Annals of Applied Biology ISSN 0003-4746

RESEARCH ARTICLE

Transcriptome analysis of seed development in apomictic *Paspalum notatum*

S.A. Felitti[1], C.A. Acuña[2], J.P.A. Ortiz[1,2] & C.L. Quarin[2]

1 Laboratorio de Biología Molecular, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Zavalla, Argentina
2 Instituto de Botánica del Nordeste (IBONE), CONICET, Facultad de Ciencias Agrarias, Universidad Nacional del Nordeste, Corrientes, Argentina

Keywords
Endosperm; endosperm balance number; extracellular ATP; *Paspalum notatum*; pseudogamy; transcriptome.

Correspondence
Dr S.A. Felitti, Laboratorio de Biología Molecular, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Parque Experimental Villarino, S2125ZAA Zavalla, Provincia de Santa Fe, Argentina. Email: sfelitti@unr.edu.ar

Received: 24 August 2014; revised version accepted: 1 January 2015.

doi:10.1111/aab.12206

Abstract

The seed developmental process involves various tissues with several ploidy levels and different genetic origins. Therefore, its characterisation at the transcriptome level is certainly a challenge. The hypothesis of endosperm balance number (EBN) postulates that each species has an effective number that is important for normal endosperm and seed development to occur. Understanding endosperm formation in apomictic plants is crucial for the perspective of transferring apomixis to sexual species of agronomic interest. Since sexual tetraploid *Paspalum* plants fit the EBN premise, the EBN insensitivity observed in apomictic plants might be a requirement for the spread of pseudogamous apomixis. Crosses using several cytotypes of *Paspalum notatum* were made in order to induce the development of seeds with different maternal/paternal genomic ratios in the endosperm. A transcriptome characterisation of ovaries 3 h after pollination was performed using cDNA-AFLP methodology. Forty-six of the 100 differentially expressed transcript-derived fragments (DETDFs) were specifically found in crosses in which apomictic plants were used as the female parent and presented a predicted m : p ratio in the endosperm that was different to the 2:1 requirement of the EBN. Moreover, 12 of the DETDFs presented identity with proteins that were differentially expressed in response to changes in the levels of extracellular ATP (eATP) in Arabidopsis cell suspension cultures. eATP is an important molecular switch in plants that tightly controls organellar energy metabolism and activates gene expression controlling specific growth and developmental programmes. The results suggest that eATP-mediated signalling could be involved in the regulation of endosperm development.

Introduction

Seed formation is a key process in the life cycle of flowering plants. Seed development is initiated by the double fertilisation event with one of the sperm cells fertilising the egg cell and the other the central cell, leading to the formation of the embryo and the endosperm, respectively. Coordinated development of the embryo and endosperm, each of them with different ploidy levels, takes place within the embryo sac, which is surrounded by layers of nucellar cells and integuments that are somatic tissues of maternal origin (Chaudhury *et al.*, 2001; Berger & Chaudhury, 2009). Therefore, the interrelationship of the different genomes is what determines the successful development of the seed. Given that the seed developmental process involves various tissues with several ploidy levels and different genetic origins, its characterisation at the transcriptome level is certainly a challenge (Chaudhury *et al.*, 2001).

In most diploid species the endosperm is a triploid tissue inheriting two maternal genomes and one paternal genome. It can probably best be termed a *new structure*, one of complex morphological nature, characteristic of the angiosperms only (Eames, 1961). Endosperm develops rapidly compared to the embryo. It is often partially or fully resorbed by the embryo early in development resulting in seeds that are exalbuminous, or it may persist until germination resulting in albuminous seeds (Brown

& Lemmon, 2007). Plant species that are members of the *Poaceae*, form seeds that are albuminous: the endosperm is a prominent part of the mature seed. Each single-seeded fruit, or caryopsis, is produced from an individual carpel, which itself contributes seed coats and pericarp tissue to the mature seed (Milligan *et al.*, 2004). The endosperm plays the vital role of 'gatekeeper' in the angiosperm life cycle by mediating nutrient transfer from mother to offspring (Brown & Lemmon, 2007). It is also a major site of gene imprinting and a tissue where inappropriate hybridizations/polyploidy are detected, thus preventing a futile investment in unfit seeds (Gehring *et al.*, 2004).

It is currently accepted that most angiosperms require certain balance of maternal and paternal genomic contributions for the successful development of the endosperm. The hypothesis of endosperm balance number (EBN) postulates that each species has an effective number which may not necessarily be a direct reflection of its ploidy level (Johnston *et al.*, 1980). The EBN must maintain a 2:1 ratio (maternal genome : paternal genome) for normal endosperm and seed development to occur.

Mutations in Arabidopsis genes *HAIKU* (*IKU*) cause a decrease in the size of the endosperm and consequently, in the size of the embryo and seed (Garcia *et al.*, 2003). The smaller size observed in the endosperm of *iku* mutants is accompanied by a reduction of cell elongation of the integuments, indicating a communication between these two genetically distinct components of the seed (Garcia *et al.*, 2005). Moreover, studies conducted in carrot and corn indicated that the endosperm is a source of signals involved in embryogenesis (Opsahl-Ferstad *et al.*, 1997; van Hengel *et al.*, 1998). The mechanism by which the embryo adjusts the degree of cell proliferation to the available space is not understood. However, this observation provided evidence of the existence of a communication from the endosperm to the embryo. Furthermore, the existence of signals from the embryo towards the endosperm was revealed by the analysis of a *cyclin dependent kinase* (*cdc2a*) mutant (Nowack *et al.*, 2006). In cdc2a mutant pollen, only one sperm cell, instead of two, was produced. Mutant pollen was viable but could fertilise exclusively the egg cell in the embryo sac, allowing for a genetic dissection of the double fertilisation process. The development of the embryo started divisions of the unfertilized central cell, suggesting that a previously unrecognised positive signal from the fertilisation of the egg cell initiated proliferation of the central cell (Nowack *et al.*, 2006). While the embryo evolved to the globular stage, the development of the endosperm in *CDKA1* seeds stopped after a few cycles of cell division. Seeds obtained were small because there was no expansion of the integuments. These observations provided evidence of the essential role of the endosperm compared to relatively limited effect exerted by the embryo on the control of seed size.

Paspalum notatum Flüggé is used as a model for studies of species that reproduce by apomixis. The species belongs to the *Poaceae* family and presents albuminous endosperm. It has races with different ploidy levels and linked reproductive characteristics. Diploid plants (2n = 2x = 20) reproduce sexually (Burton, 1955) whereas polyploids (3×, 4×, 5×, 6×) reproduce asexually by apomixis. The tetraploid cytotype (4n = 4x = 40) is the most common and widely distributed while the other polyploid cytotypes are rare or were experimentally obtained (Gates *et al.*, 2004). In tetraploid apomictic *P. notatum* plants, endosperm formation requires fertilisation of polar nuclei (pseudogamy) (Burton, 1948; Quarin, 1999). The fertilisation of two joined unreduced polar nuclei (2n + 2n) by a reduced sperm nucleus (n) produces a 4:1 maternal to paternal (m : p) genome ratio which is different from the ratio in sexual angiosperms (Grimanelli *et al.*, 1997).

Several research groups have made significant contributions to knowledge of the biology and inheritance of apomixis in different species (reviewed by Ozias-Akins & van Dijk, 2007; Tucker & Koltunow, 2009; Pupilli & Barcaccia, 2012). Most of these studies have focused on identifying genes involved in the clonal embryo formation in different apomictic species. However, there have been very few studies aiming to comprehend endosperm development in these systems (Quarin, 1999; Albertini *et al.*, 2004; Polegri *et al.*, 2010; Sharbel *et al.*, 2010). An understanding of the formation of endosperm in apomicts is crucial for the perspective of transferring apomixis to species of agronomic interest. Quarin (1999) investigated the effect of different sources and ploidy levels of pollen on endosperm formation and seed production in aposporous tetraploid (2n = 4x = 40) *P. notatum*. Results indicated that apomictic 4× *P. notatum* is a pseudogamous species with effective fertilisation of the two unreduced (2n) polar nuclei by a reduced (n) sperm. Endosperm development and seed production occurred independently of the species or the ploidy level of the pollen donor. Quarin observed that in sexual *P. notatum* plants the EBN is effective and seed is produced only when the m : p ratio is 2:1. However, there is an evident EBN insensitivity in the apomictic plants. The EBN insensitivity observed in apomictic plants could have arisen as an imprinting consequence of a high maternal genome contribution (Quarin, 1999).

The aim of this work was to characterise the transcriptome during EBN-insensitive endosperm formation, benefitting from a set of experimental crosses designed to generate a diversity of endosperm genome dosages and seed set outcomes including both sexual and apomictic

Table 1 *Paspalum notatum* plants identified according to accessions, ploidy levels, reproductive systems and their source

Plant Identification[a]	Accession	Ploidy Level And Chromosome Number	Type of Reproduction	Source
2x S1	H398	2x = 20	Sexual	Empedrado, Corrientes, Argentina
2x S2	Tifton 9	2x = 20	Sexual	An individual plant of the commercial cultivar Tifton 9
2x S3	Q4084-8	2x = 20	Sexual	An individual plant collected in a natural population at Cayastá, Santa Fe, Argentina
4x A1	Q4117	4x = 40	Apomictic	Unknown specific locality, state of Río Grande do Sul, Brazil
4x A2	Q3775	4x = 40	Apomictic	Municipality of Gómez, Tamaulipas, Mexico
4x S1	Q4188	4x = 40	Sexual	Sexual hybrid, parents: Q3664 (4x predominantly sexual, Tifton, USA) x Q3853 (Capivarí, Rio Grande do Sul, Brazil)
4x S2	Q4205	4x = 40	Sexual	Obtained by self-pollination of the facultative apomictic 4x accession Q3664
4x S3	C4-4x	4x = 40	Sexual	Induced tetraploid derived from a chromosome-duplicated callus sector obtained by tissue culture and colchicine treatment of a diploid
6x A	Q4308	6x = 60	Apomictic	B_{III} hybrid obtained by 2n + n fertilisation of apomictic 4x accession Q4023

[a]Plants are classified by ploidy levels (2x = diploid, 4x = tetraploid, 6x = hexaploid) and also by reproductive system (S = sexual, A = apomictic).

scenarios in *P. notatum*. The study was conducted taking advantage of the availability of *P. notatum* races with different ploidy levels as pollen sources.

Materials and methods

Plant material and crosses

Experimental crosses using several races of *P. notatum* were made in order to induce the development of seeds with different maternal/paternal genomic ratios in the endosperm. The plants differed in ploidy levels and reproductive systems: two were apomictic tetraploids: 4× A1 and 4*x* A2; three were sexual tetraploids: 4× S1, 4× S2 and 4× S3; three were sexual diploids: 2× S1, 2× S2 and 2× S3; and one was an apomictic hexaploid: 6× A. The origin of these plants and the different combinations of crosses are described in Tables 1 and 2, respectively. The parental combinations of crosses were selected based on the information generated by Quarin (1999). In that previous study, sexual diploid and tetraploid plants of *P. notatum* only produced seed when there was an endosperm genome ratio of 2:1 (maternal genome : paternal genome) as it was expected for sexual species considering the EBN theory. However, apomictic tetraploid plants of *P. notatum* did produce seed in spite of the pollinator being diploid or polyploid (Quarin, 1999). In *P. notatum* double fertilisation occurs between 30 min and 2 h after pollination (C.L. Quarin, personal communication). Thus, ovaries were isolated 3 h after pollination in order to detect differences during the early formation of *P. notatum* seeds. A minimum of 20 ovaries for each cross were harvested and frozen in liquid nitrogen.

