

Characterization of retrotransposon sequences expressed in inflorescences of apomictic and sexual *Paspalum notatum* plants

Ana Claudia Ochogavía · José Guillermo Seijo · Ana María González · Maricel Podio · Erica Duarte Silveira · Ana Luiza Machado Lacerda · Vera Tavares de Campos Carneiro · Juan Pablo A. Ortiz · Silvina Claudia Pessino

Received: 19 September 2010 / Accepted: 15 February 2011 / Published online: 11 March 2011
© Springer-Verlag 2011

Abstract Apomixis, an asexual mode of reproduction through seeds, holds much promise for agricultural advances. However, the molecular mechanisms underlying this trait are still poorly understood. We previously isolated several transcripts representing novel sequences differentially expressed in reproductive tissues of sexual and apomictic plants. Here, we report the characterization of two of these unknown RNA transcripts (experimental codes N17 and N22). Since original fragments showed no significant homologies to sequences at databases, preliminary genomic PCR experiments were carried out to discard possible contaminations. RACE extension on flanking regions provided longer sequences for the candidates and additional related transcripts, which revealed similarity to LTR retrotransposons carrying short transduplicated segments of protein-coding genes. Interestingly, some transduplicated segments corresponded to genes previously associated with apomictic development. Gene copy number estimations revealed a moderate representation of the elements in the

genome, with significantly increased numbers in a sexual genotype with respect to an apomictic one. Genetic mapping of N17 showed that a copy of this particular element was located onto *Paspalum notatum* linkage group F3c, at a central non-recombinant region resembling a centromere. Expression analysis showed an increased activity of N17 and N22 sense strands in ovules of the sexual genotypes. A retrotransposon-specific differential display analysis aimed at detecting related sequences allowed the identification of a complex family, with the majority of its members represented in the sexual genotype. Our results suggest that these elements could be participating in regulatory pathways related to apomixis and sexuality.

Keywords Apomixis · Apospory · Asexual reproduction · Gene expression regulation · Retrotransposon

Introduction

A classic definition for apomixis in angiosperms is “asexual (agamic) reproduction by seeds, i.e., agamospermy” (Nogler 1984). Apomictic plants generate seeds containing embryos that are genetically identical to the maternal parent, in the absence of both meiosis and fertilization. As apomixis produces genetic clones of a given individual via seeds, it has a great potential for use in agriculture. If the trait could be transferred to major crops, it would allow the fixation of heterosis or any given genetic combination. Exploitation of apomixis in crop plants would facilitate the indefinite use of hybrid vigor, the survival of combined genetic resources, the production of true seed from crops currently propagated vegetatively, and the establishment of more rapid breeding programs in response to changing needs and environments (Spillane et al. 2004).

Communicated by J.S. Heslop-Harrison.

A. C. Ochogavía · M. Podio · J. P. A. Ortiz · S. C. Pessino (✉)
Laboratorio Central de Investigaciones, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Parque Villarino, S2125ZAA, Zavalla, Santa Fe, Argentina
e-mail: pessino@arnet.com.ar

J. G. Seijo · A. M. González · M. Podio · J. P. A. Ortiz
Instituto de Botánica del Nordeste IBONE-CONICET,
Universidad Nacional del Nordeste, Sargento Cabral 2131,
3400 Corrientes, Argentina

E. Duarte Silveira · A. L. Machado Lacerda ·
V. Tavares de Campos Carneiro
Embrapa Recursos Genéticos e Biotecnologia, Universidade de Brasília, Parque Estação Biológica, Brasília, DF, Brazil

Agamospermy comprises two major types: adventitious embryony and gametophytic apomixis. While adventitious embryony involves the formation of somatic embryos directly from ovular tissues outside the embryo sac, gametophytic apomixis is a two-step process involving the production of non-meiotically reduced megagametophytes and the generation of an embryo by parthenogenesis. In gametophytic apomicts, embryo sacs with non-reduced nuclei ($2n$) can arise through variable mechanisms. When the non-reduced megagametophyte originates from the megaspore mother cell (MMC) itself, after lack or failure of meiosis, the process is classified as diplospory. Instead, when the $2n$ embryo sacs originate from companion nucellar cells, it is known as apospory. Aposporous individuals frequently form multiple non-reduced embryo sacs within the same ovule that can co-exist with the reduced meiotic ones that originate from the megaspore mother cell. The endosperm can be generated autonomously or after polar nuclei fertilization (pseudogamy), depending on the species (Crane 2002).

Paspalum notatum Flüggé (bahiagrass) is a subtropical perennial *Poaceae* native to South America. The common form of this species is a tetraploid ($2n = 4x = 40$), aposporous, pseudogamous, self-compatible biotype, whose populations spread from the north of Patagonia to Central Mexico, including the Caribbean islands (Burton 1948; Bashaw et al. 1970). Wild diploid ($2n = 2x = 20$), self-incompatible, sexual biotypes grow in a limited area in north-east Argentina (Burton 1946, 1955, 1967; Daurelio et al. 2004). Occasionally, triploid and pentaploid apomictic genotypes have been collected from natural populations (Quarin et al. 1989; Tischler and Burson 1995). Tetraploid sexual plants are not found in nature, but were produced at the laboratory (Quarin et al. 2001). Segregating progenies obtained from crosses between artificial sexual tetraploid mother plants and apomictic tetraploid pollen donors were used for determining the inheritance of apospory (Martínez et al. 2003) as well as for the construction of a frame genetic map of the species, where the locus responsible for apospory was mapped onto a large non-recombinant genomic region (Stein et al. 2004, 2007).

In a previous work, a comprehensive survey of the *P. notatum* inflorescence transcriptome led us to the identification of 65 unigenes activated or repressed during aposporous development (Laspina et al. 2008). Out of them, 45 corresponded to protein-coding plant sequences. The remaining 20 could not be successfully annotated, since they showed no significant homologies to sequences stored at the databases (Laspina et al. 2008). The objective of this work was to investigate the nature of 2 out of these 20 unknown candidates, named N17 and N22. After full characterization of the structure, genome representation and in situ expression, we arrived to the conclusion that

these sequences are putative LTR retrotransposons which carry transduplicated copies of protein-coding gene fragments, some of them corresponding to candidates previously associated with apomixis and sexual development. Several other related putative retrotransposon elements were identified.

Materials and methods

Plant material

The plant material used consisted of two tetraploid sexual genotypes of experimental origin (C4-4x and Q4188: $2n = 4x = 40$) (Quarin et al. 2001, 2003) and a natural aposporous plant (Q4117: $2n = 4x = 40$) (Ortiz et al. 1997). A total of 62 individuals from a pseudo-testcross F_1 population segregating for apospory, which was generated by crossing Q4188 \times Q4117, were used for mapping experiments. The F_1 mapping population had been previously used to produce a full genetic map of the species (Stein et al. 2007).

