Current Biology 20, 506–512, March 23, 2010 ©2010 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2010.01.051

Report

# *Arabidopsis* Female Gametophyte Gene Expression Map Reveals Similarities between Plant and Animal Gametes

Samuel E. Wuest,[1,2] Kitty Vijverberg,[1,8] Anja Schmidt,[1] Manuel Weiss,[3,4,5,6] Jacqueline Gheyselinck,[1,9] Miriam Lohr,[7] Frank Wellmer,[2] Jörg Rahnenführer,[7] Christian von Mering,[3,6] and Ueli Grossniklaus[1,4,*]

[1]Institute of Plant Biology and Zürich-Basel Plant Science Center, University of Zürich, Zollikerstrasse 107, 8008 Zürich, Switzerland

[2]Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland

[3]Institute of Molecular Life Sciences, University of Zürich

[4]Center for Model Organism Proteomes

[5]PhD Program in Molecular Life Sciences

[6]Swiss Institute of Bioinformatics Winterthurerstrasse 190, 8057 Zürich, Switzerland

[7]Fakultät Statistik, Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany

## Summary

**The development of multicellular organisms is controlled by differential gene expression whereby cells adopt distinct fates. A spatially resolved view of gene expression allows the elucidation of transcriptional networks that are linked to cellular identity and function. The haploid female gametophyte of flowering plants is a highly reduced organism: at maturity, it often consists of as few as three cell types derived from a common precursor [1, 2]. However, because of its inaccessibility and small size, we know little about the molecular basis of cell specification and differentiation in the female gametophyte. Here we report expression profiles of all cell types in the mature *Arabidopsis* female gametophyte. Differentially expressed posttranscriptional regulatory modules and metabolic pathways characterize the distinct cell types. Several transcription factor families are overrepresented in the female gametophyte in comparison to other plant tissues, e.g., type I MADS domain, RWP-RK, and reproductive meristem transcription factors. PAZ/Piwi-domain encoding genes are upregulated in the egg, indicating a role of epigenetic regulation through small RNA pathways—a feature paralleled in the germline of animals [3]. A comparison of human and *Arabidopsis* egg cells for enrichment of functional groups identified several similarities that may represent a consequence of coevolution or ancestral gametic features.**

## Results and Discussion

The plant life cycle alternates between a diploid sporophyte and a haploid gametophyte generation. During evolution, the gametophyte generation has been reduced in size and complexity [1, 2]. Because of its simple structure, the female gametophyte of flowering plants is an ideal system to determine a complete expression map of an organism. However, its small number of cells and inaccessibility have made molecular and genome-wide studies difficult. To determine cell-type-specific expression profiles in the female gametophyte of *Arabidopsis*, we combined laser-assisted microdissection (LAM) of individual cells with the Affymetrix ATH1 GeneChip, a microarray platform commonly used in *Arabidopsis* research. LAM allowed us to dissect the cells of the mature female gametophyte with little cross-contamination (Figures 1A–1C; see Figure S1 available online). RNA isolated from 300 to 800 cells per sample was amplified via a linear amplification protocol and hybridized to ATH1 GeneChips (Table S1). Cell-type-specific transcriptomes were obtained for the synergids and the two female gametes, the egg and central cell (Figures 1D–1G; Figures S1A–S1N).

A consequence of linear amplification in combination with small input amounts of RNA is the predominant amplification of 3′ mRNA ends, resulting in a loss of signal at the 5′ mRNA end. The default algorithm generally used to test whether a gene is detectable above background levels performs poorly on data from amplified samples [4]. Therefore, we applied a novel algorithm to create present/absent calls, hereafter denoted *At*PANP (Supplemental Experimental Procedures). We assessed *At*PANP with measures of precision, such as overlaps between biological replicates and accuracy, i.e., by checking its predictive power for genes known to be expressed in the female gametophyte (69 genes; Table S2). Our new algorithm outperforms the default method on our data set (Figures S1I–S1R). Using this robust statistical method, we estimate the mature female gametophyte to express about 8,850 of the 20,777 genes present on the array (conservative estimate; Table S1). This is slightly lower than our conservative estimate of 9,220 genes expressed in pollen (male gametophyte) and sperm [5]. Because of complexity reduction during amplification, we may slightly underestimate the real transcriptome size such that mature male and female gametophytes have similar transcriptional activities.

To validate the microarray data, we used alternative approaches (Table S2; Supplemental Results) such as (1) in situ hybridization (Figures 2A–2C), (2) analysis of putative *cis*-regulatory elements driving the *GUS* reporter gene (Figures 2D–2F; Figures S2A and S2B), (3) characterization of gene and enhancer traps (Figures S2C and S2D), (4) comparison to published data (Figure 2G; Table S2), and (5) comparison to maize egg cell EST data (Figures S2E and S2F). Based on these extensive validations, we conclude that the data set reported here is accurate and can be used to predict preferential expression of genes in the female gametophyte at the level of its specific cell types.