A day before anthesis, rooted culms bearing panicles due to flower the following day were collected and immediately placed in a jar with water. The collected culms were then placed in an artificial fog chamber which started misting the next day, 2 h before sunrise. Anthesis occurred around sunrise and the high level of air humidity prevented the dehiscence of anthers. Since floral maturation along an inflorescence varies, spikelets that were at anthesis were selected based on stigma and anther protrusion and marked with a waterproof permanent marker. Sharp pointed tweezers were used to remove the anthers. Emasculated inflorescences were dusted with the desired pollen. Inflorescences were covered with glassine bags after pollination to prevent contamination with pollen from undesired sources. The jars were then placed in a shaded corner of the glasshouse where they remained for 3 h after pollination.

Emasculated inflorescences of the control plant (Q4117, 4× A, emasculated and unpollinated) were

Table 2 Crosses between different genotypes of *Paspalum notatum* with different ploidy levels and reproductive systems

Female Parent[a]	Pollinator[a]	Expected Ploidy of the Embryo	Expected Ploidy of the Endosperm	Expected m : p Genome Ratio in the Endosperm
4x A1	2x S1[b]	4x (2n + 0)	9x (2n + 2n + n)	8:1
	2x S2[b]	4x (2n + 0)	9x (2n + 2n + n)	8:1
	4x A2[b]	4x (2n + 0)	10x (2n + 2n + n)	4:1
	4x S3[b]	4x (2n + 0)	10x (2n + 2n + n)	4:1
	6x A[b]	4x (2n + 0)	11x (2n + 2n + n)	8:3
4x S1	2x S1[c]	3x (n + n)	5x (n + n + n)	4:1
	4x A2[b]	4x (n + n)	6x (n + n + n)	2:1
4x S2	2x S2[c]	3x (n + n)	5x (n + n + n)	4:1
	6x A[c]	5x (n + n)	7x (n + n + n)	4:3
4x S3	2x S3[c]	3x (n + n)	5x (n + n + n)	4:1

[a]Parents are classified by ploidy levels (2x = diploid, 4x = tetraploid, 6x = hexaploid) and also by reproductive system (S = sexual, A = apomictic).
[b]Crosses that are expected to produce seed normally.
[c]Crosses that are not expected to produce seed.

covered with a clean glassine bag to prevent contamination with pollen from undesired sources. The jars were then placed in a shaded corner of the glasshouse where they remained for 3 h. Ovaries of the control plant were harvested 3 h after emasculation.

RNA isolation and cDNA synthesis

Total RNA was extracted from frozen ovaries using SV Total RNA Isolation System (Promega, Madison, WI, USA), according to the manufacturer's protocol. RNA quality and yield were assessed by spectrophotometry and agarose gel electrophoresis as described by Sambrook & Russell (2001). Double stranded cDNA was synthesised and purified as indicated in Vuylsteke *et al.* (2007) with minor modifications. Briefly, 10 µL total RNA for each sample (approximately 0.3 µg) were used for first strand cDNA synthesis with a biotinylated oligo-dT primer and the SuperScript Reverse Transcriptase (Invitrogen Life Technologies, California, CA, USA) according to the manufacturer's instructions. Then, 140 µL second strand cDNA synthesis mix [16 µL 10× second strand buffer, 3 µL 10 mM dNTP mix, 0.16 µL ribonuclease H (10U μL^{-1}), 3.5 µL DNA polymerase I (10U μL^{-1}) and 117 µL H_2O] were added to this reaction mixture and incubated for 1 h at 12°C, followed by 1 h at 22°C.

cDNA-AFLP analysis

A cDNA-AFLP analysis was performed as described by Vuylsteke *et al.* (2007) and Xiao *et al.* (2009) with some modifications. Transcript-derived fragments (TDFs) were generated by digestion of the ds cDNAs with *CviA*II and *Taq*I restriction enzymes (New England Biolabs, Ipswich, MA, USA) as indicated in Stölting *et al.* (2009). After generation of double stranded cDNA, 320 µL of buffer NT (Nucleospin Extract II kit, Macherey-Nagel, Düren, Germany) were mixed with 160 µL of second strand reaction mixture. A NucleoSpin Extract II column was placed into a 2-mL Nucleospin collecting tube and loaded with the sample. Then, it was centrifuged for 1 min at 11 000 *g* and the flow-through discarded. The column was washed with 600 µL of NT3 buffer and the sample was eluted using 30 µL of NE elution buffer. The cDNA (20 µL) was digested with the first restriction enzyme, *CviA*II in a final volume of 40 µL, for 2 h at 25°C. After this time, 40 µL (100 µg) of resuspended streptavidin-coated Dynabeads (Promega) were added to the digested cDNA fragments to give a final volume of 80 µL and incubated at room temperature for 30 min. The immobilised biotinylated 3′-terminal cDNA fragments were collected with a magnet and the supernatants removed with a pipette. The beads were released from the magnet and washed 5 times with 100 µL 1× STEX buffer (1 M NaCl, 10 mM Tris–HCl pH 8.0, 1 M EDTA pH 8.0 and 1 mL/100 mL Triton X-100). Finally, the beads were collected with the magnet, buffer STEX was removed with a pipette and the beads were resuspended in 30 µL TE buffer (10:0.1). For the second digestion step, 10 µL of the second digestion mix (10× buffer, 10U *Taq*I and H_2O) were added to the 30 µL bead suspension, and incubated for 2 h at 65°C with gentle agitation to ensure that the beads are resuspended. After that time, the beads were collected using the magnet and the supernatant (containing the liberated template fragments) transferred to a fresh tube for adapter ligation. The TDFs were then subjected to adaptor ligation as described by Vuylsteke *et al.* (2007), except that the ligation was used as template in the pre-amplification reaction without dilution. The sequences of primers and adaptors used for cDNA-AFLP analyses are listed in Table S1, Supporting Information.

A pre-amplification reaction was performed using the adaptor ligated double stranded cDNA fragments as templates and oligonucleotides complementary to the corresponding adaptors as primers as described by Xiao *et al.* (2009). In a second round of amplification, independent subpopulations of the pre-amplified cDNA fragments were selectively re-amplified. The selective primers used during this step were identical to the pre-amplification primers but were extended by one nucleotide at the 3′ end. Pre-amplified cDNA fragments were diluted one fourth and used as templates in selective amplification reactions. PCR profiles were as follows: 95°C, 5 min; 12 cycles of 94°C, 30 s; 65°C (decrease of 0.7°C each cycle), 30 s; 72°C, 1 min, and 24 cycles of 94°C, 30 s; 55°C, 30 s; 72°C, 1 min; followed by a final extension step (72°C, 10 min) as indicated by Xiao *et al.* (2009). Amplification reactions were supplemented with 10 µL of loading buffer (98% m v^{-1} formamide, 10 mM EDTA pH 8.0, bromophenol blue and xylene cyanol) denatured at 96°C for 3 min and loaded onto 6% denaturing polyacrylamide gels. A molecular weight marker (100–1500 bp; New England Biolabs) was loaded into the polyacrylamide gels in order to estimate the molecular weights of the bands obtained by cDNA-AFLP. Electrophoresis was carried out at 60 W for approximately 3 h using 0.5× TBE and 1× TBE pH 8.0 buffers (100 mM Tris–HCl, 90 mM boric acid, 1 mM EDTA pH 8.0) in the upper and lower tanks, respectively. Gels were stained using the DNA Silver Staining System (Promega). To assess the reproducibility of the electrophoresis results, two aliquots of several PCR reactions were run in different gels.

Isolation and sequencing of DETDFs

Gels run to isolate DETDFs were loaded with the amplification reactions obtained using the most informative

primer combinations and technical replicates (separate PCR reactions employing the same selective primer pair). Gel slices containing DETDFs (bands present, or showing significant higher intensity, in at least one cross compared to the other crosses in the same gel) were identified and excised from the polyacrylamide gel, crushed with a micropipette tip and incubated in 30 µL of elution buffer (0.5 M NH_4Ac, 1 mM EDTA pH 8.0) for 4 h at 37°C with occasional vortexing. The eluted DNA fragments were then PCR re-amplified using 1 µL of the eluted sample as a template, the respective primer combination used during the second round of amplification and the conditions described for the selective amplification reactions. The resulting PCR products were checked on 2% ($m\,v^{-1}$) agarose gels as described by Sambrook and Russell (2001). Those presenting with the correct size and quality were sent to Macrogen for sequencing analysis (Macrogen, Seoul, Korea). Sequence data from this article have been deposited in the GenBank database under the accession numbers JZ515891 to JZ515970.

Gene function analysis

High quality, unique sequences were analysed for best homology to infer protein functions. Database searches were performed using the BLAST Network Service (NCBI, National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/BLAST). Sequences were blasted on the non-human, non-mouse expressed sequence tag (EST) database (BLASTn) and to a non-redundant protein sequence database (BLASTx). The best EST hits were then blasted on a non-redundant protein sequence database using BLASTx. To infer the most likely biological function, the sequences of the best matching protein found in all searches were blasted against the Arabidopsis informative resource (TAIR) protein database using BLASTp. Best-hit proteins were submitted to the GO annotation tool at TAIR to search for putative molecular functions and related biological processes. *In silico* mapping onto rice chromosomes was performed with the best hits on The Institute for Genomic Research (TIGR) database using the BLASTn option.

Embryo sac observations

Inflorescences were fixed in FAA (18 70% v/v ethanol : 1 37% v/v formaldehyde : 1 glacial acetic acid) 48 h after pollination. After 24 h, inflorescences were transferred and stored until the analysis in 70% v/v ethanol. The pistils were then dissected out of the florets and cleared for approximately 2 h using 3% v/v H_2O_2. The rest of the clearing procedure was accomplished following the technique developed by Young *et al.* (1979). Ovules were rinsed for 30 min in each of the following solutions: 50% v/v ethanol, 70% v/v ethanol, 95% v/v ethanol and finally, absolute ethanol twice. Then, the samples were incubated for 30 min in each of the following mixes: 50% v/v methyl salicylate plus 50% v/v ethanol, 75% v/v methyl salicylate plus 25% v/v ethanol, 85% v/v methyl salicylate plus 15% v/v ethanol and finally, 100% v/v methyl salicylate overnight. The ovaries were observed in order to determine the degree of development of embryo and endosperm using a microscope equipped with a differential interference contrast (DIC) system. A minimum of 20 ovules were observed from at least two different inflorescences.

Real-time PCR

Total RNA was obtained using SV Total RNA isolation system (Promega) according to the instructions provided by the manufacturer in two independent experiments. Two biological replicates (different RNA extractions from two experimental crosses) were used in real-time PCR experiments: one of the RNA samples was the same that was used in the cDNA-AFLP analysis, and the second RNA sample was obtained from ovaries of a replicated cross of *P. notatum* genotypes. This isolation system included a step for DNAse treatment. cDNA was synthesised from 200 ng total RNA with the Superscript II enzyme (Invitrogen Life Technologies) and d(T)15 oligonucleotide, following the manufacturer's instructions.