Spikelets collected from immature inflorescences were used for RNA isolation and in situ hybridization experiments. Typically, the *Paspalum notatum* inflorescence has two rachises at the apex of the flowering culm which produce a Y shape. Each rachis carries two rows of spikelets. The developmental stage at which spikelets were collected was evaluated by analyzing the macromorphology of inflorescences and the stage of pollen development, following the methods and the reproductive calendar reported in Laspina et al. (2008).

Genomic PCR amplifications

Genomic DNA was extracted from 6 g of young leaves using the CTAB method (Shagai-Marouf et al. 1984). PCR were carried out in a final volume of 25 μ l including 60 ng DNA, 0.2 μ M specific primers, $1\times$ PCR buffer (Promega), 2.5 mM $MgCl_2$, 200 μ M dNTPs, and 1.5 U Taq polymerase (Promega). Cycling was done in an MJ Research thermocycler and consisted of the following steps: 5 min at 94°C, 40 cycles of 30 s at 94°C, 1 min at 55–69°C (depending on the primer pair used), and 1 min at 72°C. A final elongation of 10 min at 72°C was also included. Amplified fragments were loaded onto 5% polyacrylamide gels, electrophoresed in TBE 1X buffer, and silver-stained.

Construction of a non-cloned cDNA library

A Marathon non-cloned cDNA library (BD Biosciences Clontech) was prepared from total RNA obtained from Q4117 spikelets at late pre-meiotic developmental stage I,

following the *P. notatum* reproductive calendar reported in Laspina et al. (2008). This stage is immediately previous to the onset of the aposporous initials. Total RNA was isolated from inflorescences using the SV Total RNA Isolation System (Promega). Poly(A) + RNA was purified from RNA samples using Dynabeads (Dyna[®]), according to manufacturer's protocol. Poly(A) + RNA was reverse-transcribed, double-strand cDNA was synthesized and linked to Marathon cDNA adaptors following the Marathon cDNA amplification kit (BD Biosciences Clontech) protocol.

RACE reactions

Two pairs of nested reverse-oriented gene-specific primers were designed for each candidate sequence in order to perform 3'- and 5'-RACE (Rapid Amplification of cDNA Ends) experiments (Chenchic et al. 1996) following the recommendations of the Marathon cDNA amplification kit (BD Biosciences Clontech). Primer sequences are shown in Table 1. Oligonucleotides were 23–28 nucleotides (nt) long and had 50–70% GC content with a melting

temperature $\geq 67^\circ\text{C}$. PCR were carried out in a 50 μl final volume containing 2 μl of Marathon library product (BD Biosciences Clontech), 1 \times GoTaq activity buffer (Promega), 200 μM dNTPs, 0.2 μM of the gene-specific primer, 0.2 μM of adaptor-specific primer (AP1 or AP2), and 1.5 U of GoTaq DNA polymerase enzyme (Promega). Initial PCR conditions were the following: 94 $^\circ\text{C}$ for 1 min followed by 30 cycles of 30 s at 94 $^\circ\text{C}$ and 4 min at 68 $^\circ\text{C}$ (both annealing and polymerization temperatures were 68 $^\circ\text{C}$). To obtain the final 3' or 5' RACE product, 2–4 PCR rounds were performed. The cycle temperatures were optimized according to the T_m of the particular set of primers used for amplification. Positive and negative controls were included. Positive controls consisted of amplifications with two specific oligonucleotides matching in opposite sense, which amplified a small segment within the original sequence fragment. Negative controls consisted of amplification reactions using specific and adaptor-complementary primers in the absence of template DNA. After examination by gel electrophoresis, products were isolated by using the SV Wizard Gel and PCR Clean Up System (Promega). The transformation protocol was taken from the Molecular Cloning Laboratory Manual (Sambrook et al. 1989). Plasmids were purified with Wizard Plus SV Minipreps (Promega). Insert verification was done by PCR using the M13 forward and reverse primers and the following amplification conditions: 94 $^\circ\text{C}$ for 1 min, 25 cycles of 94 $^\circ\text{C}$ for 30 s, 55–68 $^\circ\text{C}$ for 1 s, 72 $^\circ\text{C}$ for 1 min. Sequencing of the 5' and 3' RACE clones was done by Macrogen Inc (Korea). Both strands were sequenced for several candidates by using M13 Forward and Reverse primers, in order to control sequencing quality.

Bioinformatic analysis

Vector contaminations were identified and removed by using the VecScreen tool at the National Center for Biotechnology Information (NCBI) webpage (<http://www.ncbi.nlm.nih.gov/VecScreen/>). Alignments between the overlapped 3' and 5' RACE amplification products and the original sequence were done with Clustal W2 on the EBI-EMBL website (<http://www.ebi.ac.uk/Tools/clustalw2>). Analysis of DNA similarity was carried out using the BLASTn and BLASTx packages at NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>), the Arabidopsis Information Resource website (<http://www.arabidopsis.org/blast/>), and the Gramene web site (<http://www.gramene.org/>). Open reading frames were searched with the ORF finder at NCBI (<http://www.ncbi.nlm.nih.gov/projects/gorf/>). Search for repetitive sequences was done against the Gramineae repeats TIGR database by using the tools available at the J. Craig Venter Institute webpage (<http://www.tigr.org/db.shtml>) (Ouyang and Buell 2004).

Table 1 List of oligonucleotides used as primers in RACE, differential display, real-time PCR, and mapping experiments

Primer name ^a	Sequence
3R-17 upper 1	5'-TTGTGAGTATAGAGTGC GGAGTCG-3'
3R-17 upper 2	5'-CAGTCTAGGATTGAGTAAAGTTC-3'
5R-17 lower1	5'-AAGTAAGGATGAAATAACAAAGTAAGA-3'
5R-17 lower 2	5'-GCATGCACAAGGTTCCCGTAAA-3'
3R-22- upper 1	5'-CCAACAGAAAGAGGGGCACTGGACACAA-3'
3R-22- upper 2	5'-CTGCTGGAGGGAGCTGGTAACTTCATCC-3'
5R-22- lower 1	5'-AGGGTTCAGCCGCTCGAATCAGC-3'
5R-22- lower 2	5'-ATCAGCCGTCGCCAGAAATCATCAACAG-3'
AP1	5'-CCATCTAATACGACTCACTATAGGGC-3'
AP2	5'-ACTCACTATAGGGCTCGAGCGGC-3'
N20 upper	5'-GTGTGGCAGTAGCGTTGTTGTTC-3'
N20 lower	5'-TACGGCCGACCCACCATTAGCA-3'
GAPDH upper	5'-TGAATCTAGTCCATCCGCTTG-3'
GAPDH lower	5'-TCATCAGGCAGGGAAGCTA-3'
TSDD-22-Fw	3'-CTGGTAACTTCATCC-3'
TSDD-22-Rv	5'-CCAGAAATCATCAAC-3'