Female gametophytic cells are closely related with regard to their cell lineage yet play distinct roles during reproduction [1, 2]. We found 1345 differentially expressed genes at a low-stringency cutoff of an unadjusted analysis of variance p value below 0.01, and 431 at a false discovery rate below 0.05. The majority of these were enriched in only one cell type, as shown by either subgrouping through pairwise t tests or hierarchical

*Correspondence: grossnik@botinst.uzh.ch

[8]Present address: Keygene N.V., P.O. Box 216, 6700 AE Wageningen, The Netherlands

[9]Present address: Department of Plant Molecular Biology, University of Lausanne, 1015 Lausanne, Switzerland

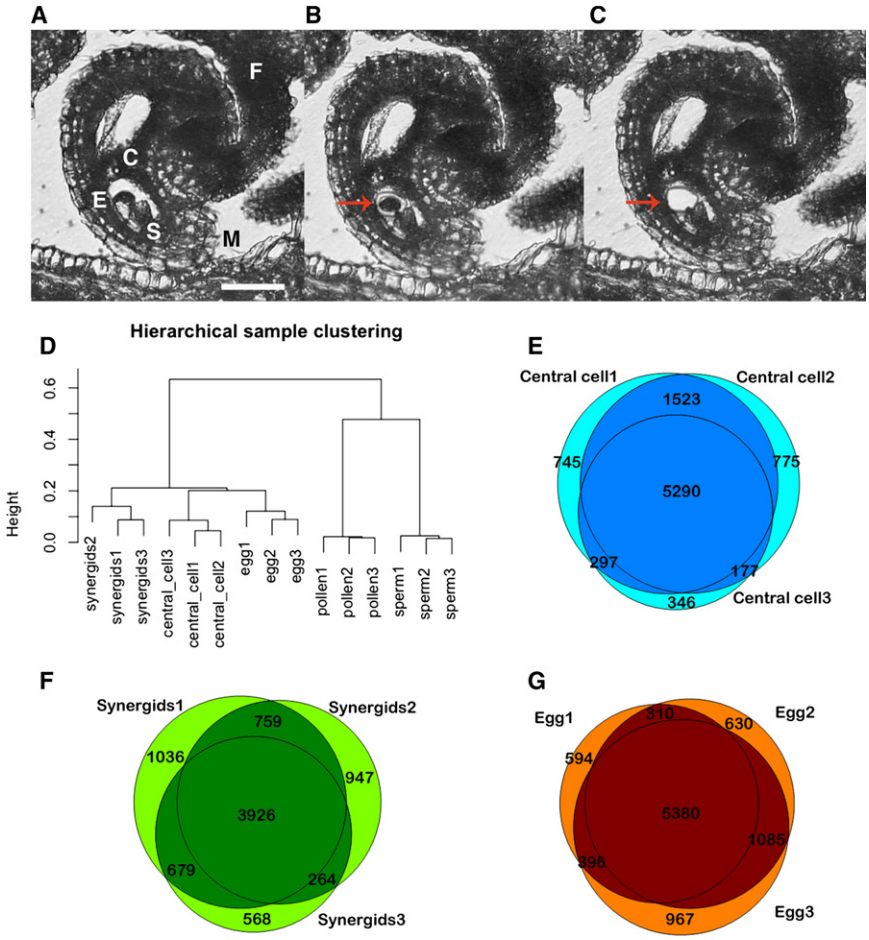



Figure 1. Laser-Assisted Microdissection and Subsequent Analysis of Transcriptomes from Populations of Individual Female Gametophytic Cell Types

(A) Dissection of the egg cell from a mature embryo sac; 8 μm section through an ovule bearing a mature embryo sac before laser microdissection with the MMI SL μCut instrument. The following abbreviations are used: M, micropyle; S, synergids; E, egg cell; C, central cell; F, funiculus. Scale bar represents 37 μm.

(B) The ultraviolet irradiation laser beam has been applied in order to isolate the egg cell (arrow). The laser cut has a diameter of 1–2 μm.

(C) The egg cell has been removed with an MMI isolation cap. After the isolation of the egg, the two remaining cell types (central cell and synergids) were collected on separate isolation caps (see also Figure S1).

(D) Hierarchical agglomerative sample clustering (correlation distance) of male and female gametophytic cell types; note that biological replicates are grouped together, demonstrating that the data are reproducible. Arrays from female gametophytic cell types form a close cluster when compared to male gametophytic cells [5].

(E–G) Overlaps of predictions of gene expression (present calls) when determined by a novel, empirical approach (*At*PANP). The algorithm determines whether a gene is expressed on an array by comparing its signal against a background distribution calculated by the use of negative probes. The Venn diagrams show present call overlaps in the three biological replicates for *At*PANP present calls (p value cutoff = 0.02; see also Table S1). Genes whose present call p values were below the cutoff in at least two of three replicates were considered present in a given cell type (darker areas): egg cell, 7171 genes; central cell, 7287 genes; synergids, 5628 genes. (E) versus (F) versus (G) are not to scale (see also Figure S1).

agglomerative clustering (Figure 3A; Table S3). This agrees with a recent study on the expression of 43 genes in the female gametophyte, of which 41 were strongly enriched in one cell type [6]. A functional gene classification and higher-level analysis suggests that the cells of the mature female gametophyte exhibit differential gene expression in distinct posttranscriptional and epigenetic regulatory mechanisms and metabolic pathways (Figure 3B). Recently, it was shown that auxin patterns the female gametophyte [7], but how this positional information controls cell specification is unknown. Our higher-level analysis did not suggest cell-type-specific differences in auxin readout. However, we identified candidates that could function in developmental events triggered by auxin: the auxin response factor *ARF17* and the polar auxin transport regulator *MKK7* genes exhibited elevated levels in egg and central cell, respectively, whereas the auxin-responsive gene *AT2G16580* was enriched in the entire gametophyte (Figure S3A; Table S3).

A dominant feature of egg cells is the relatively high expression of genes encoding the double-stranded RNA-binding factors DCL1, HYL1, and AT4G00420, a paralog of RNASE THREE-LIKE PROTEIN 1, in addition to RISC components such as AGO1. PAZ and Piwi domain-encoding genes are highly enriched among differentially expressed genes with predominant expression in the egg (Figure S3B; Table S3). In contrast, *SGS3*, involved in various gene silencing pathways, shows elevated expression in the central cell and is possibly involved in small interfering RNA (siRNA) production [8]. This suggests an important role of RNA-based silencing mechanisms in the female gametes, and the large diversity of recently discovered maternal siRNAs in developing *Arabidopsis* seeds [9] may, in part, be explained by maternal deposition of siRNAs in the female gametophyte.