DETDF-specific primer sequences were designed with Primer3 v.0.4.0 software (Rozen and Skaletsky 2000). Primer sequences used to study the expression of selected DETDFs and the expected sizes of PCR products are indicated in Table S2. Real-time PCR analysis was performed using the Rotor-Gene Q (Qiagen®, Hilden, Germany) thermal cycler. Reactions were performed on two biological replicates (different RNA extractions from two experimental crosses), using six technical replicates. The reaction contained 1× SYBR Green PCR Master Mix (Mezcla Real®, Biodynamics, Buenos Aires, Argentina), 400 nM of the forward and reverse gene-specific primers and 1 µL cDNA (1:5 dilution) in a final volume of 15 µL. No-template controls were also included. Each cycle consisted of denaturation for 15 s at 94°C, annealing for 60 s at 58°C or 60°C (depending on the primer pair), and extension for 60 s at 72°C. Specificity of the amplification reactions was assessed by melting curve analyses, which were run at 95°C for 15 s and 60°C for 15 s followed by an increase in temperature from 60 to 85°C ($0.2°C\,s^{-1}$). Quantification cycle (Cq) and efficiency (E) for each amplicon were obtained from the Comparative Quantitation software supplied by Corbett

Research for Rotor Gene. *β-tubulin* was probed to be stable throughout the reproductive developmental stages in sexual and apomictic *P. notatum* (Felitti *et al.*, 2011; Podio *et al.*, 2014). Therefore, *β-tubulin* was selected as a suitable reference gene to analyze gene expression levels in *P. notatum* ovaries (Pfaffl *et al.*, 2004). Normalised expression value for each gene was calculated based on amplification efficiency (*E*) and Cq in comparison to the reference gene according to Simon's formula (Simon, 2003). Data were tested for statistical significance using the Kruskal Wallis test. Statistical analyses were performed using agricolae package of R software (http://www.r-project.org/).

Results

Analysis of differentially expressed genes during seed formation in *Paspalum notatum*

In order to characterise the transcriptome at the start of embryogenesis and seed formation, crosses involving several races of *P. notatum* were performed to obtain a diversity of endosperm genome dosages and seed set outcomes including both sexual and apomictic scenarios. The races used differed in ploidy levels and reproductive systems (Table 2). Ovaries were isolated 3 h after pollination and the transcriptome characterisation of the samples was conducted using cDNA-AFLP methodology.

A total of 16 primer combinations were used for this analysis (Table S1). The expression profiles were highly reproducible. Each combination of primers produced, on average, patterns of approximately 30 fragments, which ranged from 100 to 800 bp. A section of a typical cDNA-AFLP polyacrylamide gel profile is shown in Fig. 1. The control was obtained from an emasculated tetraploid apomictic plant (4*x* A1) (Table 1) that was then not pollinated and ovaries isolated 3 h after emasculation. Since seed development did not start in this control plant, the expression pattern obtained should reflect the transcriptome present in the ovule at the onset of the process, just prior to double fertilisation. A very low number of bands was observed in this sample compared to most of the crosses analysed (Fig. 1).

A total of 3960 scorable TDFs were detected. Of these, 423 bands were isolated from gels loaded with five to six crosses using the most informative primer combinations and technical replicates, PCR re-amplified and fragment sizes were confirmed in agarose gels (data not shown). These TDFs differed in either presence/absence or intensity along the different crosses and therefore were considered to be differentially expressed (DETDFs). Hundred DETDFs were successfully sequenced (Table 3), presenting a size range of 102–667 bp with an average length of 384 bp.

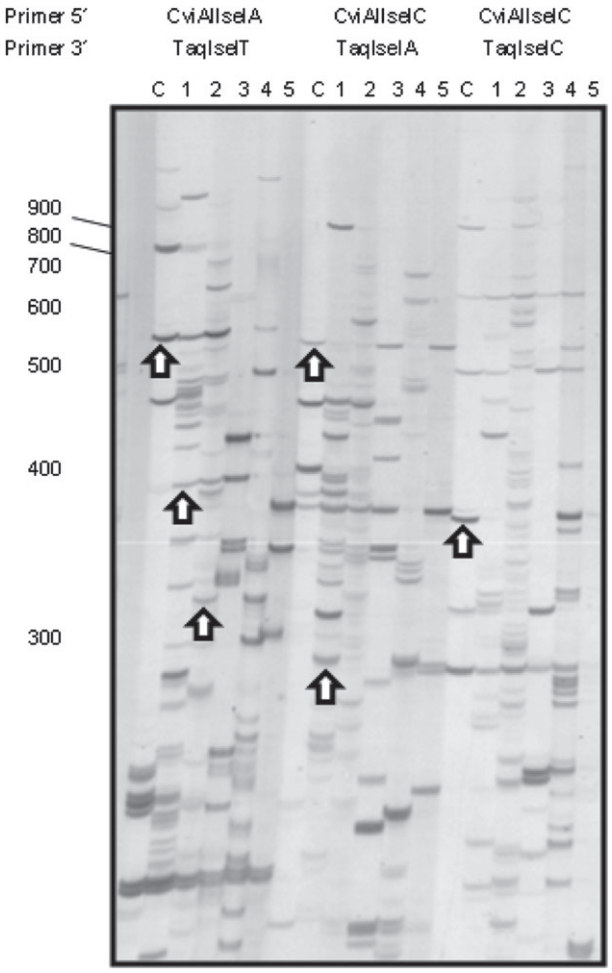

Figure 1 Example of the results obtained using the cDNA-AFLP analysis. TDFs resulting from selective amplification by using three combinations of primers (indicated in the top of the figure). The primers used in these reactions included the forward primers CviAIIselA and CviAIIselC and three different reverse primers: TaqIselT, TaqIselA and TaqIselC. Rows correspond to five crosses and control sample: control (C), 4*x* A1 × 4*x* A2 (1), 4*x* S1 × 4*x* A2 (2), 4*x* A1 × 2*x* S1 (3), 4*x* S1 × 2*x* S1 (4) and C4-4*x* × Q4084-8 (5). Arrows indicate examples of some of the DETDFs that were selected in the analysis.

Classification of DETDFs

The 100 DETDFs were sorted into 12 functional categories. The largest sets of genes corresponded to the functional groups involved in transcription (10%), signal transduction, metabolism and cell structure (8% each). A minor number of DETDFs belonged to the protein destination and storage (7%), intracellular traffic (6%) and regulation of gene expression (4%). The remaining 12% of the sequences were classified as being involved in disease/defence, transporters, secondary metabolism and protein synthesis. Finally, 40% of the analysed sequences were included in the unknown category (Fig. 2).

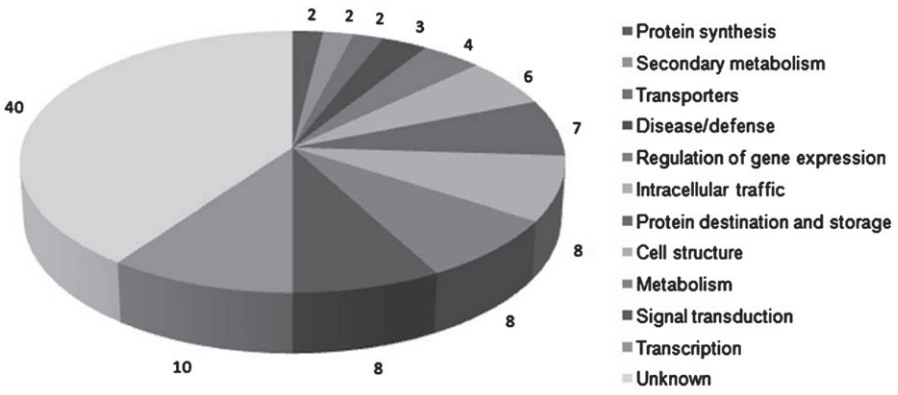

Figure 2 Functional categories assigned to 100 DETDFs (differentially expressed transcript-derived fragments) identified by cDNA-AFLP analysis. The transcripts were isolated and sequenced from ovaries obtained from different crosses of *Paspalum notatum* 3 h after pollination. The sequences of the best matching protein found in all searches were blasted against the Arabidopsis informative resource (TAIR) protein database using BLASTp. Best-hit proteins were submitted to the GO annotation tool at TAIR to search for putative molecular functions and related biological processes. The numbers indicate percentages of DETDFs grouped into each functional category.

Five DETDFs were expressed only when the predicted m : p ratio in the endosperm was 2:1 using a tetraploid sexual plant as a female parent (Table 3, Class A). This case could be considered as the normal situation and therefore, seeds could be expected to be formed. The sequences in this group (AC13, AC14, CC11, CC15 and TG1) were classified as being involved in cell structure, metabolism, intracellular traffic and unknown categories (Table S3). Also, 21 DETDFs were expressed only when the predicted m : p ratio in the endosperm was 4:1 using a tetraploid sexual plant as the female parent (Table 3, Class B). In this case, the m : p ratio was different from the 2:1 ratio necessary to fulfil the EBN hypothesis and no seeds were expected to be formed. Five sequences (AA5, AA7, AA9, AA14 and GG11) probably regulate transcription either directly or by modifying chromatin structure (Table S3).

Table 3 Classes of DETDFs[a]

			APO[b]			SEX[b]		
Expression Pattern	Seed Set	Class	4:1[c]	8:1[c]	8:3[c]	2:1[c]	4:1[c]	Total of DETDFs
Sexuals only (EBN-sensitive)	Yes	A				X		5
	No	B					X	21
	Yes/No	C				X	X	2
Apomicts only (EBN-insensitive)	Yes	D	X					25
	Yes	E		X				2
	Yes	F			X			1
	Yes	G	X	X				4
	Yes	H		X	X			14
Apomicts and sexuals	Yes	I	X			X		4
	Yes	J	X	X		X		1
	Yes	K			X	X		2
	Yes/No	L	X				X	2
	Yes/No	M		X			X	2
	Yes/Yes/No	N	X	X			X	3
	Yes/Yes/No	O		X	X		X	3
	Yes/Yes/No	P		X		X	X	2
	Yes/Yes/Yes/No	Q	X	X		X	X	5
Control only	No	R						2
Total								100

[a]*Paspalum notatum* DETDFs classification based on the comparative gene expression patterns among the various reproductive scenarios.
[b]Reproductive mode of the plant used as the female parent in the cross.
[c]Predicted m : p ratio in the endosperm.

Also, CC14 and GG1 belong to the protein destination and storage category with cysteine peptidase and ubiquitin protein ligase activities, respectively. AA6, a lectin-like receptor kinase, and CG3, a cysteine-rich receptor-like protein kinase, were included in the signal transduction category. Finally, GA5 presented identity with a transmembrane amino acid permease transporter and CA10 with a glutamine synthetase (Table S3). Finally, ten of these 21 TDFs were proteins with no assigned function (Table S3). Only two DETDFs were expressed when the predicted m : p ratios in the endosperm were 2:1 and 4:1 using a tetraploid sexual plant as the female parent (Table 3, Class C). One of these transcripts encodes a ribosomal protein and the second a predicted protein with unknown function (Table S3). Since these DETDFs could be inducing or repressing pathways that result in seed abortion in sexual *P. notatum* plants with m : p ratios different to the 2:1 relation found in most angiosperms, they are interesting candidates for functional analyses.

On the other hand, 46 transcripts were expressed in EBN-insensitive seed set when a tetraploid apomictic plant was used as a female parent and the predicted m : p ratio in the endosperm is different from 2:1 present in most angiosperms (Table 3, Classes D to H). Out of them, 25 DETDFs were expressed only when the predicted m : p ratio in the endosperm was 4:1 using a tetraploid apomictic as the female parent (Table 3, Class D). A considerably lower number of DETDFs (3) were expressed only when m : p predicted ratios were 8:1 or 8:3, using a tetraploid apomictic as the female parent (Table 3, Classes E and F). Since tetraploid apomictic plants are insensitive to the EBN, seeds are formed in the three as situations described above. In addition to that, four and 14 DETDFs were expressed in common when m : p ratios were 4:1/8:1 and 8:1/8:3, respectively (Table 3, Classes G and H). Of these 46 sequences, five transcripts (AC2, CA6, CC6, GA3 and TA4) were probably involved in intracellular signal transduction, and another six (CA7, CA12, GA1, GG10, GG15 and GG16) were predicted to function in transcription and regulation of gene expression (Tables 4a and 4b). Also, three transcripts were associated with the organisation of the cytoskeleton (AC3, AG2 and CC2) and one (AC12) with the plant cell wall (Tables 4a and 4b). Finally, 16 gene products (AG6, AG7, CC4, CC10, CC16, CC17, GC1, GG6, GG9, GA4, AA3, GA8, AC10, GG13, GG17 and TA5) were possibly functioning in protein synthesis, protein destination and storage, intracellular traffic, disease/defence and participating in several metabolic pathways (Tables 4a and 4b). The remaining 15 sequences were classified as unknown function.