^a Primers 5R and 3R were used in 5' and 3' RACE nested amplifications, respectively. Primers 3R-17-upper1/5R-17-lower2 and 3R-22-upper1/5R-22-lower1 were used in cDNA and/or genomic real-time PCR, polymorphism detection and/or mapping experiments. AP1 and AP2 are oligonucleotides complementary to the Marathon library adapters. N20 upper and N20 lower were used to amplify the LORELEI-like GPI-anchored protein used as reference gene in gene copy number analysis. GAPDH upper and GAPDH lower were used to amplify the glucose-6-phosphate dehydrogenase used as control of equal expression in reproductive tissues of sexual and apomictic plants. TSDD primers were designed on the N22 sequence to be used in the transposon-specific differential display assays

Real-time PCR experiments

To determine gene copy numbers, RT-PCR were carried out in a final volume of 25 μ l containing 200 nM gene-specific primers (Table 1), 1 \times Realmix qPCR (Biodynamics), and the corresponding quantity of recombinant plasmid or genomic DNA (see below). Calibration curves were obtained by including standards of 5, 50, 500, 5000, and 50000 copies of N17 and N22 fragments cloned in pGEMTeasy (Promega). Genomic DNA was extracted from young leaves by using the CTAB method (Shagai-Marooof et al. 1984). Four different dilutions of genomic DNA containing 30×10^{-9} g/ μ L, 30×10^{-11} g/ μ L, 30×10^{-12} g/ μ L, and 30×10^{-13} g/ μ L were tested in triplicate. Only those values comprised in the lineal range of the curve were considered for analysis. Gene N20, which encodes a *P. notatum* GPI-anchored LORELEI-like protein, was included in the assays as an equal copy number reference. Non-template controls were incorporated to the assays. Amplifications were performed in an Rotor-Gene Q thermocycler (Qiagen), programmed as follows: 2 min at 94°C, 45 cycles of 15 s at 94°C, 30 s at 63°C, 40 s at 72°C. A melting curve (10 s cycles from 72 to 95°C, where temperature was increased by 0.2°C after cycle 2) was produced at the end of the cycling.

RT-PCR was also used in quantitative expression analysis. Total RNA was extracted from spikelets collected at early pre-meiosis (stage 0), late pre-meiosis/meiosis (stage I/II), and post-meiosis (stage IV/V) according to Laspina et al. (2008), by using the SV Total RNA Isolation kit (Promega). The derived cDNA was synthesized with Superscript II (Invitrogen). RT-PCR were carried out in a final volume of 25 μ l containing 200 nM gene-specific primers (Table 1), 1 \times Realmix qPCR (Biodynamics), and 20 ng of reverse-transcribed RNA (prepared by using Superscript II, Invitrogen-Life Technologies). Specific PCR primer pairs were designed by using Primer 3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). Glucose-6-phosphate dehydrogenase was used as an equal-expression reference, since this housekeeping gene was reported to be one of the most stable candidates in reproductive organs of related apomictic and sexual *Brachiaria brizantha*, a related aposporous grass (Duarte Silveira et al. 2009). Oligonucleotides were synthesized by IDT (Integrated DNA technologies, <http://www.idtdna.com/Home/Home.aspx>). RT (–) and non-template controls were incorporated to the assays. All apomictic and sexual samples and controls were amplified in triplicate. Amplifications were performed in a Rotor-Gene Q thermocycler (Qiagen), programmed as follows: 2 min at 94°C, 45 cycles of 15 s at 94°C, 30 s at 63°C, 40 s at 72°C. A melting curve (10 s cycles from 72 to 95°C, where temperature was increased by 0.2°C after cycle 2) was

produced at the end of the cycling. Relative quantitative expression was estimated by using REST-RG (Relative Expression Software Tool V 2.0.7 for Rotor Gene, Corbett Life Sciences), considering the take off values and amplification efficiency for each particular reaction.

Mapping procedure

Specific primers (Table 1) were used to amplify the extended fragments from genomic DNA, and search for mapping useful polymorphisms between the parents of the mapping population (Q4188 and Q4117). Amplification products were electrophoresed in 5% polyacrylamide gels and silver-stained. The presence/absence of polymorphic fragments between Q4188 and Q4117 genotypes was visually determined and recorded for each progeny. Segregation data were analyzed as described by Stein et al. (2007). A χ^2 test was used to determine the goodness of fit (at $P \leq 0.05$) between the observed and the expected number of genotypes for each class of segregation ratio. Bands that showed segregation ratios corresponding to single-dose markers (1:1) were incorporated to previously generated marker-segregation matrixes of the species (Stein et al. 2007). Map units in centimorgan (cM) were derived from the Kosambi (1944) mapping function. Linkage analysis was carried out by using MAPMAKER/EXP 3.0 (Lander et al. 1987) at a minimum LOD threshold of 3.0 and a recombination frequency $\theta = 0.30$.

Reproductive tissue in situ hybridization

In situ hybridizations were performed on reproductive tissues from Q4117 (apomictic genotype) and C4- 4x (sexual genotype) at developmental stage I (late premeiotic) (Laspina et al. 2008). Spikelets were fixed in 4% paraformaldehyde/0.25% glutaraldehyde in 0.01 M phosphate buffer pH 7.2, dehydrated in an ethanol series, and embedded in paraffin. Specimens were cut into 7 μ m thin sections and placed onto slides treated with poly-L-lysine 100 μ g/ml. Paraffin was removed with a xylene series. Plasmids containing the selected clones were linearized using restriction enzymes *NcoI* or *SalI* (Promega). Probes were labeled with the Roche Dig RNA Labeling kit (SP6/T7), following the manufacturers' instructions. Template digested with *SalI* restriction enzyme was used to produce a probe from the T7 transcription start. Template digested with *NcoI* restriction enzyme was used to produce a probe from the SP6 transcription start. Probes were hydrolyzed to 150–200 bp fragments. Prehybridization was carried out in a buffer of 0.05 M Tris–HCl pH 7.5 containing 1 μ g/ml proteinase K in a humid chamber at 37°C for 10 min. Hybridization was carried out overnight in a humid chamber at 42°C, in buffer containing 10 mM Tris–HCl pH 7.5, 300 mM NaCl, 50% formamide (deionized), 1 mM

EDTA pH 8, 19 Denhardt, 10% dextranulphate, 600 ng/ml total RNA, and 60 ng of the corresponding probe. Detection was performed following the instructions of the Roche Dig Detection kit, using anti-DIG AP and NBT/BCIP as substrates.