In order to relate the female gametophyte transcriptome to other plant tissues, we compared it with data from 59 different tissues of the *Arabidopsis* sporophyte and male gametophyte (Table S1; Supplemental Experimental Procedures) [4, 5, 10, 11]. First we used sample clustering on binary present/absent calls to assess the overall structure of the female gametophytic cell transcriptomes. They were comparable in size and/or composition to the transcriptomes of male gametophytes or laser-captured embryos but distinct from sporophytic tissues or cell types from root and shoot, which exhibit higher expression activities (Figure 3C). Thus, male and female gametophytes share transcriptome characteristics that are distinct from those of sporophytic tissues. Interestingly, the embryo shares more characteristics with the gametophytes from which it is derived than with the adult sporophyte. Future comparisons with transcriptomes of gametophytes and sporophytes from haploid-dominant plants, such as mosses, may reveal differential gene family expansion or transfer of transcriptional modules from the haploid to the diploid generation [12].

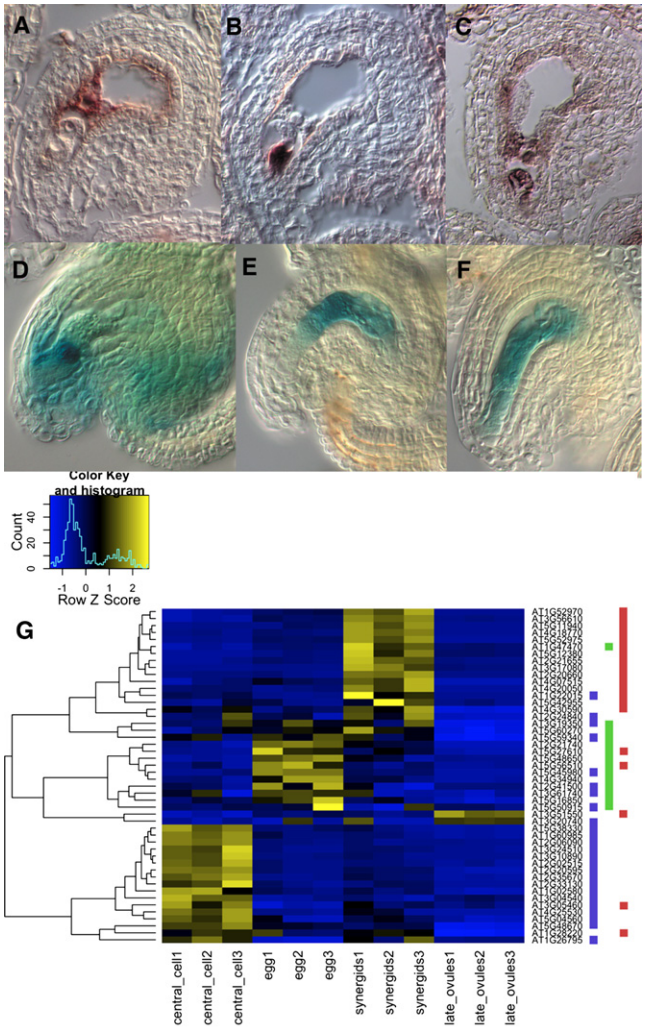



Figure 2. Data Validation by In Situ Hybridization, Promoter GUS Fusions, and Comparison to the Literature

(A–C) In situ hybridization of genes with enriched expression in the female gametophyte.

(A) *AT2G20595*, a gene with unknown function, is highly expressed in the central cell and expressed at low levels in the egg and synergids.

(B) *AT3G17080*, a self-incompatibility-related gene, is highly expressed in the synergids.

(C) The *Arabidopsis* telomerase gene *AT5G16850* shows increased expression in the egg cell and a lower expression level in the central cell.

(D–F) GUS activity in ovules expressing the *GUS* reporter gene under 5′ upstream elements (5′UE) of different genes enriched in the female gametophyte.

(D) The 5′UE of *AT5G48650*, encoding a nuclear transport factor, shows highest activity in the egg.

(E) The 5′UE of *RALF18*, a gene with putative signaling function, is highly expressed in the central cell and at much lower levels in the egg cell and synergids.

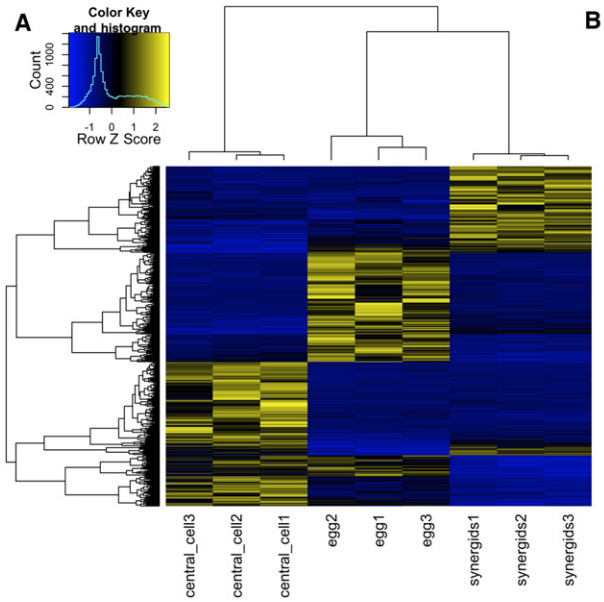
(F) The 5′UE of the MYB64 transcription factor gene *AT5G11050* shows high activity in the whole gametophyte but not in the surrounding sporophytic tissue.

