities using The Arabidopsis Information Resource (TAIR) BLAST-P. Potential candidates were also screened by the motif-pattern-based dendrogram (SALAD database) to rule out the possibility of identifying a different family member as the best Arabidopsis match. We considered candidates that passed both BLAST-P and SALAD database screening as homologs (Supplemental Tables S5 and S6). We excluded genes for which we could not determine corresponding relationships between species due to too many similar genes in the genome.

Initially, we compared Arabidopsis homologs of rice egg cell- or synergid cell-enriched genes (44 or 56 genes) with reported Arabidopsis egg cell- or synergid cell-enriched genes (163 or 144 genes; available in supplemental table S3 of Wuest et al., 2010). This showed that only a few genes were shared by Arabidopsis homologs of rice egg/synergid cell-enriched genes and Arabidopsis egg/synergid cell-enriched genes. The shared genes comprised three egg cell-enriched genes: a gene coding a Cys peptidase active site domain-containing protein (Os04g0599600); a gene coding a ubiquitin-protein ligase/zinc ion-binding protein (Os08g0451900); and a gene coding protein of unknown function belonging to the DUF1278 family (ECA1-like2, Os11g0168000), and two synergid cell-enriched genes: a gene coding a protein similar to MYB98 (Os06g0162700); and a gene coding a glycosyl transferase, family 31 protein (Os09g0452900).

Next, we surveyed a larger number of genes and examined whether Arabidopsis homologs of rice egg cell- or synergid cell-enriched genes were expressed in the Arabidopsis egg or synergid cell by examining a list of Arabidopsis egg cell- or synergid cell-expressed genes (7,171 and 5,628 genes; available in supplemental table S1 of Wuest et al., 2010; Fig. 7). Among the rice egg cell-enriched genes (44 genes), 33 Arabidopsis homologs were found (total of “Available for comparison” and “No probes” in Fig. 7A), and 16 homologs were present in the list of Arabidopsis egg cell-
expressed genes, indicating that 15 homologs were not expressed in the Arabidopsis egg cell (Supplemental Table S5). Among the rice synergid cell-enriched genes (56 genes), 43 Arabidopsis homologs were found (Fig. 7B), and 18 homologs were present in the list of Arabidopsis synergid cell-expressed genes, indicating that 20 homologs were not expressed in the Arabidopsis synergid cell (Supplemental Table S6). These results show that a large number of the Arabidopsis homologs of rice egg cell- or synergid cell-enriched genes did not have the same expression pattern in the Arabidopsis egg or synergid cell.

Figure 5. Expression analysis of genes expressed in egg and synergid cells from transgenic rice. In each panel, the column on the left shows differential interference contrast (DIC) images, the second column shows autofluorescence (AF) and GFP signals, and the third column shows merged images. A, Autofluorescence observed in the wild-type ovule (white arrows). B, ECA1-like1 pro::GFP expression in the egg cell (enclosed in the white dotted line). C, Os04g0289600 pro::GFP expression in the synergid cell (enclosed in the white dotted line). Black arrows and white arrowheads indicate the egg cell and the synergid cell, respectively. EC, Egg cell; EV, egg cell vacuole; SC, synergid cell. Bars = 30 μm.

Figure 6. Gene classification based on GO in the egg cell, synergid cell, and ovary. The 2,000 most highly expressed genes in each sample were classified by GO analysis. In this ontology, cellular component, biological process, and molecular function are treated as independent attributes. Asterisks indicate remarkable relationships, if the P value of the Pearson χ² test between the gene numbers of the two samples is below the significance level of 0.05. EC, Egg cell; Ova M+, ovary with mannitol treatment; SC, synergid cell.
From these comparisons, we conclude that some of the homologous genes have conserved functions in rice and Arabidopsis and that nearly half of the homologous genes obtained or lost their roles in the egg or synergid cell through species-specific evolution.