Retrotransposon-specific differential display experiments

Total RNA was obtained from Q4117 (apomictic) and Q4188 (sexual) panicles by using the SV Total RNA Isolation kit from Promega. Reverse transcription was performed using Superscript II reverse transcriptase (Invitrogen-Life Technologies) as indicated by the manufacturers. Differential display experiments were conducted under the general protocol reported by Liang and Pardee (1992) with minor modifications. The anchored oligonucleotides used were named DDT1, DDT2, DDT3, DDT4, and corresponded to the sequence 5'T₍₁₂₎(ACg)X3', where X was A, C, G, or T, respectively. Specific primers corresponding to the N22 sequence (Table 1) were used to create primer pair combinations with the anchored oligonucleotides. PCR were carried out in a final volume of 25 μ l containing 1 \times Taq activity buffer (Promega), 1.5 mM MgCl₂, 50 μ M dNTPs, 0.70 μ M arbitrary primer, 2.5 μ M anchored primer, 2 U of Taq DNA polymerase enzyme (Promega), and 2.5 μ l of the reverse transcription reaction (previously diluted 1/20). All samples (including controls) were processed in duplicate. Negative control reactions were performed as follows: (1) using total non-reverse-transcribed RNA as template, to verify the absence of chromosomal DNA in the RNA preparations; (2) using sterile distilled water instead of template, to discard contaminations from the reagents. The cycling program consisted of an initial step of 3 min at 94°C, 40 cycles of 20 s at 94°C, 20 s at 38°C, and 30 s at 72°C, followed by a final step of 5 min at 72°C. Samples were mixed with denaturing loading buffer and were treated for 3 min at 95°C and separated in a 5% (w/v) polyacrylamide gels. Amplification products were silver-stained. Bands were scored only in the middle portion of the gel, where resolution was maximal and profiles were fully reproducible. Differentially expressed bands were cut, eluted, and cloned. Sequencing was done by Macrogen Korea. Sequences were used in Blastn surveys at NCBI (green plants non-redundant databases) and TIGR (Gramineae plant repeats databases). Sequences fragments (corresponding only to repetitive elements) were analyzed with the Phylogeny platform (<http://www.phylogeny.fr>). The multiple alignment program T-Coffee was used with the default parameters. For tree construction, the program Fast-Dist/ProtDist + BioNJ/Neighbor (PHYLYP) 3.66 was used setting 1000 bootstraps. The TreeDyn program was used to visualize the tree.

Results

Genomic amplification controls

Original N17 and N22 cDNA fragments showed no significant homologies in plant databases Blast searches. In order to investigate the presence of the corresponding genomic sequences and discard contaminations, oligonucleotides that amplified short internal segments of each clon were used to produce PCR amplicons from both aposporous and sexual tetraploid *P. notatum* genomic DNA. All amplification reactions generated fragments of the predicted molecular weight, as well as a few additional

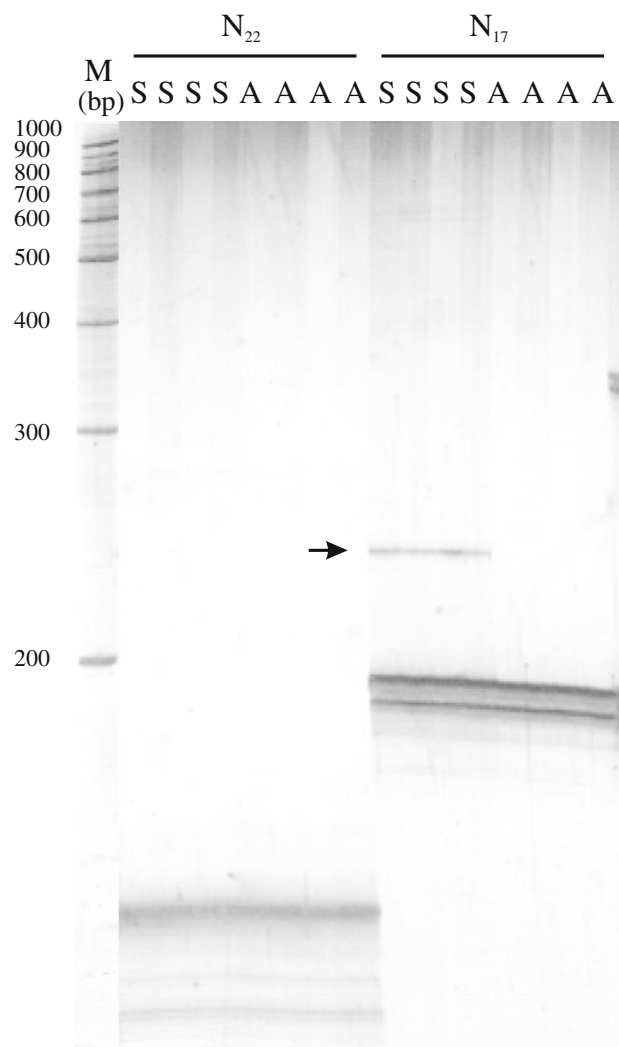


Fig. 1 Genomic DNA PCR amplification of selected candidates N17 and N22 from the sexual Q4188 (S) and the apomictic plant Q4117 (A). Two technical replicates were amplified by duplicate. All amplicons showed a moderate number of bands, suggesting the presence of different related sequences in the *P. notatum* genome. A band originated from N17 was specifically amplified in the sexual genotype (arrow). M: 100 kb ladder (number of base pairs is indicated)

bands (Fig. 1). These results indicated that both sequences occurred in the *P. notatum* genome and probably involve a moderate number of repetitive copies. Interestingly, amplification of N17 originated a polymorphic band between parental plants Q4188 and Q4117, which was used for localizing the transcript in the *P. notatum* genome (see below).

RACE experiments

Both N17 and N22 fragments isolated through differential display experiments had been amplified from apomictic samples by using random decamers (Laspina et al. 2008) and were 476 and 303 nucleotide long, respectively. In order to extend the flanking regions, two pairs of nested oligonucleotides (upper and lower) were used in RACE (Rapid Amplification of cDNA Ends) experiments (Chenchic et al. 1996) (Table 1). The specific matching position for each of the primers as well as the expected size of the amplification products are indicated in Fig. 2.