(G) Heat map of expression signals of genes with described differential expression within the female gametophyte. Yellow denotes high expression, blue denotes low expression. Sample/genes were clustered via correlation distance and hierarchical agglomerative clustering, and colors are scaled per row. The color code panels on the right indicate the described preferential expression of a gene according to the literature: preferential expression in synergids, red; egg, green; central cell, blue. Note that apart from <8% disagreement out of 96 contrasts examined, the array data mirrors preferential expression within these cell types (see also Table S2).

The construction of a comprehensive tissue atlas allowed us to identify genes exhibiting enriched expression in female gametophytic cells (Supplemental Experimental Procedures). At stringent conditions, we found 420 genes significantly enriched in one of the cell types (Table S3; Figure S3C), including several genes playing a role in gametophyte development and function: *MYB98* [13] in the synergids and *FIS2* [14], *DME* [15], *CKI1* [16], *UNE6*, and *EDA28* [17] in the central cell. Thus, genes enriched in the three cell types are likely involved in cell-type-specific functions and constitute an important resource for reverse and forward genetic approaches. Recently, small, cysteine-rich defensin-like proteins (DEFLs) were implicated as signaling molecules required for pollen tube guidance in *Zea mays* and *Torenia fournieri* [18, 19]. Of the 33 (of 317) DEFLs [20] present on the array, we found seven highly enriched in the female gametophyte (Figure S3D). Six were predominantly expressed in the central cell but not the synergids, which produce the guidance signal. Whether these DEFLs act as signals remains to be examined, but they might contribute to the recently discovered role of the *Arabidopsis* central cell in pollen tube guidance [21].

We next searched for gene families or groups of genes containing a Pfam domain (Pfam groups) that are globally enriched in female gametophytic cells. Five of the ten gene sets previously found enriched in the female gametophyte [22] were also significantly enriched when examined in the more comprehensive context of our tissue atlas. Seventy-four Pfam groups and 32 gene families were enriched in at least one of the cell types or in the entire female gametophyte ($p < 0.01$; Table S3). Enriched Pfam groups contain a high number of domains of unknown function (DUFs), highlighting the lack of characterization for genes expressed in the female gametophyte (seen: 20 from 74; expected: 7.5; chi-square $p < 0.001$). Gene sets involved in transcriptional, posttranscriptional, and epigenetic regulation, signaling, and cell wall modification were enriched (Figures S3E–S3H; Table S3). Expansins were overrepresented in the transcriptomes of both male and female gametophytes (Figure S3E), as may be expected given their rapid growth, necessitating cell wall biosynthesis. Three groups of transcription factors (TFs) were overrepresented in the whole female gametophyte transcriptome, namely the RWP-RK domain, the MADS domain (predominantly type I), and the reproductive meristem TF families. It was shown that several members of these families are important in sexual plant reproduction [23–28]. Type I MADS-domain TFs were exclusively enriched in reproductive tissues, i.e., male and female gametophytes, and developing embryos (Figure S3F). Of the 28 type I MADS-domain TFs on the array, seven had highest expression in the female gametophyte—including AGL23 [26], AGL61 (DIANA) [27], and AGL80 [24], known to play a functional role in this tissue—whereas AGL62—required for endosperm cellularization [28]—exhibited highest expression in the central cell. Other family members showed highest expression in the male gametophyte (10 genes), embryo (9 genes), or seed (2 genes). These expression patterns suggest a predominant role of type I MADS-domain TFs in sexual reproduction, which may explain their highly dynamic evolutionary history [29]. Other enriched gene families, e.g., those encoding F box or leucine-rich repeat domains (Figures S3G–S3H), have also undergone rapid evolution, possibly correlated with their putative role in reproduction.

Our analysis highlighted that genes encoding PAZ, Piwi, and DUF1785 domains, mainly associated with the *Arabidopsis* Argonaute and Dicer proteins, were globally enriched in the




B

| Biol. Process | Term | Signif. | Exp. | P value |
|---|---|---|---|---|
| GO:0016246 | RNA interference | 7 | 1.3 | 0.0002 |
| GO:0006414 | translational elongation | 7 | 1.56 | 0.0006 |
| GO:0009084 | glutamine family aa biosynthetic process | 6 | 1.17 | 0.0007 |
| GO:0019243 | methylglyoxal catabolic process to D-lactate | 3 | 0.32 | 0.0025 |
| GO:0006635 | fatty acid beta-oxidation | 6 | 1.56 | 0.0036 |
| GO:0006560 | proline metabolic process | 3 | 0.39 | 0.0047 |
| GO:0006979 | response to oxidative stress | 23 | 11.86 | 0.0058 |
| GO:0009627 | systemic acquired resistance | 6 | 1.75 | 0.0067 |
| GO:0009644 | response to high light intensity | 7 | 2.33 | 0.0075 |
| GO:0035195 | gene silencing by miRNA | 3 | 0.45 | 0.0078 |
| GO:0009616 | virus-induced gene silencing | 3 | 0.45 | 0.0078 |
| GO:0007021 | tubulin complex assembly | 3 | 0.45 | 0.0078 |
| GO:0042128 | nitrate assimilation | 3 | 0.45 | 0.0078 |
| GO:0009408 | response to heat | 13 | 6.22 | 0.0089 |
| **Mol. Function** | **Term** | **Signif.** | **Exp.** | **P value** |
| GO:0016844 | strictosidine synthase activity | 5 | 0.93 | 0.0015 |
| GO:0051087 | chaperone binding | 4 | 0.6 | 0.0018 |
| GO:0004416 | hydroxyacylglutathione hydrolase activity | 3 | 0.33 | 0.0026 |
| GO:0003725 | double-stranded RNA binding | 4 | 0.66 | 0.0029 |
| GO:0009983 | tyrosine aminopeptidase activity | 2 | 0.13 | 0.0044 |
| GO:0010385 | double-stranded methylated DNA binding | 2 | 0.13 | 0.0044 |
| GO:0004356 | glutamate-ammonia ligase activity | 3 | 0.46 | 0.0083 |