Moreover, seven DETDFs were expressed when seeds are formed regardless of the EBN and may represent the basal seed formation expression set. These transcripts were found in the following situations: four sequences in common between S2:1 and A4:1 (Table 3, Class I); one in S2:1, A4:1 and A8:1 (Table 3, Class J) and two sequences in common between and S2:1 and A8:3 (Table 3, Class K). They were classified as being part of transcription, cell structure, disease/defence, metabolism, signal transduction and unknown categories (Table S3). All these transcripts seem to participate in cellular processes that are essential for cell survival, such as methyltransferases that modify histones, proteins that are enzymes of key metabolic pathways or participate in signal transduction.

In order to estimate the percentages of these DETDFs that were expressed in the different parts of the developing seed, we performed an *in silico* comparison with EST collections available in public databases (ESTdb NCBI). Local BLASTn (blast-2.2.28+) searches (Zhang *et al.*, 2000) were performed in order to compare the sequences. Twenty-six DETDFs presented identity with sequences expressed in barley, maize, rice and wheat endosperm. In contrast, 20 *P. notatum* sequences showed identity with barley, maize, rice and wheat embryo-specific sequences (Table S4). Hence, even though the 100 sequences analysed in this work are expressed in whole developing seeds, and at this stage only the binucleated central cell could be observed, we can assume that at least 26% of them are transcripts that could be associated with the development of the endosperm in *P. notatum.*

Endosperm and embryo development (DIC microscopy)

The development of the endosperm and the embryo were studied in order to determine if both happen in a coordinated manner when containing different ploidy levels. The crosses analysed were: $4x$ A1 × $2x$ S1, $4x$ S2 × $2x$ S2 and $4x$ S1 × $2x$ S2, 48 h after pollination (Tables 1 and 2). A minimum of 20 ovules from at least two different inflorescences were observed using a microscope equipped with a DIC system. It was determined that 48 h after pollination, the endosperm from the sexual $4x \times 2x$ cross was small, occupying a small portion of the ovule, and has stopped growing (Fig. 3). In contrast, the endosperm from the apomictic $4x \times 2x$ cross occupied most of the ovule and continued growing (Fig. 3).

Expression analysis

The expression patterns of some DETDFs were analysed by real-time PCR. The selected sequences were AC10, AC5, GG17, TC2 and CG1. Since transcripts expressed in the EBN-insensitive seed set are the most interesting transcripts for further characterisation, four out of the five (AC10, AC5, GG17 and CG1) genes analysed

Table 4a Transcripts expressed in EBN-insensitive seed set (predicted m : p ratio in the endosperm 4:1) and sequence search results of clones with significant homologies

Class	DETDF	Direct Protein by BLASTx ID; name (*e*-value)	TIGR CDS with BLASTn ID; name (*e*-value)	EST with BLASTn ID; name (*e*-value)	Protein by BLASTx of the EST Best Match ID; name (*e*-value)	TAIR BLASTp[a] ID (GO function)	Processes
D	AA12	n	n	n	–	–	Unknown
D	AC7	XP_003638451; hypothetical protein [*Medicago truncatula*] ($3e^{-14}$)	LOC_Os09g01000; expressed protein ($8e^{-109}$)	DV470205; *Brachypodium distachyon* callus EST library ($3e^{-48}$)	XP_002450733; hypothetical protein SORBIDRAFT_05g016477 [*Sorghum bicolor*] ($3e^{-47}$)	n	Unknown
D	AC12	ACA23876; Pas n1 allergen precursor [*Paspalum notatum*] ($5e^{-08}$)	LOC_Os10g40090; expansin precursor ($1e^{-66}$)	FL847655; *Panicum virgatum* late flowering buds + seed development ($3e^{-23}$)	XP_004983734; PREDICTED: expansin-B9-like [*Setaria italica*] ($5e^{-101}$)	AT1G65680, *EXPB2* (Beta-expansin)	Cell structure (cell wall)
D	AG5	n	n	n	–	n	Unknown
D	AG6	n	LOC_Os03g01960; expressed protein ($2e^{-29}$)	FL770192; *Panicum virgatum* early floral buds + reproductive tissue ($4e^{-23}$)	XP_004982817; PREDICTED: uncharacterised protein LOC101776194 [*Setaria italica*] ($1e^{-31}$)	AT1G65720 (unknown function)	Intracellular traffic (vacuole)
D	AG7	n	LOC_Os03g01960; expressed protein ($1e^{-28}$)	FL925235; *Panicum virgatum* root (L) cDNA clone ($7e^{-14}$)	XP_004982817; PREDICTED: uncharacterised protein LOC101776194 [*Setaria italica*] ($6e^{-15}$)	AT1G65720 (unknown function)	Intracellular traffic (vacuole)
D	AG8	n	n	n	–	–	Unknown
D	CA4	n	n	n	–	–	Unknown
D	CA6	XP_002447513; hypothetical protein SORBIDRAFT_06g002515 [*Sorghum bicolor*] ($4e^{-11}$)	LOC_Os02g18370; transposon protein ($9e^{-27}$)	CN127783; Acid- and alkaline-treated roots *Sorghum bicolor* cDNA clone ($1e^{-34}$)	XP_002447513; hypothetical protein SORBIDRAFT_06g002515 [*Sorghum bicolor*] ($3e^{-123}$)	AT4G38180, *FRS5* (zinc ion binding)	Signal transduction
D	CA7	XP_002451633; hypothetical protein SORBIDRAFT_04g004940 [*Sorghum bicolor*] ($2e^{-12}$)	LOC_Os02g07780; OsSPL4-SBP-box gene family member ($2e^{-37}$)	FL927913; *Panicum virgatum* root cDNA clone ($2e^{-34}$)	XP_004951747; PREDICTED: squamosa promoter-binding-like protein 11-like [*Setaria italica*] ($3e^{-139}$)	AT1G27370, *SPL10* (DNA binding)	Transcription
D	CA12	ADB92637; putative histone acetyltransferase ELP3 [*Hordeum vulgare*] ($3e^{-06}$)	LOC_Os04g40840; elongator complex protein 3 ($8e^{-35}$)	EB661798; pAPO *Cenchrus ciliaris* cDNA 3′, mRNA sequence ($3e^{-35}$)	ACN34993; unknown [*Zea mays*] ($9e^{-38}$)	AT5G50320, *ELP3* (histone acetyltransferase activity)	Transcription (chromatin modification)
D	CC2	AFW62846; hypothetical proteinZEAMMB73_521065 [*Zea mays*] ($2e^{-58}$)	LOC_Os02g38340; actin ($4e^{-53}$)	JG795098; *Panicum virgatum* AP13 normalised full root cDNA clone ($2e^{-108}$)	XP_004952989; PREDICTED: actin-related protein 3-like isoform X2 [*Setaria italica*] ($3e^{-69}$)	AT1G13180, *ARP3* (ATP binding, actin binding)	Cell structure (cytoskeleton)
D	CC3	XP_002516640; conserved hypothetical protein [*Ricinus communis*] ($2e^{-03}$)	n	FG612270; stem cDNA library *Papaver somniferum* ($2e^{-33}$)	n	n	Unknown
D	CC4	NP_001130331; uncharacterised protein LOC100191426 [*Zea mays*] ($7e^{-27}$)	LOC_Os01g35040; ZOS1-09-C2H2 zinc finger protein ($4e^{-60}$)	DV639223; Sugarcane root library *Saccharum* hybrid cultivar ($2e^{-52}$)	XP_002455751; hypothetical protein SORBIDRAFT_03g023980 [*Sorghum bicolor*] ($4e^{-94}$)	AT4G15420 (zinc ion binding)	Protein destination and storage (proteolysis)

Table 4a Continued

Class	DETDF	Direct Protein by BLASTx ID; name (*e*-value)	TIGR CDS with BLASTn ID; name (*e*-value)	EST with BLASTn ID; name (*e*-value)	Protein by BLASTx of the EST Best Match ID; name (*e*-value)	TAIR BLASTp[a] ID (GO function)	Processes
D	CC6	EMS48658; Serine/threonine-protein kinase tricorner [*Triticum urartu*] ($1e^{-05}$)	LOC_Os10g33640; ACG kinases include homologues to PKA, PKG and PKC ($2e^{-89}$)	EB641411; *Zea mays* cDNA ($5e^{-08}$)	DAA49648; TPA: putative AGC protein kinase family protein [*Zea mays*] ($6e^{-131}$)	AT2G20470 (ATP binding)	Signal transduction (kinases)
D	CC7	n	n	n	–	–	Unknown
D	CC10	XP_004978559; PREDICTED: uncharacterised protein LOC101767807 [*Setaria italica*] ($2e^{-29}$)	LOC_Os12g01930; lipase-related ($2e^{-87}$)	FL859340; *Panicum virgatum* etiolated seedlings cDNA clone ($2e^{-60}$)	XP_002442668; hypothetical protein SORBIDRAFT_08g000930 [*Sorghum bicolor*] ($8e^{-123}$)	AT2G44970 (hydrolase activity)	Protein destination and storage
D	CC16	NP_001130331; uncharacterised protein LOC100191426 [*Zea mays*] ($9e^{-24}$)	LOC_Os01g35040; ZOS1-09-C2H2 zinc finger protein ($6e^{-99}$)	CA096774; *Saccharum* hybrid cultivar SP80-3280 cDNA clone ($4e^{-54}$)	XP_002455751; hypothetical protein SORBIDRAFT_03g023980 [*Sorghum bicolor*] ($4e^{-146}$)	AT4G15420 (zinc ion binding)	Protein destination and storage (proteolysis)
D	CC17	XP_006654859; bidirectional sugar transporter SWEET5-like [*Oryza brachyantha*] ($2e^{-12}$)	LOC_Os05g51090; nodulin MtN3 family protein ($1e^{-27}$)	FL834082; *Panicum virgatum* late flowering buds + seed development ($3e^{-30}$)	XP_004961041; PREDICTED: bidirectional sugar transporter SWEET5-like [*Setaria italica*] ($7e^{-30}$)	AT3G28007, *SWEET4* (sugar transmembrane transporter activity)	Transporters (sugars)
D	GA1	XP_004984387; probable nucleolar protein 5-2-like [*Setaria italica*] ($1e^{-22}$)	LOC_Os03g22730; nucleolar protein NOP5-1 ($3e^{-106}$)	CA281262; *Saccharum* hybrid cultivar (mixed) cDNA clone (1e-55)	XP_002456868; hypothetical protein SORBIDRAFT_03g044260 [*Sorghum bicolor*] (2e-131)	AT3G05060 (DNA binding)	Transcription
D	GA6	n	n	JG772215; full-length cDNA library of inflorescence of *Eulaliopsis binata* ($5e^{-13}$)	n	n	Unknown
D	GC1	n	LOC_Os06g13680; B12D protein ($9e^{-12}$)	JZ199471; SSH cDNA library *Saccharum* hybrid cultivar Co 93009 cDNA clone ($2e^{-13}$)	XP_004965125; PREDICTED: uncharacterised protein LOC101762161 [*Setaria italica*] ($3e^{-18}$)	AT3G48140 (unknown function)	Metabolism (mitochondrion)
D	GG6	n	LOC_Os07g48244; ubiquinol-cytochrome c reductase complex 6.7 kDa protein ($1e^{-14}$)	CN140965; Oxidatively stressed leaves and roots *Sorghum bicolor* cDNA clone ($5e^{-34}$)	XP_002463395; hypothetical protein SORBIDRAFT_02g043020 [*Sorghum bicolor*] ($2e^{-14}$)	AT2G40765 (unknown function)	Metabolism (mitochondrion)
D	GG9	XP_002461681; hypothetical protein SORBIDRAFT_02g006440 [*Sorghum bicolor*] ($4e^{-22}$)	LOC_Os07g11120; hydrolase, NUDIX family ($1e^{-44}$)	CA087922; *Saccharum* hybrid cultivar SP80-3280 cDNA clone ($1e^{-48}$)	XP_002461681; hypothetical protein SORBIDRAFT_02g006440 [*Sorghum bicolor*] ($2e^{-63}$)	AT3G12600, *NUDT16* (hydrolase activity)	Disease/defence (stress responses)
D	GG10	n	LOC_Os07g43980; DEAD-box ATP-dependent RNA helicase ($6e^{-36}$)	CA219615; *Saccharum* hybrid cultivar (mixed) cDNA clone ($3e^{-03}$)	XP_002463240; hypothetical protein SORBIDRAFT_02g040380 [*Sorghum bicolor*] ($2e^{-39}$)	AT1G16280, *RH36* (ATP binding)	Transcription (embryo sac development)