**DISCUSSION**

Despite previous studies in plant reproduction, the molecular machineries that underlie the development of the female gametophyte and the interactions between the female and male gametophytes have remained elusive. To bring understanding to the molecular basis for these machineries, we performed microarray analysis using pure rice egg and synergid cells. As a result, we found that the cell types at the micropylar end of the female gametophyte had already differentiated with distinct genome-wide expression patterns (Fig. 2C). Our results suggest that these cells are specified rapidly and drastically after the female gametophytic cell fate is conferred by positional information, such as by an auxin gradient. In particular, we found that the gene expression profile of the synergid cell was very different from that of other samples (Fig. 2D). This may be because the synergid cell plays unique roles in many steps of the angiosperm fertilization process, including guiding pollen tube growth toward the female gametophyte. A recent study determined the cell type-specific expression profiles in the female gametophyte of Arabidopsis by using LAM (Wuest et al., 2010). Using their data sets, we calculated the correlation coefficient between gene expression in the Arabidopsis egg and that in the synergid cell. This revealed that the correlation coefficient between the egg and synergid cell was higher in Arabidopsis ($r = 0.75$) than that in rice ($r = 0.29$; Fig. 2C). These results indicate that there is no significant difference between the gene expression profiles of the egg cell and that of the synergid cell in Arabidopsis, unlike in rice. This is a major difference between the two studies and might reflect fundamental differences in gene expression of the female gametophyte between Arabidopsis and rice or cross-contamination by neighboring cells, which is difficult to avoid in LAM.

We identified a large number of highly expressed genes in egg and synergid cells as well as genes whose expression profiles were specific to the egg or synergid cell. The cell type-dependent GFP expression in transgenic rice confirmed that our strategy was successful in identifying genes expressed in the egg or synergid cell. Although the genes listed in Tables I and II have no reported function in rice, they might be involved in reproductive events that correlate with their expression patterns in these plants and animals. In particular, we found that the genes included in two gene families, the ECA1 family and the allergen V5/Tpx-1-related protein family, show female gametophyte cell type-dependent expression patterns. The proteins derived from each gene family may have a high degree of functional redundancy. These genes cannot be identified or analyzed using traditional forward-genetic approaches, which rely on the screening of loss-of-function mutants. Our results revealed that allergen V5/Tpx-1-related proteins, which belong to a family of Cys-rich secretory proteins called the CAP superfamily, are abundant in the rice synergid cell. In *Torenia fournieri*, two secreted, Cys-rich polypeptides in a subgroup of defensin-like proteins are abundantly and predominantly expressed in the synergid cell and act as pollen tube attractants (Okuda et al., 2009). These rice CAP superfamily genes have no known function and therefore may play a role in pollen tube guidance. For the rice MYB98-like gene (Os06g0162700), each of three independent *Tos17* insertion lines of this gene exhibits a 50% sterility phenotype (http://tos.nias.affrc.go.jp/), as does the Arabidopsis *myb98* mutant (Kasahara et al., 2005). It is possible, therefore, that this gene encodes the rice ortholog of Arabidopsis MYB98. The common features of gametophytic cells in different species suggest that these essential factors for plant reproduction are conserved in angiosperms and could be involved in the molecular mechanisms of not only fertilization but also cell differentiation.

From the GO analysis, we found that a large number of genes related to mitochondrial function were expressed in the rice egg cell. This is consistent with our previous results that the egg cell has a large amount of mitochondrial DNA and that egg cell mitochondria are active in rice (Takanashi et al., 2010). In agreement with a previous report (Wuest et al., 2010), the GO

---

**Figure 7.** Comparison of rice and Arabidopsis egg and synergid cell gene expression. We compared the data with reported Arabidopsis data (Wuest et al., 2010). We investigated whether Arabidopsis homologs of the rice egg cell-enriched genes were present in a list of Arabidopsis egg cell-expressed genes (A) and whether Arabidopsis homologs of rice synergid cell-enriched genes were present in a list of Arabidopsis synergid cell-expressed genes (B). Available for comparison means genes fulfilling both of the following conditions: (1) the homolog can be found, and (2) probes for the homolog exist on the microarray chip. No probes means that there were no probes of the homolog on the microarray chip. EC, Egg cell; SC, synergid cell. Values show number of genes, and values in parentheses show their percentages.