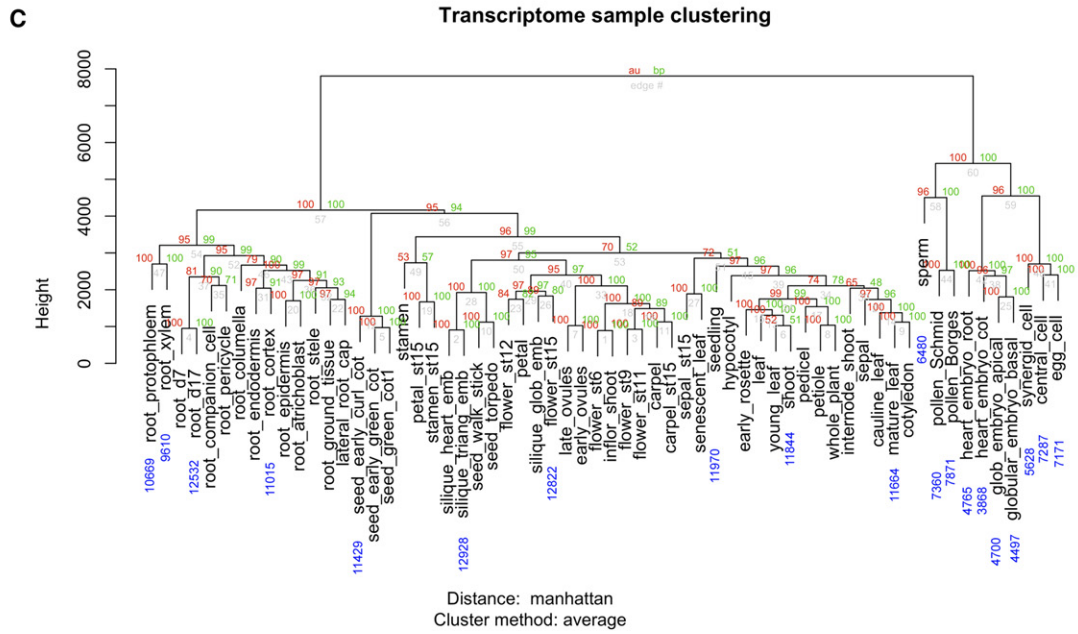



Figure 3. Female Gametophytic Cells Have Distinct Transcriptomes and Exhibit Differential Expression of Translational and Epigenetic Control Factors and Metabolic Pathways

(A) Heat map of differentially expressed genes identified by an analysis of variance (unadjusted $p < 0.01$) as a first screening method. Yellow denotes high expression, blue denotes low expression. Samples/genes were clustered via correlation distance and hierarchical agglomerative clustering. Colors are scaled per row. Note that most genes exhibit elevated expression in one cell type only, as opposed to elevated expression in two cell types.

(B) Gene ontology (GO) term enrichment table showing the significantly enriched biological processes and molecular functions among differentially expressed genes. The following abbreviations are used: signif., number of significant genes in a given term; exp., number of expected genes in a given term.

(C) Transcriptome clustering based on binary expression values (present/absent). The female gametophytic cells group together and are comparable in overall transcriptome sizes and/or compositions; female gametophyte, male gametophyte, and laser-captured embryo transriptomes form an outgroup to sporophytic tissues and cell types. Blue numbers denote overall transcriptome sizes (see also Figure S3 and Table S3).

egg. We found predominant expression of a subgroup of PAZ domain-encoding genes in the egg: among *DCL1*, *AGO1*, *AGO2*, and *AGO5*, the two functionally uncharacterized paralogs *AT5G21150* and *AT5G21030* (Figure 4A). These data suggest that small RNA pathways are a dominant feature of the generative female gamete of *Arabidopsis*. This could be important for protection against selfish genetic elements, as in the male gamete [30], or to regulate stem cell fate, paralleling epigenetic regulation in the germline of *Drosophila* and mammals (Figure 4B) [3]. In light of the recently reported genome-wide DNA hypomethylation [31, 32] and elevated production of siRNAs in the endosperm [9], we propose that