Primers designed for sequence N17 allowed to extend it toward the 3' end (from 476 to 907 bp). The opposite extension (5' RACE) did not produce any amplicon. Primers designed for sequence N22 did not allow to extend the clone, but amplified several related sequences that could not be treated as a consensus. Even when the original oligonucleotides were found located at the edge of the amplified sequences, the internal overlapped regions did not match exactly with N22. Instead, they were considered to be different members of a related transcript family. Five different related sequences (A43, A44, A45, A46, and A47, which were 849, 827, 784, 643, and 802 bp long, respectively) were isolated from N22 RACE experiments.

Bioinformatic surveys of original cDNAs and extended fragments

Each one of the RACE-derived sequences was used to interrogate public databases, in order to detect similarities with genes from related and model plant species.

N17 surveys at NCBI, Gramene, and TAIR revealed partial homology to retrotransposons. When compared against the Gramineae Plant Repeats database at TIGR, the consensus sequence presented similarity with the LTR region of ORSiCMCM00100010 Above centromeric LTR-retrotransposon ($eV = 9.2e^{-06}$), from nt 139 to 334. However, there was a conserved sector within the sequence (from nt 816 to 878) that presented similarity to gene Os11g0269700, a hypothetical protein ($eV = 2e^{-11}$, identities 56/64, 87%). Therefore, N17 seemed to be a centromeric retrotransposon carrying a 62 nt sequence similar to a plant protein-coding gene. More recent N17 surveys at TIGR revealed homology with the long terminal repeats region (LTR) of SRSiTERTOOT00011 gil13400380|gbl AZ922097.11AZ922097 HRCot2G11 *Sorghum bicolor*, similar to *Sorghum bicolor* Retrosor-6 retroelement LTR (E-value 1.7×10^{-4}).

As already mentioned, RACE extension of clone N22 had rendered different related segments that did not overlap perfectly (A43, A44, A45, A46, A47). Therefore, they could not be assembled in a single contig, but BLAST surveys were performed for each member independently. Low-stringency searches at the TIGR Gramineae repetitive database showed that the original N22 sequence was similar to the gag-protein-coding sector (5648-5944 nt) of ORSiTERT00200152, a Rire8 *Ty3-gypsy*-like retrotransposon (E-val: 0.97). Clone A43 showed homology to a Gypsy20 retrotransposon (gblEU558521.11 *Arabidopsis lyrata* clone Gypsy20 transposon-insertion display, $eV = 8e^{-65}$) between nt 198–756, while it was highly similar to *Poa pratensis serk2* gene ($eV = 3e^{-28}$) between nt 43–188. Clone A44 showed homology to the same LTR-retrotransposon (gblEU558521.11, $eV = 8e^{-35}$) between nt 421–685, while it showed similarity to *Zea mays* cytochrome P450 (*cyt P450*) between nt 1–382 ($eV = 6e^{-100}$). Clone A45 revealed homology with the same retrotransposon (gblEU558521.11, $eV = 1e^{-16}$), but no homologies to plant genes were detected. Clone A46 showed homology with the same retrotransposon (gblEU558521.11,

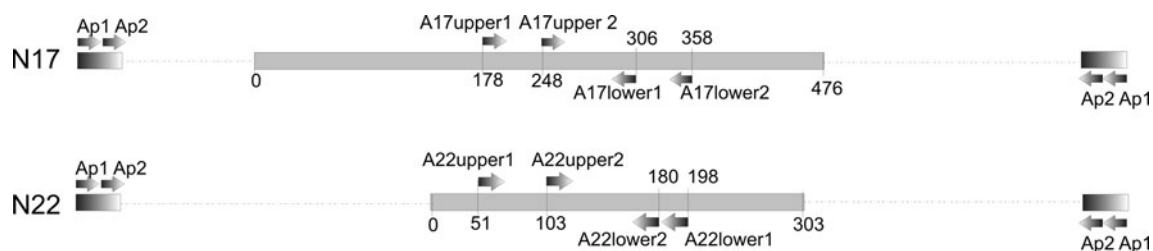


Fig. 2 RACE amplification strategy. Schematic representation of candidate initial fragments and oligonucleotides designed to perform RACE experiments. AP1 and AP2 indicate the position of the primers

matching the Marathon library terminal adaptors. Nucleotide number positions corresponding to primer locations were indicated below

$eV = 2e^{-60}$) in the 4–60 and 273–569 intervals, and to the *P. pratensis serk2* ($eV = 2e^{-54}$) between nt 76 and 255. Clone A47 presented homology at the TIGR Gramineae plant repeat database to another retrotransposon (ZRSiT-ERTO00146, dagaf_576C20-1 retrotransposon, $eV = 3e^{-8}$). These results indicate that all N22-related fragments might correspond to family members showing common similarity to retrotransposons, some of them including transduplicated segments of plant protein-coding genes. Interestingly, the plant protein-coding fragments detected corresponded to genes previously associated with apomixis. Both *serk* and *cytP450* had been reported to be differentially regulated during apomictic development in *Poa pratensis*, *Paspalum notatum*, and *Eragrostis curvula* (Albertini et al. 2004, 2005; Laspina et al. 2008; Cervigni et al. 2008). A diagram showing features of the N22-related sequences is shown in Fig. 3.