**A** EC enriched genes

**B** SC enriched genes
analysis also indicated that not only transcriptional regulation but posttranscriptional regulation and epigenetic regulation contribute to the differentiation of the rice egg and synergid cells. Furthermore, the rice profiles and reported Arabidopsis profiles were compared to screen for genes that exhibit female gametophytic cell type-dependent expression in rice or Arabidopsis. Although the developmental process of the female gametophyte is similar in Arabidopsis and rice (both are of the polygonum type), our comparisons revealed that most of the enriched genes in the egg or synergid cell of rice were distinct from those in Arabidopsis. While differences in the methods of cell isolation and statistical analysis used in this study in rice and those used in the Arabidopsis study (Wuest et al., 2010) may have led to some discrepancies in the data, the results nonetheless suggest that the widely different sets of genes function in the egg/synergid cell between rice and Arabidopsis.

CONCLUSION

To our knowledge, this is the first comprehensive genome-wide transcriptional analysis of female gametophyte cells in monocots. The analysis showed that the gene expression profiles of rice egg and synergid cells are distinct despite their neighboring locations and that complicated machineries exist for the specification of these cells besides the auxin gradient. Our findings enabled us to decipher the differences between individual female gametophyte cell types and to identify new genes expressed in these cells. We are convinced that the transcriptional profiles of female gametophyte cells will help us understand how female gametophyte cells become specified and acquire their unique features and functions. Future research should be directed toward identifying the functions of the egg-cell- and synergid cell-enriched genes before, during, and possibly after double fertilization.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Wild-type rice plants (Oryza sativa ‘Nipponbare’) were grown in pots under natural conditions. Transgenic rice plants were grown and maintained in a biohazard greenhouse at 30°C under natural conditions. Transgenic rice plants were grown and maintained (Molecular Devices).

RNA Samples for Microarrays and RT-PCR

Ovaries were harvested from flowers in stages ranging from heading to just before anthesis (Fig. 1A). The egg and synergid cells were isolated from ovaries by nonenzymatic procedures based on an isolation method for rice egg cells described by Takanashi et al. (2010). Ovary samples were used for RNA extraction steps after cutting and removing the egg and synergid cells from the basal portion of the lower part of the cut ovary. Mannitol treatment consisted of soaking five ovaries at a time in mannitol solution for 8 h. Shoots and roots of 2-week-old rice plants were used as whole plant samples. Total RNA of each cell type and tissue was extracted with a PicoPure RNA isolation kit (Molecular Devices).

Microarray

Microarray experiments were performed as described by Suwabe et al. (2008). Briefly, extracted total RNA was quantified with the Quanti-IT Ribogreen RNA reagent and kit (Invitrogen). Three biological replicates were prepared for each sample. A rice 44K oligomicroarray (Agilent Technologies) with 42,000 oligonucleotides based on the Rice Annotation Project was used. Annotation lists available in the Rice Annotation Project database (http://rapdb.dna.affrc.go.jp/, build 5) were used in Tables I and II and Supplemental Tables S3 to S6 (Tanaka et al., 2008). Fluorescent probe labeling using the oligo (dT)-T7 strand-specific amplification method and hybridization were performed in accordance with the manufacturer’s instructions (Agilent Technologies) with slight modifications (The cDNA synthesis reaction time was extended to 6 h [2 h is the manufacturer’s instruction]). The slide images were scanned with a DNA microarray scanner (G2565BA; Agilent Technologies) using the manufacturer’s Feature Extraction software.

Data Analysis

Raw expression data sets were scaled with 75th percentile scaling (per chip) using R software (http://www.r-project.org/). To identify differentially expressed genes, the scaled data were analyzed by one-way ANOVA using one-way ANOVA in the R package (http://cran.r-project.org/). To adjust P values, a Benjamini and Hochberg (1995) correction of the false discovery rate was applied (adjusted \( P < 0.000001 \)) using R library multtest (http://cran.r-project.org/web/packages/multtest/multtest.pdf), leading to a list of differentially expressed gene sets (3,044 probes; Supplemental Table S1). The scaled data sets of the selected 3,044 probes were transformed to binary log values for further analysis using Microsoft Excel. To identify genes enriched in the egg cell, we performed four patterns of t test (EC versus SC, EC versus Ova M+, EC versus Ova M−, and EC versus WP) and fold change calculations in each pairing using Microsoft Excel. We applied a filter that considers both the fold change (log2 ratio > 3.5) and the P value (\( P < 0.01 \)) in each pairing, and we considered the genes that fulfilled these criteria among all of the pairings as egg cell-enriched genes (44 genes in 48 probes; Supplemental Table S3). Under similar methods, the gene enrichment (with 48 probes; Supplemental Table S4) were considered as synergid cell-enriched genes (Supplemental Table S4).