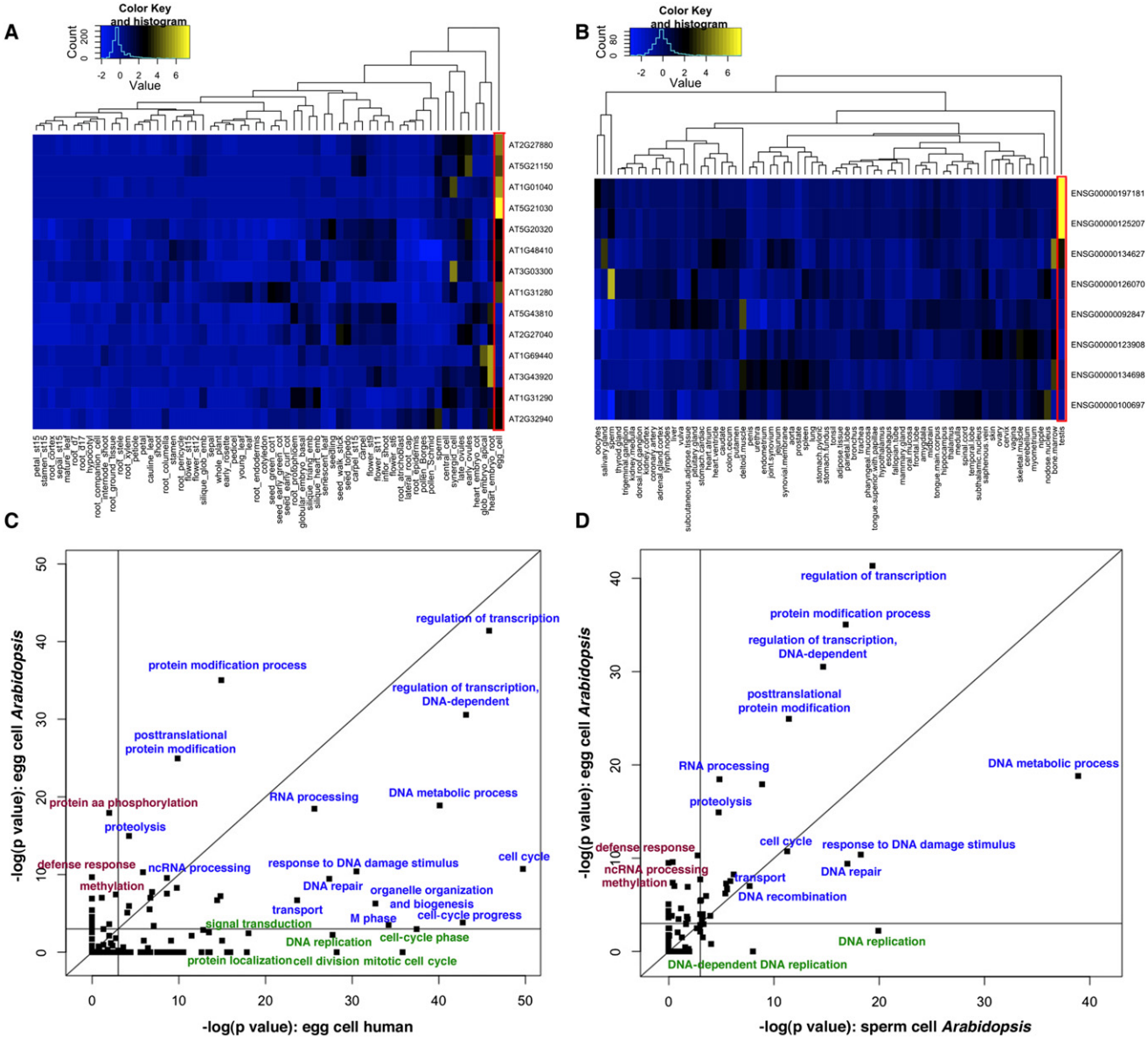



Figure 4. Gamete and Germ Cell Features in the *Arabidopsis* Egg Cell Transcriptome

(A and B) Expression of PAZ domain-encoding genes across the *Arabidopsis* and human tissue atlas. Yellow denotes high expression, blue denotes low expression. Genes/tissues are clustered via hierarchical agglomerative clustering (Euclidean distance), and signals are Z score normalized across rows.

(A) Heat map representation of mean expression across the *Arabidopsis* tissue atlas. Egg cells (red box), embryo tissues, and sperm exhibit elevated levels of expression of several genes encoding a PAZ domain when compared to the rest of the plant body.

(B) Heat map representation of mean expression across the human tissue atlas. Note that there is enriched expression of *Miwil2*, *Piwil1*, and *Piwil4* in testis (red box), as expected from their roles in germline development [3].

(C) Functional map of biological process terms comparing upregulation of GO functions in female gametes of *Arabidopsis* and humans. Negative logarithms of Bonferroni-adjusted p values from a Kolmogorov-Smirnov test for shifts toward higher signal values within a GO group are plotted. Vertical/horizontal lines indicate an adjusted p value of 0.001. Red and green denote terms that are significant in only one gamete; blue denotes terms that are significant in both. Selected data points are annotated in the figure; a full annotation can be found in Table S4.

(D) Functional map of biological process terms comparing *Arabidopsis* egg cells and sperm (as in C) (see also Figure S4 and Table S4).

embryo and endosperm development involves differential expression of epigenetic pathway components already established in the egg and central cell. Our data support a model in which siRNAs produced in the central cell and endosperm effect epigenetic gene regulation in the egg and embryo, respectively. This model could also explain egg-enriched expression of *RDM4* (Table S3), encoding a factor necessary for RNA-directed DNA methylation (RdDM) during development [33].

Sexual reproduction evolved in eukaryotes before the divergence of plants and animals. Thus, molecular aspects of gamete (syngamy) and nuclear fusion (karyogamy) may be conserved in the two lineages. Additionally, reproduction of angiosperms evolved several parallels to mammalian reproduction: (1) both lineages evolved anisogamy, (2) female gametes develop in a maternal environment providing nutrients, (3) mature gametes arrest prior to fertilization, (4) parental

imprinting evolved in both groups, and (5) selection based on male-male competition occurs in the prezygotic phase [34–36]. Thus, although it is generally not possible to compare plant and animal cell types, an interkingdom comparison of gamete transcriptomes may reveal basic molecular similarities. Therefore, we also constructed a tissue atlas of human transcriptomes, including oocytes and sperms (Supplemental Experimental Procedures), and compared expression signals within gametes against several tissues of the human body. We tested whether there is common up- or downregulation of orthologous pairs in human and *Arabidopsis* gametes but did not find a global trend (data not shown). However, from 7289 orthologous pair relations, we identified a total of 68 pairs with signals more than three standard deviations above the population mean in the eggs of both species. The latter are good candidates for genes that perform ancestral gamete functions, such as syngamy and karyogamy; however, little functional information is available for most of these genes (Table S4). The AtDRM1-HsDNMT3A orthology pair, encoding de novo DNA methyltransferases, is enriched in female gametes of both species. AtDRM1 is required for RdDM [37], providing a possible link between siRNA pathways and genome integrity maintenance that could function in both animal and plant female gametes.