Table 4b Transcripts expressed in EBN-insensitive seed set (predicted m : p ratio in the endosperm 8:1, 8:3 or a combination of these ratios with 4:1) and sequence search results of clones with significant homologies

Class	DETDF	Direct protein by BLASTx ID; name (*e*-value)	TIGR CDS with BLASTn ID; name (*e*-value)	EST with BLASTn ID; name (*e*-value)	Protein by BLASTx of the EST Best Match ID; name (*e*-value)	TAIR BLASTp[a] ID (GO function)	Processes
E	CG2	n	n	CA245133; *Saccharum* hybrid cultivar (mixed) cDNA clone ($7e^{-07}$)	n	n	Unknown
E	GA4	AFW73308; hypothetical proteinZEAMMB73_115168 [*Zea mays*] ($4e^{-25}$)	LOC_Os02g50340; membrane attack complex component/perforin/complement C9 ($1e^{-166}$)	CT845849; *Oryza sativa* Indica Group cDNA clone ($3e^{-37}$)	EAY87441; hypothetical protein Osl_08850 [*Oryza sativa* Indica Group] ($1e^{-161}$)	AT1G14780 (unknown function)	Disease/defence
F	GC2	n	LOC_Os05g29920; expressed protein ($6e^{-56}$)	CA293091; *Saccharum* hybrid cultivar (mixed) cDNA clone ($5e^{-15}$)	XP_002439653; hypothetical protein SORBIDRAFT_09g017990 [*Sorghum bicolor*] ($3e^{-104}$)	AT4G02370 (unknown function)	Unknown (chloroplast, vacuole)
G	AA3	XP_004961679; 60S ribosomal protein L30-like isoform X2 [*Setaria italica*] ($6e^{-62}$)	LOC_Os05g41110; ribosomal protein L7Ae ($3e^{-58}$)	CN611628; Sugarcane mature stem library *Saccharum* hybrid cultivar CoS 767 cDNA clone ($1e^{-123}$)	NP_001278524; LOC100285769 [*Zea mays*] ($9e^{-66}$)	AT1G77940 (structural constituent of ribosome)	Protein synthesis (ribosomal proteins)
G	CC5	n	LOC_Os01g33670; expressed protein ($9e^{-10}$)	FL842635; *Panicum virgatum* late flowering buds + seed development ($2e^{-14}$)	XP_004971878; PREDICTED: uncharacterised protein LOC101764772 [*Setaria italica*] ($6e^{-36}$)	n	Unknown
G	GA3	n	LOC_Os01g57940; tyrosine protein kinase domain containing protein ($2e^{-37}$)	EB165638; *Zea mays* cDNA ($7e^{-14}$)	NP_001136610; putative protein kinase superfamily protein [*Zea mays*] ($1e^{-96}$)	AT3G20530 (ATP binding)	Signal transduction (kinases)
G	GA8	n	LOC_Os03g38020; mps one binder kinase activator-like 1A ($6e^{-66}$)	GD013686; abiotic and biotic stress treatments *Panicum virgatum* cDNA clone ($9e^{-33}$)	XP_004972757; PREDICTED: MOB kinase activator-like 1-like [*Setaria italica*] ($1e^{-66}$)	AT5G45550, *MOB1-LIKE* (unknown function)	Intracellular traffic (vesicular)
H	AC2	DAA42683; TPA: CK2 protein kinase alpha 1 [*Zea mays*] ($3e^{-20}$)	LOC_Os03g55389; casein kinase II subunit alpha-1 ($9e^{-57}$)	EE189502; *Zea mays* cDNA clone ($9e^{-49}$)	AFW67542; putative casein kinase II alpha subunit family protein [*Zea mays*] ($9e^{-80}$)	AT2G23080, *CKA3* (ATP binding)	Signal transduction (calcium-mediated signalling)
H	AC3	n	LOC_Os03g51600; tubulin/FtsZ domain containing protein ($2e^{-26}$)	CA295462; *Saccharum* hybrid cultivar (mixed) cDNA clone ($5e^{-26}$)	KCW87067; hypothetical protein EUGRSUZ_B03604 [*Eucalyptus grandis*] ($1e^{-17}$)	AT5G19780, *TUA5* (GTP binding)	Cell structure (cytoskeleton)
H	AC5	EPS74717; hypothetical protein M569_00042 [*Genlisea aurea*] ($2e^{-22}$)	LOC_Os01g57960; retrotransposon protein ($3e^{-05}$)	JZ199264; subtracted cDNA library *Saccharum* hybrid cultivar Co 93009 cDNA clone ($1e^{-71}$)	EPS74717; hypothetical protein M569_00042 [*Genlisea aurea*] ($1e^{-29}$)	n	Unknown
H	AC8	XP_003627732; ATP synthase subunit beta [*Medicago truncatula*] ($4e^{-29}$)	n	HX821495; *Brachypodium distachyon* cDNA clone ($1e^{-95}$)	AGV54793; hypothetical protein [*Phaseolus vulgaris*] ($3e^{-39}$)	n	Unknown
H	AC10	NP_001182963; uncharacterised protein LOC100501272 [*Zea mays*] ($7e^{-46}$)	LOC_Os07g42960; phospho-2-dehydro-3-deoxyheptonate aldolase ($3e^{-66}$)	FL942121; *Panicum virgatum* root cDNA clone ($2e^{-113}$)	DAA41496; TPA: hypothetical protein ZEAMMB73_920697 [*Zea mays*] ($1e^{-84}$)	AT1G22410 (3-deoxy-7-phosphoheptulonate synthase activity)	Metabolism (aromatic amino acids)

Table 4b continued

Class	DETDF	Direct protein by BLASTx ID; name (*e*-value)	TIGR CDS with BLASTn ID; name (*e*-value)	EST with BLASTn ID; name (*e*-value)	Protein by BLASTx of the EST Best Match ID; name (*e*-value)	TAIR BLASTp[a] ID (GO function)	Processes
H	AG2	n	LOC_Os01g59130; proline-rich family protein ($2e^{-111}$)	JG948277; *Panicum virgatum* cv. Alamo-AP13 cDNA clone ($4e^{-30}$)	XP_002458653; hypothetical protein SORBIDRAFT_03g037470 [*Sorghum bicolor*] ($2e^{-142}$)	AT2G26770, *SCAB1* (actin binding)	Cell structure (cytoskeleton)
H	CG1	ACG25350; hypothetical protein [*Zea mays*] ($4e^{-60}$)	LOC_Os02g55200; expressed protein ($4e^{-109}$)	CA198293; *Saccharum* hybrid cultivar (mixed) cDNA clone ($2e^{-125}$)	XP_002444897; hypothetical protein SORBIDRAFT_07g001060 [*Sorghum bicolor*] ($2e^{-150}$)	AT3G54360, *NCA1* (zinc ion binding)	Unknown
H	GG13	ACF84459; unknown [*Zea mays*] ($1e^{-11}$)	LOC_Os05g06430; mOsPDIL2-1 protein disulphide isomerase PDIL2-1 ($2e^{-25}$)	JK603074; floral spike *Setaria italica* cDNA clone ($4e^{-30}$)	XP_004977081; PREDICTED: protein disulphide isomerase-like 2-1-like [*Setaria italica*] ($5e^{-28}$)	AT2G47470, *PDIL2* (protein disulphide isomerase activity)	Protein destination and storage (modification)
H	GG14	XP_003614396; hypothetical protein MTR_5g051150 [*Medicago truncatula*] ($5e^{-19}$)	n	HX819322; *Brachypodium distachyon* cDNA clone ($2e^{-84}$)	EXC10888; hypothetical protein L484_002552 [*Morus notabilis*] ($6e^{-73}$)	n	Unknown
H	GG15	EEC78946; hypothetical protein OsI_19395 [*Oryza sativa* Indica Group] ($2e^{-06}$)	LOC_Os04g36810; LSM domain containing protein ($8e^{-10}$)	GW883531; drought stress library *Phleum pratense* cDNA ($4e^{-31}$)	BAK05591; predicted protein [*Hordeum vulgare* subsp. vulgare] ($7e^{-11}$)	AT5G53060, *RCF3* (RNA binding)	Regulation of gene expression
H	GG16	ABX71498; glyceraldehyde-3-phosphate dehydrogenase [*Secale cereale*] ($8e^{-04}$)	LOC_Os04g36810; LSM domain containing protein ($8e^{-10}$)	GW883531; drought stress library *Phleum pratense* cDNA ($3e^{-16}$)	BAK05591; predicted protein [*Hordeum vulgare* subsp. vulgare] ($7e^{-11}$)	AT5G53060, *RCF3* (RNA binding)	Regulation of gene expression
H	GG17	XP_004975032; protein transport protein Sec24-like At3g07100-like [*Setaria italica*] ($1e^{-71}$)	LOC_Os04g04020; protein transport protein Sec24-like ($1e^{-136}$)	JK553215; whole seedlings *Setaria italica* cDNA clone ($9e^{-137}$)	XP_004975032; PREDICTED: protein transport protein Sec24-like At3g07100-like [*Setaria italica*] ($6e^{-169}$)	AT3G07100, *SEC24A* (zinc ion binding)	Intracellular traffic (vesicular)
H	TA4	n	LOC_Os06g01850; ferredoxin-NADP reductase ($3e^{-23}$)	GW883549; drought stress library *Phleum pratense* cDNA ($2e^{-57}$)	CAD30024; ferredoxin-NADP(H) oxidoreductase [*Triticum aestivum*] ($3e^{-29}$)	AT1G20020; *LFNR2* (NADP binding)	Signal transduction (chloroplast)
H	TA5	XP_002466978; hypothetical protein SORBIDRAFT_01g017790 [*Sorghum bicolor*] ($3e^{-09}$)	LOC_Os10g35960; NAD-dependent malic enzyme, mitochondrial precursor ($7e^{-73}$)	CA257130; *Saccharum* hybrid cultivar (mixed) cDNA clone ($6e^{-32}$)	ACR34641; unknown [*Zea mays*] ($5e^{-91}$)	AT4G00570, *NAD-ME2* (ATP binding)	Metabolism (mitochondrion)

[a]BLASTp on TAIR database of the best hit.