GO terms for the 2,000 most highly expressed genes in the egg cell, synergid cell, and ovary after mannitol treatment were obtained from the database site AgriGO (http://bioinfo.cau.edu.cn/agriGO/index.php), which was developed on the foundation of EasyGO (Zhou and Su, 2007), and GO terms were analyzed with the Web Gene Ontology Annotation Plotting tool WEGO using default settings (http://wego.genomics.org.cn/cgi-bin/wego/index.pl; Ye et al., 2006). A scaled data set with 75th percentile scaling was used in the scatterplot analysis (Fig. 2, A–C) and correlation plot analysis (Fig. 2D). Z-score-transformed data sets (per chip, using R software) were used in the heat map analysis, and the data were viewed with Java Tree View software (http://jtreeview.sourceforge.net/). Whole microarray data can be found at the National Center for Biotechnology Information Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/; accession no. GSE21074) and in Supplemental Table S1.

RT-PCR Analysis

Total RNA (5 ng) from each sample (egg cell, synergid cell, ovary with or without mannitol treatment, and whole plant) was amplified using the SMART mRNA Amplification Kit (Clontech) in accordance with the manufacturer’s instructions. First-strand cDNA was synthesized using 100 ng of each amplified mRNA, 25 μM L-1 oligo(dT)c primer, 0.5 μM deoxyribonucleotide triphosphate mix, 1× first-strand buffer, 0.01 μM dithiothreitol, 40 units of RNaseOUT (Invitrogen), and 200 units of SuperScript III reverse transcriptase (Invitrogen) in accordance with the manufacturer’s instructions. Each RT-PCR used 1 μL of first-strand cDNA as template in a 20-μL reaction with 0.2 μM of each primer. The primers used for RT-PCR and cycle numbers are listed in Supplemental Table S7.

Analysis of Promoter::GFP Fusions

Genomic sequences of promoter regions about 3 kb upstream of the translation start site of the two egg cell-enriched genes, ECA1-like1 (Os11g0187600) and ECA1-like2 (Os11g0168000), and the two synergid cell-enriched genes, allergen V5/Tpx-1-related gene (Os04g0289600) and Ornyza
Compared to the genome expression profiles of rice and Arabidopsis Egg and Synergid Cells

We compared the egg and synergid cell transcriptome of rice with the genes reported to be expressed in Arabidopsis (Arabidopsis thaliana) egg and synergid cells (Wuest et al., 2010) by examining their homologs. According to Wuest et al. (2010), 163 genes were Arabidopsis egg cell-enriched genes; 5,380 genes were Arabidopsis egg cell-expressed genes; 144 or 3,926 genes were Arabidopsis synergid cell-enriched or -expressed genes. In rice, 44 or 56 genes were rice egg cell- or synergid cell-enriched genes (Supplemental Tables S5 and S6). Homologs of rice egg cell- or synergid cell-enriched genes were present in a list of Arabidopsis egg cell- or synergid cell-enriched or -expressed genes (Supplemental Tables S5 and S6).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure S1. How the isolated cells were collected and washed.

Supplemental Figure S2. Assessments of the integrity of the RNA extracted from egg cells and synergid cells.

Supplemental Figure S3. Multiple alignment and phylogenetic tree analysis of ECA1 family proteins.

Supplemental Table S1. Microarray data.

Supplemental Table S2. RNA concentrations and integries in egg and synergid cell samples.

Supplemental Table S3. The probe list of the rice egg cell-enriched genes.

Supplemental Table S4. The probe list of the rice synergid cell-enriched genes.

Supplemental Table S5. Data validation of the genes enriched in the rice egg cell by comparing with the data for genes expressed in the Arabidopsis egg cell.

Supplemental Table S6. Data validation of the genes enriched in the rice synergid cell by comparing with the data for genes expressed in the Arabidopsis synergid cell.

Supplemental Table S7. Primer sequences for RT-PCR and plasmid construction.

ACKNOWLEDGMENTS

We thank T. Higashiyama and R.D. Kasahara for excellent technical assistance.

Received October 15, 2010; accepted November 22, 2010; published November 24, 2010.

LITERATURE CITED