We searched for overlaps of enriched functions in gametes of both lineages by comparing gene signal distributions within functional groups (GO groups or Pfam groups) across the respective tissue atlas. Functions enriched in human oocytes as detected by our analysis agreed with earlier studies (Table S4). When comparing globally enriched functions across the two species, we found overlaps of 26 "biological process" groups, 26 "molecular function" groups, and three Pfam domain groups (Figure 4A; Table S4), including RNA metabolism (transfer RNA and noncoding RNA processing, RNA polymerase activity), protein degradation, and cell-cycle control. The overlapping enrichment of the latter functional terms may indicate that, in both species, factors required for early cleavage cycles are deposited in the egg, which could be a consequence of the convergent evolution of anisogamy. That half of the female gametophytic mutants recovered to date show maternal effects [1, 17] supports the notion that, in plants, egg cells store cytoplasmic products as they do in animals [38].

Whether the functions and protein families we found enriched in the female gametes of both humans and *Arabidopsis* are indicative of conserved sexual elements or are a consequence of convergent evolution remains to be elucidated. However, 70% and 80% of the commonly enriched functional groups exhibit enrichment also in *Arabidopsis* and human sperm, respectively (Figure 4D; Figure S4). This could indicate that most represent ancestral gametic functions; however, evolutionary conclusions should await the availability of more gamete transcriptomes across different taxa. Comparisons among multiple species should allow a better dissection of gametic and gametophytic transcription modules within the female gametophyte. In addition, a better temporal resolution of female gametogenesis and early embryogenesis events will shed light on the molecular evolution of sexual processes and the transition between generations in plants and animals.

### Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, four figures, and four tables and can be found with this article online at doi:10.1016/j.cub.2010.01.051.

### Acknowledgments

We thank K. Byrne, K. Hokamp, M. Fares, K. Wolfe (Trinity College Dublin [TCD]), and G. Conant (TCD/University of Missouri-Columbia) for helpful discussions, M. Curtis (University of Zürich) for providing the pMDC162 cloning vector, T. Lehman and P. Grosscurt (University of Zürich) for access to the MMI-LAM system, C. Kägi (University of Zürich/University of Tübimgen) for testing various laser capture microscope systems, N. Kerk and T. Nelson (Yale University) for encouraging discussions, and T. Casneuf (University of Ghent/European Bioinformatics Institute Cambridge) for providing the negative probe set data. We are indebted to A. Patriagni and M. Künzli-Gontarczyk (Functional Genomics Center Zürich) for help with the array hybridizations. This work was supported by the University of Zürich, the Stiftung für wissenschaftliche Forschung through a grant of the Baumgarten-Stiftung (to U.G.), grants of the Science Foundation Ireland (to F.W.) and the Swiss National Science Foundation (to U.G.), and fellowships by the Netherlands Genomics Initiative (to K.V.) and the Deutsche Forschungsgemeinschaft (to A.S.).

Received: October 1, 2009
Revised: January 8, 2010
Accepted: January 12, 2010
Published online: March 11, 2010

### References

1. Brukhin, V., Curtis, M., and Grossniklaus, U. (2005). The angiosperm female gametophyte: No longer the forgotten generation. Curr. Sci. *89*, 1844–1852.
2. Yadegari, R., and Drews, G.N. (2004). Female gametophyte development. Plant Cell *16* (*Suppl*), S133–S141.
3. Klattenhoff, C., and Theurkauf, W. (2008). Biogenesis and germline functions of piRNAs. Development *135*, 3–9.
4. Casson, S., Spencer, M., Walker, K., and Lindsey, K. (2005). Laser capture microdissection for the analysis of gene expression during embryogenesis of *Arabidopsis*. Plant J. *42*, 111–123.
5. Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijó, J.A., and Becker, J.D. (2008). Comparative transcriptomics of *Arabidopsis* sperm cells. Plant Physiol. *148*, 1168–1181.
6. Steffen, J.G., Kang, I.H., Macfarlane, J., and Drews, G.N. (2007). Identification of genes expressed in the *Arabidopsis* female gametophyte. Plant J. *51*, 281–292.
7. Pagnussat, G.C., Alandete-Saez, M., Bowman, J.L., and Sundaresan, V. (2009). Auxin-dependent patterning and gamete specification in the *Arabidopsis* female gametophyte. Science *324*, 1684–1689.
8. Kumakura, N., Takeda, A., Fujioka, Y., Motose, H., Takano, R., and Watanabe, Y. (2009). SGS3 and RDR6 interact and colocalize in cytoplasmic SGS3/RDR6-bodies. FEBS Lett. *583*, 1261–1266.
9. Mosher, R.A., Melnyk, C.W., Kelly, K.A., Dunn, R.M., Studholme, D.J., and Baulcombe, D.C. (2009). Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. Nature *460*, 283–286.
10. Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Schölkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of *Arabidopsis thaliana* development. Nat. Genet. *37*, 501–506.
11. Brady, S.M., Orlando, D.A., Lee, J.Y., Wang, J.Y., Koch, J., Dinneny, J.R., Mace, D., Ohler, U., and Benfey, P.N. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. Science *318*, 801–806.
12. Nishiyama, T., Fujita, T., Shin-I, T., Seki, M., Nishide, H., Uchiyama, I., Kamiya, A., Carninci, P., Hayashizaki, Y., Shinozaki, K., et al. (2003). Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: Implication for land plant evolution. Proc. Natl. Acad. Sci. USA *100*, 8007–8012.
13. Kasahara, R.D., Portereiko, M.F., Sandaklie-Nikolova, L., Rabiger, D.S., and Drews, G.N. (2005). *MYB98* is required for pollen tube guidance and synergid cell differentiation in *Arabidopsis*. Plant Cell *17*, 2981–2992.
14. Luo, M., Bilodeau, P., Dennis, E.S., Peacock, W.J., and Chaudhury, A. (2000). Expression and parent-of-origin effects for *FIS2, MEA*, and *FIE* in the endosperm and embryo of developing *Arabidopsis* seeds. Proc. Natl. Acad. Sci. USA *97*, 10637–10642.
15. Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J.J., Goldberg, R.B., Jacobsen, S.E., and Fischer, R.L. (2002). DEMETER, a DNA

glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell *110*, 33–42.