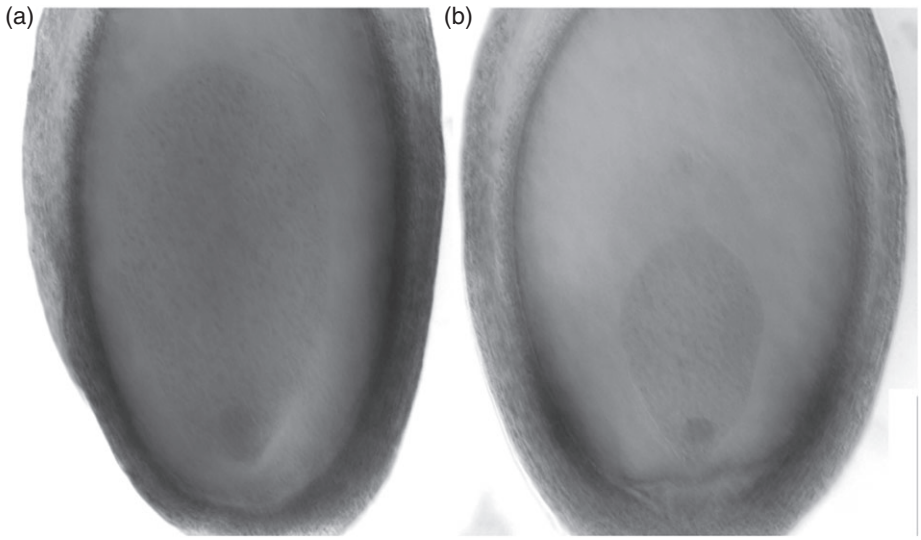

Figure 3 Embryo and endosperm observed in bahiagrass 48 h after pollination. (a) Ovaries from a 4*x* A1 plant (Q4117) pollinated by a 2*x* S2 plant (Tifton 9). (b) Ovaries from a 4*x* S2 plant that has been pollinated by a 2*x* S2 plant (Tifton 9). Magnification: ×200.

by qPCR belonged to this group of DETDFs (Table 3, Class H). Additionally, TC2 was one of the transcripts associated with EBN-sensitive (no seed set) fertilisation scenario (Table 3, Class B), the largest group containing 21 DETDFs. Transcript AC10 was putatively involved in amino acid metabolism, whereas AC5 was classified as hypothetical protein and GG17 in intracellular protein transport (Table 4b). The other two TDFs (CG1 and TC2) presented identity with sequences with no assigned function (Tables 4b and S4). The expression profiles were confirmed for TDFs AC10, AC5, GG17 and CG1. These transcripts were identified by cDNA-AFLP analysis as being expressed in crosses that were expected to form seeds (Table 3, Class H). Indeed, relative expression values observed for these TDFs were significantly higher in crosses 4*x* A1 × 2*x* S1 and 4*x* A1 × 4*x* A2 compared to 4*x* S1 × 2*x* S1 (Fig. 4). However, the expression pattern observed by cDNA-AFLP could not be confirmed by real-time PCR for TC2 (Fig. 4). TC2 was originally isolated from crosses that did not form seeds (Table 3, Class B).

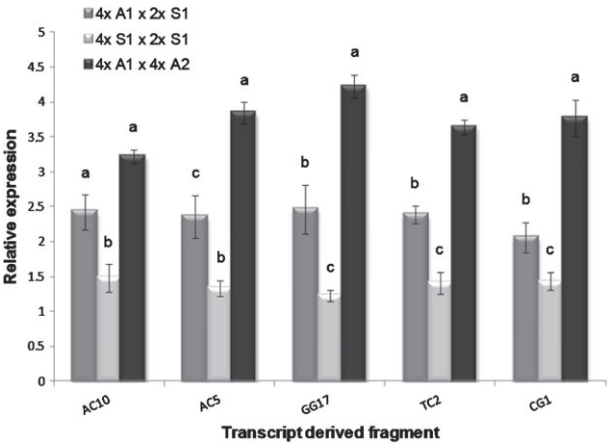

Figure 4 Real-time PCR showing the expression levels for selected TDFs that are indicated below the axis. Each TDF was analysed separately. Columns show relative expression values (normalised expression relative to the lowest normalised value for each gene) and bars indicate errors. Columns with the same letter are not significantly different. Crosses are indicated in the figure. The experiment was performed on two biological replicates (different RNA extractions from two experimental crosses), using six technical replicates.

Discussion

Seeds are the major resource for human nutrition and animal feed throughout the world. They are also raw material for industry and the production of alternative energy. The understanding of seed formation is essential to improve grain properties with respect to yield, nutritional value or industrial usage (Guillon *et al.*, 2012). Despite its wide distribution in the flowering plants, apomixis is present only in some species of agronomic interest, which include a group of forage grasses, cassava, apple, citrus, mango and strawberries (Dwivedi *et al.*, 2010). Transferring apomixis to major crops worldwide would produce a huge impact on agriculture since it would enable the establishment of hybrid combinations, and their maintenance and propagation by seed (Vielle-Calzada *et al.*, 1996; Dwivedi *et al.*, 2010). The prospect of cloning genotypes has great potential due to increasing commercialization of hybrids in different crops and the enormous cost that represents the

annual generation of hybrid seed. This applies to both, cross-pollinated crops such as corn and sunflower and also for autogamous crops such as rice and cotton. It would also be an important aid for farmers in developing countries, enabling them to sustain high yields year after year using part of their own harvested seeds.

In this work we identified DETDFs associated with the formation of seeds in *P. notatum*. A transcriptome analysis of ovaries (3 h after pollination) allowed the characterisation of 100 DETDFs that showed differential expression patterns during the initial phase of seed development. The spatial and temporal regulation of seed maturation requires the concerted action of several signalling pathways integrating information from genetic programmes, from hormonal signals, as well as from metabolic signals (Weber *et al.*, 2010). During the initial pre-storage phase, which is characterised by maternal control, organogenesis and morphogenesis occur on the basis of high cell division activity (Weber *et al.*, 2010). Whereas the embryo quickly organises meristems, the endosperm is essentially without meristems and matures into a short-lived, terminal structure specialised for food storage (Brown & Lemmon, 2007). Even when there could be differences in pollen tube growth rates depending on pollen ploidy, it could be assumed, based on the information available for other species of the same genus (Burson, 1987), that double fertilisation occurs before 3 h of pollination in *P. notatum*. Furthermore, there were no visible early endosperm formation differences between sexual and apomictic fertilisation programmes 3 h after pollination. Only the binucleated central cell and the egg cell could be observed at this stage in the pollinated ovaries (Acuña *et al.*, 2009).

An important percentage of the transcripts identified in this work belonging to the metabolism and cell structure functional categories, were predicted to function in amino acid metabolism and transport, cell wall and cytoskeleton organisation (Fig. 2). These three processes play fundamental roles during the active cell division that occurs in the initial phase of seed development. Especially, the cytoskeleton is a key factor for the unique developmental pathway of cereal endosperm. It includes a specific programme of changes in the cell and microtubule cycles leading to an assemblage of plant tissues specialised for uptake of metabolites (transfer cells), storage of food reserves (starchy endosperm), and enzyme synthesis (aleurone) (Brown & Lemmon, 2007). It could be hypothesised that the transcripts detected at this stage in *P. notatum* pollinated ovaries are accumulating to prepare the central cell for the active cell division that occurs at the beginning of endosperm formation. It is worth mentioning that by using local tBLASTx search (blast-2.2.28+) only one of the 100 DETDFs presented identity with genes known to be expressed in the central cell (Liu *et al.*, 2010). CC6 transcript presented low identity (*e* value 0.008) with AT4G24974, a self-incompatibility protein-like (Liu *et al.*, 2010). This result could be consequence of the cDNA-AFLP technique, in which the choice of the restriction enzymes used to digest the cDNAs is critical. Nevertheless, even using the best combination of restriction enzymes, a percentage of the transcripts present in the samples will be missed in the cDNA-AFLP analysis.

The endosperm has a major influence on seed mass, and because it makes up most of the cereal seed, genes that affect endosperm growth have direct effects on seed size (Ohto *et al.*, 2007). A model states that the final seed size is determined by the product of the seed growth rate (SGR) and the seed fill duration (Egli, 1998). In addition to that, SGR is influenced by a genetic component, regulated by the seed itself through the number of cells in the endosperm (or the cotyledons), and environmental components, such as temperature and water availability (Egli, 1998). The termination of seed growth when assimilate is still available requires a regulatory mechanism in the seed (Egli, 2004). However, more research is needed to find out the reasons why seeds stop growing and the factors determining the seed number.

Observations of *P. notatum* ovaries 48 h after pollination indicated that the volume of the cellularized endosperm is considerably bigger in ovaries of the apomictic plant compared with the sexual plant (Fig. 3) when male parents are 2x plants. Genes expressed differentially in these tissues might be interesting candidates to be further investigated in relation to the number of cells in the endosperm and, therefore, to the final seed size. Particularly in this species, the reproductive mode adds another level of complexity to the process. The endosperm of a sexual plant has a maternal genomic contribution in concordance with its ploidy level, while an apomictic plant provides two unreduced polar nuclei for endosperm formation (Quarin, 1999). Therefore, the maternal contributions are different. In fact, several genetic mechanisms have been proposed to explain the parent-of-origin effects observed during seed development, including the disproportionate maternal contribution to the endosperm, plastidic and cytoplasmic inheritance, expression of genes in the gametophytes and gametes, and differential expression of parental alleles in the developing seed (Dilkes & Comai, 2004). Dilkes & Comai (2004) introduced a broader model to reinterpret parent-of-origin effects based on dosage. This model is still compatible with the EBN hypothesis and states that differentially contributed dosage-sensitive components interact to produce a viable endosperm.

Crossing plants with different ploidy levels often alters seed development, generating reciprocal phenotypes

depending on the direction of the cross. In general, an increased ratio of paternally to maternally contributed genomes in the seed ('paternal excess') is associated with increased growth of endosperm, while an increased ratio of maternal to paternal genomes ('maternal excess') inhibits endosperm growth (Haig & Westoby, 1991; Gehring *et al.*, 2004). A widely accepted interpretation of interploidy cross phenotypes is that they disrupt the balance in the seed of active copies of parentally imprinted genes, which, depending on the particular gene, are expressed from only the maternal or only the paternal alleles (Tiwari *et al.*, 2010). Fifty-eight of the 100 transcripts identified in this work were expressed in crosses that set seeds in *P. notatum* (Table 3). Moreover, 46 of the 58 sequences were specifically found in crosses in which apomictic plants were used as the female parent and presented a predicted m : p ratio in the endosperm that was different to the 2:1 requirement of the EBN (Table 3, Classes A, D-K; Tables 4a and 4b). Tiwari *et al.* (2010) studied the transcriptional profiles underlying parent-of-origin effects in seeds of *A. thaliana* using microarrays. These authors stated that over-expression of genes observed in their experiments is much more likely to reflect genuine changes in developmental programmes than simply the presence of extra copies of genes in seeds with increased ploidy. Only two of the 46 DETDFs of *P. notatum*, CC4 and CC16 (both similar to a PRLI-interacting factor K participating in ubiquitin-dependent protein catabolism, AT4G15420; Table 3, Class D; Table 4a) were found in common with the list of up-regulated genes in crosses with maternal excess (Tiwari *et al.*, 2010). However, it is worth highlighting that these 46 transcripts were isolated from crosses of *P. notatum* plants with maternal excess in the seeds, and therefore they should have shown inhibited endosperm growth. Strikingly, these crosses were EBN-insensitive, presented endosperm cellularization (Figure 3A) and produced seeds. It is possible that our results are also a consequence of altered expression programmes. Nonetheless, since a 'paternalized seed' phenotype was observed in crosses with maternal excess in *P. notatum*, there should be particular alterations in gene expression that are specifically taking place in apomictic plants.