N17 and N22 gene copy number estimation

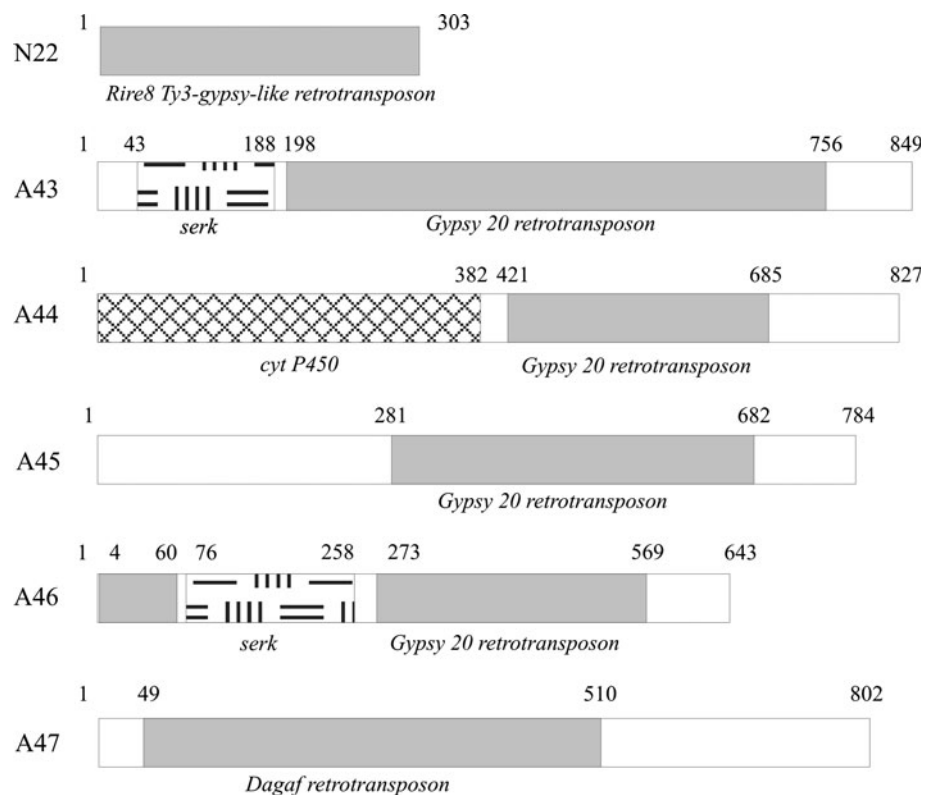
An estimation of the number of copies of N17 and N22 in the genome of sexual and apomictic *P. notatum* plants was carried out by performing quantitative real-time PCR experiments. A protein-coding gene (N20), encoding a LORELEI-like GPI-anchored protein, was used as an equal representation reference (Capron et al. 2008). In order to calculate gene copy number, a weight of 0.6 pg per monoploid *P. notatum* genome was considered, since this is

the C value for diploid *P. notatum* (Jarret et al. 1995). Occasionally, a variable reduction in genome size have been reported during polyploidization events (Pessino and Martelotto 2006). Therefore, a more accurate copy number determination would be achieved if the actual mass of the monoploid genome in tetraploid *P. notatum* were determined.

N17 showed an average copy number/monoploid genome of 50.4 in the apomictic genotype and 460.4 in the sexual one. Differences in the copy numbers between genotypes were highly significant (standard deviations of 11.8 and 49.8, respectively, sexual/apomictic representation ratio of 9.13). N22 showed an average copy number/monoploid genome of 16.3 in the apomictic genotype and 53.2 in the sexual genotype. Again, copy numbers were differential at a highly significantly level (standard deviations of 5.3 and 0.56, respectively, sexual/apomictic representation ratio of 3.26). The reference gene N20 (LORELEI-like protein) showed an average copy number/monoploid genome of 4.33 in the apomictic genotype and 2.83 in the sexual genotype. As expected, N20 copy numbers were not differential at a significant level (standard deviations of 0.98 and 0.55, respectively, sexual/apomictic representation ratio of 0.65). Our results are in agreement with previous analysis in *Arabidopsis thaliana*, where the LORELEI-like family includes only four member genes (Capron et al. 2008).

Altogether, these results confirm that N17 and N22 are members of a repetitive family that is moderately

Fig. 3 N22-related sequences amplified by RACE. Sequences displaying similarity to retrotransposons were represented as *gray boxes*. Sequences displaying similarity to plant protein-coding genes were represented as *textured boxes*. Sequences with no similarity were represented as *white boxes*. Nucleotide numbers are indicated at the top of the represented fragment. N22 sequence is similar to sector 5648–5944 of ORSiTERT00200152 (E-val 0.97), a rir8 Ty3-gypsy-like retrotransposon, which in turn keeps homology to the gag-protein-coding 56–130 sector of retrotransposon multiprotein AAX95147



represented in *P. notatum*. Interestingly, the genomic copy number is different in the apomictic and sexual plants analyzed, with a higher representation in the sexual genotype. On the contrary, gene N20 (LORELEI-like GPI-anchored protein) is represented as a low-copy family (three to four copies) in the monoploid *P. notatum* genome, and no differential representation could be observed between the sexual and the apomictic genotypes. A graphic representation of the estimated copy numbers for N17, N22, and reference gene N20 is shown in Fig. 4.

Mapping experiments

Since N17 displayed homology to centromeric retrotransposons, we conducted mapping experiments aimed at determining its location onto the *P. notatum* genome. Previous assays (Fig. 1) had revealed the occurrence of a polymorphic band useful for mapping between the parents of the mapping population available at our laboratory (Q4188 and Q4117). Polymorphism segregation was followed in 62 F₁ individuals of the pseudo-testcross F₁ mapping population segregating for apospory previously used to construct a tetraploid map of the species (Stein et al. 2007). Presence/absence data were converted to a binary matrix and processed together with the entire data file of the *P. notatum* genetic map (Stein et al. 2007) by using MAPMAKER 3.0 (Lander et al. 1987).

N17 resulted linked in coupling phase to genetic linkage group F3c, onto a central low-recombinant region (Fig. 5). Moreover, N17 was linked in coupling at a genetic distance of ~30 cM to another gene differentially expressed in inflorescences of apomictic with respect to sexual plants,

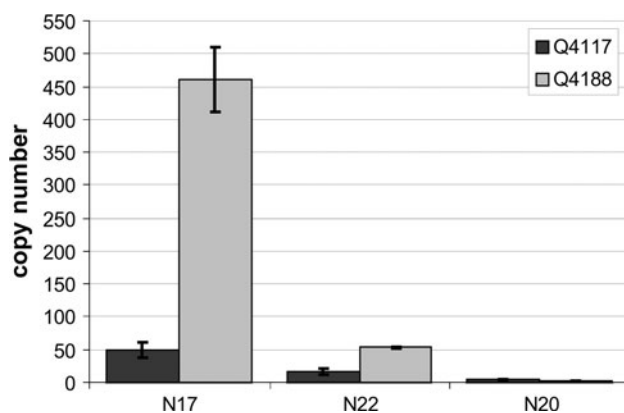


Fig. 4 Estimated gene copy number per monoploid genome in an apomictic (Q4117) and a sexual (Q4188) *P. notatum* tetraploid genotypes. Both N17 and N22 are represented at a moderate copy number typical of repetitive sequences. A higher copy number is observed in the sexual genotype. Protein-coding gene N20 (which encodes a LORELEI-like GPI-anchored protein) is equally represented at a low copy number (three to four copies) in the sexual and the apomictic genomes

the kinesin Mc10 (Stein et al. 2007). The position of N17 in a central low-recombinant chromosomal area is in agreement with sequence similarity bioinformatic surveys, which indicated homology to centromeric retrotransposons. Since N17 resulted not genetically linked to the genomic region governing apospory in *P. notatum*, it does not qualify as a primary genetic determinant of this reproductive system. However, it could be performing a role as a downstream participant in the reproductive developmental process.