16. Hejátko, J., Pernisová, M., Eneva, T., Palme, K., and Brzobohatý, B. (2003). The putative sensor histidine kinase CKI1 is involved in female gametophyte development in *Arabidopsis*. Mol. Genet. Genomics *269*, 443–453.
17. Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S., Capron, A., Xie, L.F., Ye, D., and Sundaresan, V. (2005). Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. Development *132*, 603–614.
18. Okuda, S., Tsutsui, H., Shiina, K., Sprunck, S., Takeuchi, H., Yui, R., Kasahara, R.D., Hamamura, Y., Mizukami, A., Susaki, D., et al. (2009). Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells. Nature *458*, 357–361.
19. Márton, M.L., Cordts, S., Broadhvest, J., and Dresselhaus, T. (2005). Micropylar pollen tube guidance by *egg apparatus 1* of maize. Science *307*, 573–576.
20. Silverstein, K.A., Graham, M.A., Paape, T.D., and VandenBosch, K.A. (2005). Genome organization of more than 300 defensin-like genes in *Arabidopsis*. Plant Physiol. *138*, 600–610.
21. Chen, Y.H., Li, H.J., Shi, D.Q., Yuan, L., Liu, J., Sreenivasan, R., Baskar, R., Grossniklaus, U., and Yang, W.C. (2007). The central cell plays a critical role in pollen tube guidance in *Arabidopsis*. Plant Cell *19*, 3563–3577.
22. Jones-Rhoades, M.W., Borevitz, J.O., and Preuss, D. (2007). Genome-wide expression profiling of the Arabidopsis female gametophyte identifies families of small, secreted proteins. PLoS Genet. *3*, 1848–1861.
23. Portereiko, M.F., Lloyd, A., Steffen, J.G., Punwani, J.A., Otsuga, D., and Drews, G.N. (2006). *AGL80* is required for central cell and endosperm development in *Arabidopsis*. Plant Cell *18*, 1862–1872.
24. Steffen, J.G., Kang, I.H., Portereiko, M.F., Lloyd, A., and Drews, G.N. (2008). AGL61 interacts with AGL80 and is required for central cell development in *Arabidopsis*. Plant Physiol. *148*, 259–268.
25. Franco-Zorrilla, J.M., Cubas, P., Jarillo, J.A., Fernández-Calvín, B., Salinas, J., and Martínez-Zapater, J.M. (2002). *AtREM1*, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. Plant Physiol. *128*, 418–427.
26. Colombo, M., Masiero, S., Vanzulli, S., Lardelli, P., Kater, M.M., and Colombo, L. (2008). *AGL23*, a type I MADS-box gene that controls female gametophyte and embryo development in *Arabidopsis*. Plant J. *54*, 1037–1048.
27. Bemer, M., Wolters-Arts, M., Grossniklaus, U., and Angenent, G.C. (2008). The MADS domain protein DIANA acts together with AGAMOUS-LIKE80 to specify the central cell in *Arabidopsis* ovules. Plant Cell *20*, 2088–2101.
28. Kang, I.H., Steffen, J.G., Portereiko, M.F., Lloyd, A., and Drews, G.N. (2008). The AGL62 MADS domain protein regulates cellularization during endosperm development in *Arabidopsis*. Plant Cell *20*, 635–647.
29. Nam, J., Kim, J., Lee, S., An, G., Ma, H., and Nei, M. (2004). Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. Proc. Natl. Acad. Sci. USA *101*, 1910–1915.
30. Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzić, M., Becker, J.D., Feijó, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell *136*, 461–472.
31. Gehring, M., Bubb, K.L., and Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science *324*, 1447–1451.
32. Hsieh, T.F., Ibarra, C.A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R.L., and Zilberman, D. (2009). Genome-wide demethylation of *Arabidopsis* endosperm. Science *324*, 1451–1454.
33. He, X.J., Hsu, Y.F., Zhu, S., Liu, H.L., Pontes, O., Zhu, J., Cui, X., Wang, C.S., and Zhu, J.K. (2009). A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. Genes Dev. *23*, 2717–2722.
34. Bernasconi, G., Ashman, T.L., Birkhead, T.R., Bishop, J.D., Grossniklaus, U., Kubli, E., Marshall, D.L., Schmid, B., Skogsmyr, I., Snook, R.R., et al. (2004). Evolutionary ecology of the prezygotic stage. Science *303*, 971–975.
35. Marton, M.L., and Dresselhaus, T. (2008). A comparison of early molecular fertilization mechanisms in animals and flowering plants. Sex. Plant Reprod. *21*, 37–52.
36. Randerson, J.P., and Hurst, L.D. (2001). A comparative test of a theory for the evolution of anisogamy. Proc Biol Sci *268*, 879–884.
37. Cao, X., Aufsatz, W., Zilberman, D., Mette, M.F., Huang, M.S., Matzke, M., and Jacobsen, S.E. (2003). Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr. Biol. *13*, 2212–2217.
38. Baroux, C., Autran, D., Gillmor, C.S., Grimanelli, D., and Grossniklaus, U. (2008). The maternal to zygotic transition in animals and plants. Cold Spring Harb. Symp. Quant. Biol. *73*, 89–100.