One of these 46 transcripts (GG13) presented 84% identity with a rice protein disulphide isomerase-like PDIL2-1 (Table 4b). Protein disulphide isomerases (PDI) play important roles in the maturation of secreted or plasma membrane proteins. Considering that an extracellular PDI was shown to be important for gamete fusion in mammals (Ellerman *et al.*, 2006), Wang *et al.* (2008) observed that several *PDI-like* genes were represented in EST libraries from egg, sperm and embryo sac in maize. These authors studied Arabidopsis homologues of maize PDIL proteins in order to determine if they played any role in plant reproduction using T-DNA insertion mutants. PDIL2-1 is a functional PDI that is localised in the ER and is highly expressed in the micropylar region of the ovule (Wang *et al.*, 2008). They demonstrated that certain T-DNA insertions in *Arabidopsis thaliana PDIL2-1* (AT2G47470) have reduced seed set, due to delays in embryo sac maturation. Moreover, results suggest that these truncated versions of PDIL2-1 function in sporophytic tissues to affect ovule structure and impede embryo sac development, thereby disrupting pollen tube guidance. This transcript was found in pollinated ovaries of *P. notatum* that completed the development of seeds.

Since the endosperm undergoes a brief and highly specialised pattern of development involving cell and microtubule cycles, and specific ways of wall placement, the four transcripts (AC3, AG2, CC2 and AC12, Tables 4a and 4b) predicted to be related to these processes could be interesting candidates for functional analyses. Also, some of the transcripts predicted to be involved in intracellular signal transduction, transcription and regulation of gene expression could be of interest for further studies. Particularly, CA7 presented identity with *SPL10* (AT1G27370), which encodes a specific transcription factor involved in several processes such as ovule development, response to brassinosteroid stimulus and vegetative to reproductive phase transition of meristem (Cardon *et al.*, 1999; Guo *et al.*, 2008) (Table 4a). Additionally, GG10 was similar to *RH36* (AT1G16280), a gene encoding a RNA helicase, essential for female gametogenesis (Huang *et al.*, 2010) (Table 4a).

DETDFs of *P. notatum* were compared to the differentially expressed transcripts reported in two other aposporic species, *P. simplex* (Polegri *et al.*, 2010) and *Poa pratensis* (Albertini *et al.*, 2004). Local tBLASTx (blast-2.2.28+) searches (Altschul *et al.*, 1997) were conducted against 14 sequences that were expressed after anthesis and/or in early stages of embryo and endosperm development in apomictic and sexual *P. simplex* (Polegri *et al.*, 2010). Three hits were obtained, with AA7 and CG5 showing identity with *TLP1* (AT1G76900, *e* value $1e^{-15}$) and *ATPK2* (AT3G08720, *e* value 0.01), respectively. These Arabidopsis genes are similar to two *P. simplex* ESTs, constitutively expressed in apomicts only and in all developmental stages analysed by Polegri *et al.* (2010). Also, CA7 presented identity (*e* value 0.001) with *RAPTOR1* (AT3G08850), an Arabidopsis gene similar to an EST expressed in a stage-specific manner (early stages of embryo and endosperm development and of seed formation) in apomictic *P. simplex* (Polegri *et al.*, 2010). Only CA7 was found in apomictic *P. notatum* ovaries (Table 3, Class D). However, AA7 and CG5 were found

in ovaries from *P. notatum* sexual plants (Table 3, Classes B and C, respectively). On the other hand, local BLASTx (blast-2.2.28+) searches against 32 protein sequences reported by Albertini *et al.* (2004) to be expressed in late stages of the developing reproductive organs of apomictic and sexual genotypes of *P. pratensis* rendered no hits.

Additionally, Chivasa *et al.* (2011) identified 40 spots in 2 D gels (representing 26 unique proteins) that were differentially expressed in response to changes in the levels of eATP in Arabidopsis cell suspension cultures. Twelve of the 100 DETDFs identified in our work presented identity (by local tBLASTx search using blast-2.2.28+) with some of those proteins, including peptides classified as glycolytic enzymes, ATP synthesis machinery, amino acid metabolism, protein degradation and molecular chaperones. Moreover, one of the TDFs (GG8) presented identity with the β-subunit of the mitochondrial ATP synthase. Chivasa *et al.* (2011) analysed mutant plants for this gene and postulated that in addition to its function in mitochondrial oxidative phosphorylation, ATP synthase β-subunit is a novel target for eATP in its function as a key negative regulator of plant cell death. Furthermore, programmed cell death (PCD) plays an important role in cereal endosperm development, with convincing evidence implicating hormones in the onset and progression of PCD in the endosperm (Sabelli & Larkins, 2009; Lombardi *et al.*, 2012). Moreover, GG11 was isolated only from pollinated ovaries of a sexual plant that did not set seed (m : p ratio 4:1) and was predicted to function as a transcription factor implicated in ethylene-mediated responses (AT2G33710; Table S3). Fertilisation determines the point after which the endosperm begins to develop and accumulate reserves while the nucellus becomes dispensable and dies by mean of PCD. Lombardi *et al.* (2012) demonstrated that the endosperm is indeed necessary for nucellus PCD to occur, initiating a cascade of events able to guarantee the accomplishment of the process. Since there are no anatomical connections between the two tissues, they demonstrated that ethylene, which is able to diffuse in both aqueous and lipid environments, is a regulator of nucellus PCD. It could be assumed that in pollinated ovaries of a sexual plant with a 4:1 m : p ratio, this signalling pathway is somehow deregulated, originating the earlier cell death of the endosperm in these plants. This could be one of the reasons why the endosperm in developing seeds of sexual plants that do not fit the EBN hypothesis is considerably smaller than in apomictic plants (Fig. 3).

In plants, cell communication via the extracellular matrix (apoplast) controls many processes vital for plant survival. Using this signalling route plant cells stimulate their neighbours by secreting signal molecules, such as peptides, proteins and primary or secondary metabolites. Secretion of ATP into the extracellular matrix is now recognised as a previously unknown stimulus for cell signalling with a role in many aspects of plant physiology (Tanaka *et al.*, 2010; Chivasa & Slabas, 2012). Recently, Choi *et al.* (2014) identified the first plant receptor for eATP. The level of extracellular ATP (eATP) induces several plant responses such as increased cytosolic calcium to changes in auxin transport, xenobiotic resistance, nitric oxide signalling, pollen germination and pollen tube growth, root growth and obstacle avoidance and response to microbes (Kim *et al.*, 2006; Reichler *et al.*, 2009; Chivasa & Slabas, 2012). Twelve of the DETDFs of *P. notatum* ovaries presented identity with proteins that were differentially expressed in response to changes in the levels of eATP in Arabidopsis cell suspension cultures (Chivasa *et al.*, 2011). Taking all this information into consideration, we suggest that signalling mediated by eATP could be involved in the regulation of endosperm development. This is a very interesting hypothesis but certainly needs more investigation.

In the past 20 years, significant progress has been made in defining regulatory genes that control aspects of seed development. A major challenge is to define the gene networks that operate during seed development (Ohto *et al.*, 2007). Initial progress towards this goal has been made with descriptions of the transcriptomes and proteomes of whole/parts of developing seeds (Lai *et al.*, 2004; Casson *et al.*, 2005; Grimanelli *et al.*, 2005; Day *et al.*, 2008; Ge *et al.*, 2008; Polegri *et al.*, 2010; Sharbel *et al.*, 2010; Guillon *et al.*, 2012; Pellny *et al.*, 2012; Thiel *et al.*, 2012; Schmidt *et al.*, 2014). Our group is currently investigating the transcriptome of developing seeds of *P. notatum* 24 and 48 h after pollination. Eventually, all this knowledge integrated with the physiological, biochemical and morphological description of seeds, will provide a comprehensive understanding of seed development.

Acknowledgements

This work was financed by the Secretaría de Ciencia, Tecnología e Innovación (SECTEI) de la provincia de Santa Fe grant N° 219109, the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) grant # PIP-11220090100613, the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) grant # PICT-2012-1321, and the Secretaría de Ciencia y Técnica, Universidad Nacional de Rosario (UNR) grant # AGR161. We thank Drs. Gabriela Breccia and Ana Ochogavía for their valuable advice on statistical analyses. Authors are research career members of the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). The authors declare that they have no conflict of interest.

References

Acuña C.A., Blount A.R., Quesenberry K.H., Kenworthy K.E., Hanna W.W. (2009) Bahiagrass tetraploid germplasm: reproductive and agronomic characterization of segregating progeny. *Crop Science*, **49**, 581–588.

Albertini E., Marconi G., Barcaccia G., Raggi L., Falcinelli M. (2004) Isolation of candidate genes in *Poa pratensis*. *Plant Molecular Biology*, **56**, 879–894.

Altschul S.F., Madden T.L., Schaffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

Berger F., Chaudhury A.M. (2009) Parental memories shape seeds. *Trends in Plant Science*, **14**, 550–556.

Brown R.C., Lemmon B.E. (2007) The developmental biology of cereal endosperm. In *Endosperm, Developmental and Molecular Biology*. Ed. O.-A. Olsen. Heidelberg, Germany: Springer.

Burson B.L. (1987) Pollen germination, pollen tube growth and fertilization following self and interspecific pollination of *Paspalum* species. *Euphytica*, **36**, 641–650.

Burton G.W. (1948) The method of reproduction in common Bahia Grass, *Paspalum notatum*. *Journal of American Society of Agronomy*, **40**, 443–452.

Burton G.W. (1955) Breeding Pensacola bahiagrass, *Paspalum notatum*: I. Method of reproduction. *Agronomy Journal*, **47**, 311–314.

Cardon G., Höhmann S., Klein J., Nettesheim K., Saedler H., Huijser P. (1999) Molecular characterisation of the *Arabidopsis* SBP-box genes. *Gene*, **237**, 91–104.

Casson S., Spencer M., Walker K., Lindsey K. (2005) Laser capture microdissection for the analysis of gene expression during embryogenesis of Arabidopsis. *The Plant Journal*, **42**, 111–123.

Chaudhury A.M., Koltunow A., Payne T., Luo M., Tucker M.R., Dennis E.S., Peacock W.J. (2001) Control of early seed development. *Annual Review of Cell and Developmental Biology*, **17**, 677–699.

Chivasa S., Slabas A.R. (2012) Plant extracellular ATP signalling: new insight from proteomics. *Molecular BioSystems*, **8**, 445–452.

Chivasa S., Tomé D.F.A., Hamilton J.M., Slabas A.R. (2011) Proteomic analysis of extracellular ATP-regulated proteins identifies ATP synthase β-subunit as a novel plant cell death regulator. *Molecular and Cellular Proteomics*, **10**, 1–13.

Choi J., Tanaka K., Cao Y., Qi Y., Qiu J., Liang Y., Lee S.Y., Stacey G. (2014) Identification of a Plant Receptor for Extracellular ATP. *Science*, **343**, 290–293. DOI:10.1126/science.343.6168.290.

Day R.C., Herridge R.P., Ambrose B.A., Macknight R.C. (2008) Transcriptome analysis of proliferating Arabidopsis endosperm reveals biological implications for the control of syncytial division, cytokinin signaling and gene expression regulation. *Plant Physiology*, **148**, 1964–1984.

Dilkes B.P., Comai L. (2004) A differential dosage hypothesis for parental effects in seed development. *The Plant Cell*, **16**, 3174–3180.

Dwivedi S.L., Perotti E., Upadhyaya H.D., Ortiz R. (2010) Sexual and apomictic plant reproduction in the genomics era: exploring the mechanisms potentially useful in crop plants. *Sexual Plant Reproduction*, **23**, 265–279.

Eames A.J. (1961) *Morphology of the Angiosperms*. New York, NY, USA: McGraw-Hill.

Egli D.B. (1998) *Seed Biology and the Yield of Grain Crops*. Wallingford, UK: CAB International.

Egli D.B. (2004) Seed-fill duration and yield of grain crops. *Advances in Agronomy*, **83**, 243–279.