Lack of polymorphisms produced by N22 between parental genotypes Q4117 and Q4188 frustrated attempts to map the sequence onto the *P. notatum* genome. The primers used to search for polymorphisms were located within the original N22 sequence, and therefore, they did not span transduced gene fragments (the original N22 sequence was slightly similar to a Rire8 Ty3-gypsy-like retrotransposon sequence, but no gene transduplication was detected within this sequence). SSCP or RFLP analysis should be conducted in order to generate polymorphisms useful for mapping.

Reproductive tissue in situ hybridization

Labeled mRNA probes complementary to both nucleotide strands were used in reproductive tissue in situ hybridization experiments, in order to assess differential expression of the N17 and N22 target sequences (Fig. 6). Original (non-extended) N17 and N22 clones were used for labeling. The developmental phase at which flowers were examined corresponded to stage I at the *P. notatum* reproductive calendar (Laspina et al. 2008). This stage (late pre-meiotic) immediately precedes the emergence of apospory initials and/or the MMC in the nucellus. Candidate transcripts displaying differential expression had been isolated from inflorescences of aposporous genotypes at the same stage (Laspina et al. 2008).

The antisense (*NcoI*/SP6) N17 probe originated a moderate signal in the anther tapetum and a strong signal in ovules of the sexual plant, while in the apomictic counterpart no significant signal could be distinguished. The sense (*SalI*/T7) N17 probe produced a very faint signal for both the sexual and the apomictic plants (Fig. 6, left panel). These results indicated that the sense strand of the putative N17 retrotransposon was expressed at higher levels in reproductive tissues of the sexual plant with respect to the apomictic one.

Similarly, the antisense (*NcoI*/SP6) N22 probe produced a very strong signal in anthers and ovules of the sexual genotype and a strong signal only in anthers of the apomictic one. Signal in apomictic ovules was absent. Meanwhile, the sense (*SalI*/T7) N22 probe generated a low-intensity identical signal for both the apomictic and the

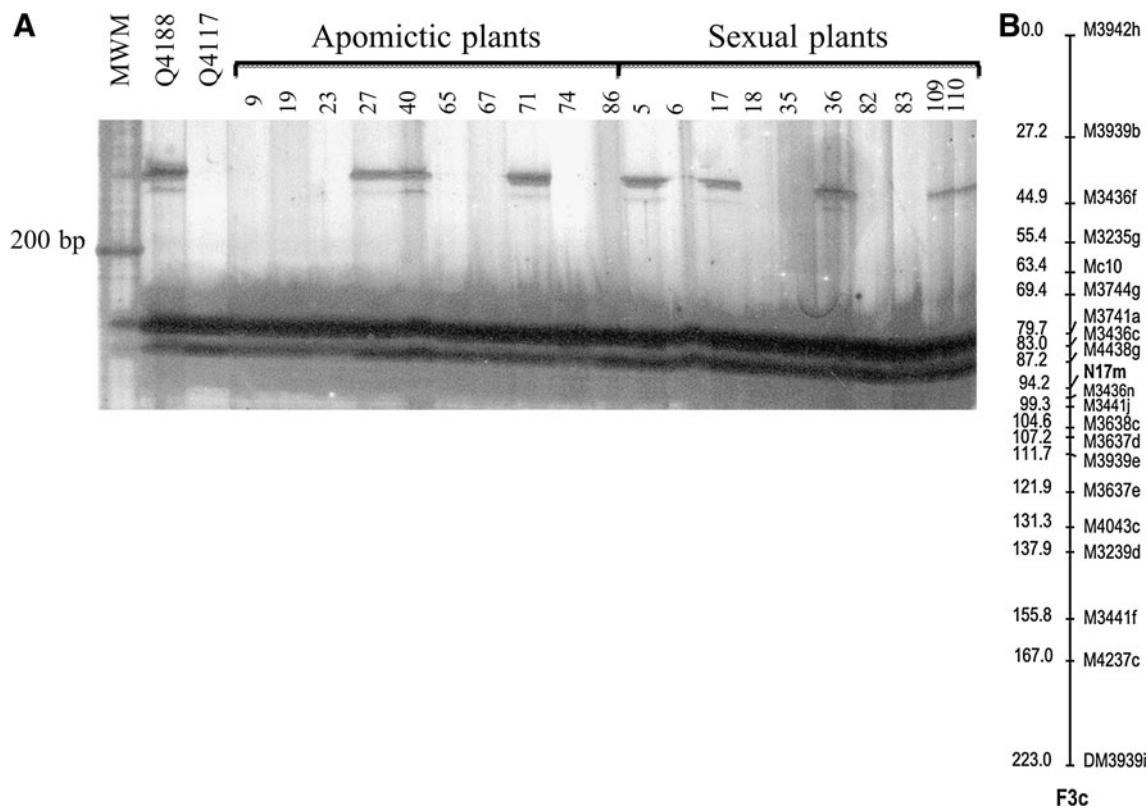


Fig. 5 Genomic location of N17 internal fragment as revealed through mapping experiments. **a** Segregation observed in a selected group of 10 apomictic and 10 sexual F_1 progeny individuals. **b** *P. notatum* linkage group F3c showing the location of the N17 fragment

sexual genotype (Fig. 6, right panel). Again, the sense strand of the putative retrotransposon was up-regulated in the sexual plant.

Quantitative expression analysis

A chronological relative quantitation of N22 expression was achieved by performing real-time PCR. Inflorescences samples were collected at early premeiotic (0), late premeiotic/meiotic (I/II), and post-meiotic (IV/V) stages, following the reproductive calendar reported by Laspina et al. (2008). Primers were designed to specifically amplify the original N22 fragment. Glucose-6-phosphate dehydrogenase (GAPDH) was used as a reference gene, since it was reported to be one of the most stable housekeeping genes in apomictic vs. sexual-related reproductive systems (Duarte Silveira et al. 2009).

Data corresponding to N22 chronological relative expression levels as revealed by real-time PCR experiments were represented in Fig. 7. Expression ratios were established with respect to the apomictic plant at early premeiosis (the stage with the lowest expression). In the apomictic genotype, the expression levels increased to 3.476 at meiosis (Std error: 1.642–8.914) and to 14.64 at

within a low-recombinant central region, close to it, another gene activated in inflorescences of apomictic plants with respect to sexual ones was detected (kinesin Mc10). Segregation data of all markers corresponded to those reported by Stein et al. (2007)

post-meiosis (Std. error: 6.924–33.006). In the sexual genotype, expression level was low and similar to that of the apomictic plant at pre-meiosis (1.447, Std error: 0.414–1.133), drastically increased to 22.187 at meiosis (Std error: 16.052–36.401) and remained with a similar value at post-meiosis (22.157, Std error: 14.625–32.269). At the meiotic stage, the expression ratio between the sexual and the apomictic genotype was 6.38, which is in agreement with the results observed in the in situ hybridization experiments. As the real-time PCR experiments do not allow distinction between sense and antisense expression, these values should be considered a putative additive contribution of both transcript types.