Ellerman D.A., Myles D.G., Primakoff P. (2006) A role for sperm surface protein disulfide isomerase activity in gamete fusion: evidence for the participation of ERp57. *Developmental Cell*, **10**, 831–837.

Felitti S.A., Seijo G., González A.M., Podio M., Laspina N.V., Siena L.A., Ortiz J.P.A., Pessino S. (2011) Expression of *lorelei-like* genes in aposporous and sexual *Paspalum notatum* plants. *Plant Molecular Biology*, **77**, 337–354.

Garcia D., Saingery V., Chambrier P., Mayer U., Jurgens G., Berger F. (2003) Arabidopsis *haiku* mutants reveal new controls of seed size by endosperm. *Plant Physiology*, **131**, 1661–1670.

Garcia D., Fitz Gerald J.N., Berger F. (2005) Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis. *The Plant Cell*, **17**, 52–60.

Gates R.N., Quarin C.L., Pedreira C.G.S. (2004) Bahiagrass. In *Warm-season (C4) Grasses*. Eds L.E. Moser, B.L. Burson and L.E. Sollenberger. Madison, WI, USA: ASA, CSSA, and SSSA.

Ge X., Chen W., Song S., Wang W., Hu S., Yu J. (2008) Transcriptomic profiling of mature embryo from an elite super-hybrid rice LYP9 and its parental lines. *BMC Plant Biology*, **8**, 1–20.

Gehring M., Choi Y., Fischer R.L. (2004) Imprinting and seed development. *The Plant Cell*, **16**, S203–S213.

Grimanelli D., Hernández M., Perotti E., Savidan Y. (1997) Dosage effects in the endosperm of diplosporous apomictic *Tripsacum* (Poaceae). *Sexual Plant Reproduction*, **10**, 279–282.

Grimanelli D., Perotti E., Ramirez J., Leblanc O. (2005) Timing of the maternal-to-zygotic transition during early seed development in maize. *The Plant Cell*, **17**, 1061–1072.

Guillon F., Larré C., Petipas F., Berger A., Moussawi J., Rogniaux H., Santoni A., Saulnier L., Jamme F., Miquel M., Lepiniec L., Dubreucq B. (2012) A comprehensive overview of grain development in *Brachypodium distachyon* variety Bd21. *Journal of Experimental Botany*, **63**, 739–755.

Guo A.-Y., Zhu Q.-H., Gu X., Ge S., Yang J., Luo J. (2008) Genome-wide identification and evolutionary analysis of the plant specific SBP-box transcription factor family. *Gene*, **418**, 1–8.

Haig D., Westoby M. (1991) Genomic imprinting in endosperm: its effect on seed development in crosses between species, and between different ploidies of the same species, and its implications for the evolution of apomixis. *Philosophical Transactions of the Royal Society B*, **333**, 1–13.

van Hengel A.J., Guzzo F., Van K.A., de Vries S.C. (1998) Expression pattern of the carrot EP3 endochitinase genes in suspension cultures and in developing seeds. *Plant Physiology*, **117**, 43–53.

Huang C.-K., Huang L.-F., Huang J.-J., Wu S.-J., Yeh C.-H., Lu C.-A. (2010) A DEAD-Box protein, *AtRH36*, is essential for female gametophyte development and is involved in rRNA biogenesis in Arabidopsis. *Plant and Cell Physiology*, **51**, 694–706.

Johnston S.A., den Nijs T.P.M., Peloquin S.J., Hanneman R.E. (1980) The significance of genetic balance to endosperm development in interspecific crosses. *Theoretical and Applied Genetics*, **57**, 5–9.

Kim S.-Y., Mayandi Sivaguru M., Stacey G. (2006) Extracellular ATP in plants. Visualization, localization, and analysis of physiological significance in growth and signaling. *Plant Physiology*, **142**, 984–992.

Lai J., Dey N., Kim C.-S., Bharti A.K., Rudd S., Mayer K.F.X., Larkins B.A., Becraft P., Messing J. (2004) Characterization of the maize endosperm transcriptome and its comparison to the rice genome. *Genome Research*, **14**, 1932–1937.

Liu Y., Yan Z., Chen N., Di X., Huang J., Guo G. (2010) Development and function of central cell in angiosperm female gametophyte. *Genesis*, **48**, 466–478.

Lombardi L., Mariotti L., Picciarelli P., Ceccarelli N., Lorenzi R. (2012) Ethylene produced by the endosperm is involved in the regulation of nucellus programmed cell death in *Sechium edule* Sw. *Plant Science*, **187**, 31–38.

Milligan A.S., Lopato S., Langridge P. (2004) Functional genomics of seed development in cereals. In *Cereal Genomics*. Eds P.K. Gupta and R.K. Varshney. Amsterdam, Netherlands: Kluwer Academic Publishers.

Nowack M.K., Grini P.E., Jakoby M.J., Lafos M., Koncz C., Schnittger A. (2006) A positive signal from the fertilization of the egg cell sets off endosperm proliferation in angiosperm embryogenesis. *Nature Genetics*, **38**, 63–67.

Ohto M.-A., Stone S.L., Harada J.J. (2007) Genetic control of seed development and seed mass. In *Seed Development, Dormancy and Germination*. Eds K.J. Bradford and H. Nonogaki. Malden, MA, USA: Blackwell Publishing.

Opsahl-Ferstad H.G., Le D.E., Dumas C., Rogowsky P.M. (1997) *ZmEsr*, a novel endosperm-specific gene expressed in a restricted region around the maize embryo. *The Plant Journal*, **12**, 235–246.

Ozias-Akins P., van Dijk P.J. (2007) Mendelian genetics of apomixis in plants. *Annual Review of Genetics*, **41**, 509–537.

Pellny T.K., Lovegrove A., Freeman J., Tosi P., Love C.G., Knox J.P., Shewry P.R., Mitchell R.A. (2012) Cell walls of developing wheat starchy endosperm: comparison of composition and RNA-Seq transcriptome. *Plant Physiology*, **158**, 612–627.

Pfaffl N.W., Tichopad A., Prgomet C., Neuvians T. (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnology Letters*, **26**, 509–515.

Podio M., Felitti S.A., Siena L.A., Delgado L., Mancini M., Seijo J.G., González A.M., Pessino S.C., Ortiz J.P.A. (2014) Characterization and expression analysis of *SOMATIC EMBRYOGENESIS RECEPTOR KINASE* (*SERK*) genes in sexual and apomictic *Paspalum notatum*. *Plant Molecular Biology*, **84**, 479–495. DOI:10.1007/s11103-013-0146-9.

Polegri L., Calderini O., Arcioni S., Pupilli F. (2010) Specific expression of apomixis-linked alleles revealed by comparative transcriptomic analysis of sexual and apomictic *Paspalum simplex* Morong flowers. *Journal of Experimental Botany*, **61**, 1869–1883.

Pupilli F., Barcaccia G. (2012) Cloning plants by seeds: inheritance models and candidate genes to increase fundamental knowledge for engineering apomixis in sexual crops. *Journal of Biotechnology*, **159**, 291–311.

Quarin C.L. (1999) Effect of pollen source and pollen ploidy on endosperm formation and seed set in pseudogamous apomictic *Paspalum notatum*. *Sexual Plant Reproduction*, **11**, 331–335.

Reichler S.A., Torres J., Rivera A.L., Cintolesi V.A., Clark G., Roux S.J. (2009) Intersection of two signalling pathways: extracellular nucleotides regulate pollen germination and pollen tube growth via nitric oxide. *Journal of Experimental Botany*, **60**, 2129–2138.

Rozen S., Skaletsky H.J. (2000) Primer3 on the WWW for general users and for biologist programmers. In *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, pp. 365–386. Eds S. Krawetz and S. Misener. Totowa, NJ, USA: Humana Press.

Sabelli P.A., Larkins B.A. (2009) The development of endosperm in grasses. *Plant Physiology*, **149**, 14–26.

Sambrook J., Russell D.W. (2001) *Molecular Cloning: A Laboratory Manual*. New York, NY, USA: Cold Spring Harbor Laboratory Press.

Schmidt A., Schmid M.W., Klostermeier U.C., Qi W., Gutho D., Sailer C., Waller M., Rosenstiel P., Grossniklaus U. (2014) Apomictic and sexual germline development differ with respect to cell cycle, transcriptional, hormonal and epigenetic regulation. *PLoS Genetics*, **10**, e1004476. DOI:10.1371/journal.pgen.1004476.

Sharbel T.F., Voigt M.-L., Corral J.M., Galla G., Kumlehn K., Klukas C., Schreiber F., Vogel H., Rotterd B. (2010) Apomictic and sexual ovules of *Boechera* display heterochronic global gene expression patterns. *The Plant Cell*, **22**, 655–671.

Simon P. (2003) Q-Gene: processing quantitative real-time RT–PCR data. *Bioinformatics*, **19**, 1439–1440.

Stölting K.N., Gort G., Wüst C., Wilson A.B. (2009) Eukaryotic transcriptomics in silico: optimizing cDNA-AFLP efficiency. *BMC Genomics*, **10**, 565–579.

Tanaka K., Gilroy S., Jones A.M., Stacey G. (2010) Extracellular ATP signaling in plants. *Trends in Cell Biology*, **20**, 601–608.

Thiel J., Hollmann J., Rutten T., Weber H., Scholz U., Weschke W. (2012) 454 transcriptome sequencing suggests a role for two-component signaling in cellularization and differentiation of barley endosperm transfer cells. *PLoS ONE*, **7**, e41867. DOI:10.1371/journal.pone.0041867.

Tiwari S., Spielman M., Schulz R., Oakey R.J., Kelsey G., Salazar A., Zhang K., Pennell R., Scott R.J. (2010) Transcriptional profiles underlying parent-of-origin effects in seeds of *Arabidopsis thaliana*. *BMC Plant Biology*, **10**, 72–94.

Tucker M.R., Koltunow A.M.G. (2009) Sexual and asexual (apomicitic) seed development in flowering plants: molecular, morphological and evolutionary relationships. *Functional Plant Biology*, **36**, 490–504.

Vielle-Calzada J.-P., Crane C.F., Stelly D.M. (1996) Apomixis: the asexual revolution. *Science*, **274**, 1322–1323.

Vuylsteke M., Peleman J.D., van Eijk M.J.T. (2007) AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. *Nature Protocols*, **2**, 1399–1413.

Wang H., Boavida L.C., Ron M., McCormick S. (2008) Truncation of a protein disulfide isomerase, PDIL2-1, delays embryo sac maturation and disrupts pollen tube guidance in *Arabidopsis thaliana*. *Plant Cell*, **20**, 3300–3311.

Weber H., Sreenivasulu N., Weschke W. (2010) Molecular physiology of seed maturation and seed storage protein biosynthesis. In *Plant Developmental Biology – Biotechnological Perspectives*. Volume 2. Eds E.-C. Pua and M.R. Davey. Heidelberg, Germany: Springer.

Xiao X., Li H., Tang C. (2009) A silver-staining cDNA-AFLP protocol suitable for transcript profiling in the latex of *Hevea brasiliensis* (Para Rubber Tree). *Molecular Biotechnology*, **42**, 91–99.

Young B.A., Sherwood R.T., Bashaw E.C. (1979) Cleared pistil and thick sectioning techniques for detecting aposporous apomixis in grasses. *Canadian Journal of Botany*, **57**, 1668–1672.

Zhang Z., Schwartz S., Wagner L., Miller W. (2000) A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, **7**, 203–214.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1. Adaptors and primers used for cDNA-AFLP analysis

Table S2. Sequences of primers used for real-time PCR experiments

Table S3. Sequence search results of DETDFs with significant identities

Table S4. Sequence search results of *Paspalum notatum* DETDFs with endosperm and embryo-specific ESTs from different species