Isolation of a family of retroelements differentially expressed in apomictic and sexual plants

Based on the results obtained from RACE and in situ hybridization experiments, we hypothesized that complex families of related retrotransposon sequences might have been subjected to differential expression regulation during aposporous and sexual development. As it was stated before, attempts to extend the N22 sequence by 5' and 3' RACE had already allowed the detection of a group of

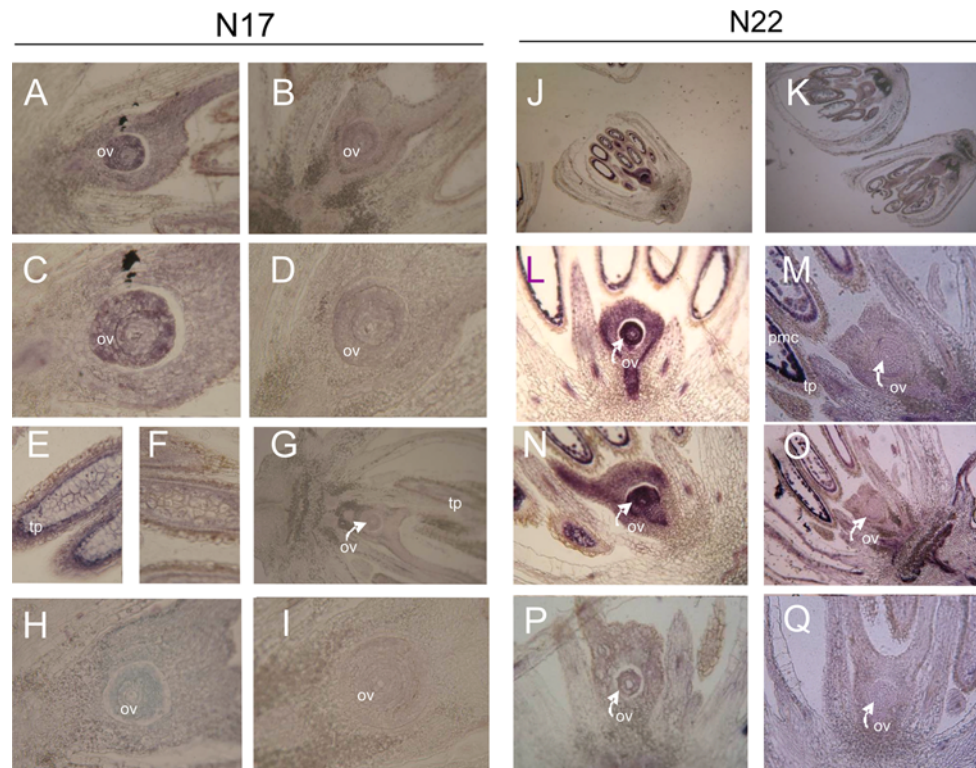


Fig. 6 Reproductive tissue in situ hybridization from sexual (C4-4X) and apomictic (Q4117) genotypes. **a** Sexual genotype hybridized with the antisense *NcoI* N17 probe showed a strong signal in the ovule. **b** Apomictic genotype hybridized with antisense *NcoI* N17 probe, with no signal detected in the ovule. **c** A closer image of *panel A*. **d** A closer image of *panel B*. **e** Anthers from sexual genotype hybridized with antisense *NcoI* N17 probe, showing a moderate signal at the tapetum. **f** Anthers from apomictic genotype hybridized with antisense *NcoI* N17 probe, with no signal detected. **g** and **i** Apomictic genotype hybridized with sense *SalI* N17 probe with no signal detected. **h** Sexual genotype hybridized with sense *SalI* N17 probe.

No signal was detected. **j** Panoramic view of the sexual genotype reproductive tissues hybridized with antisense *NcoI* N22 probe. **k** Panoramic view of the apomictic genotype reproductive tissues hybridized with antisense *NcoI* N22 probe. **l** and **n** Sexual genotype hybridized with antisense *NcoI* N22 probe (closer view). **m** and **o** Apomictic genotype hybridized with antisense *NcoI* N22 probe (closer view). **p** Sexual genotype hybridized with sense *SalI* N22 probe. No hybridization was detected. **q** Apomictic genotype hybridized with sense *SalI* N22 probe. No hybridization was detected. *Ov* ovule, *Tp* tapetum, *Pmc* pollen mother cells

related elements similar to *gypsy20* retrotransposons, which carried segments of genes already reported to be differentially expressed in apomictic and sexual plants (Albertini et al. 2004; Laspina et al. 2008; Cervigni et al.

2008). Short (15 nt-long) oligonucleotides matching the N22 sequence were used in combination with polyT-anchored primers to perform a retrotransposon-specific differential display experiment on sexual and apomictic plants, in order to identify additional related retrotransposons whose expression could be associated with apomixis or sexuality.

Amplifications conducted in duplicate on the Q4117 (apo) and Q4188 (sex) genotypes produced several apo vs. sexual polymorphic bands. From a total of 107 bands, 48 (44.85%) were polymorphic. Most of these polymorphic bands originated from the sexual genotype (34 bands, 77% of the polymorphic ones). Sequencing of several differential bands revealed that a high percentage of them corresponded to *gypsy* retrotransposons including protein-coding gene segments of variable length. A section of a differential display gel is shown in Fig. 8. Sequence similarity comparison results originated from bioinformatics surveys are summarized in Table 2. Our observations point

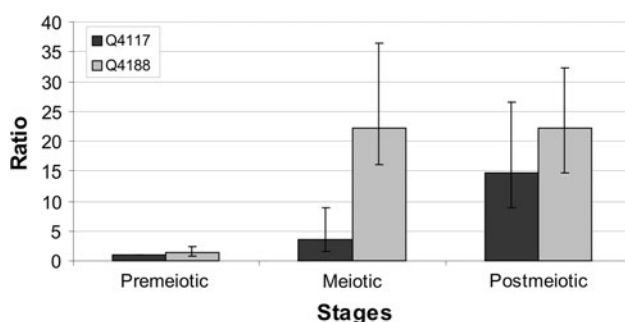


Fig. 7 N22 relative expression ratios in the apomictic (Q4117) and sexual (Q4188). *P. notatum* genotypes at different ovary developmental stages estimated by real-time PCR. All samples and controls were amplified in triplicate

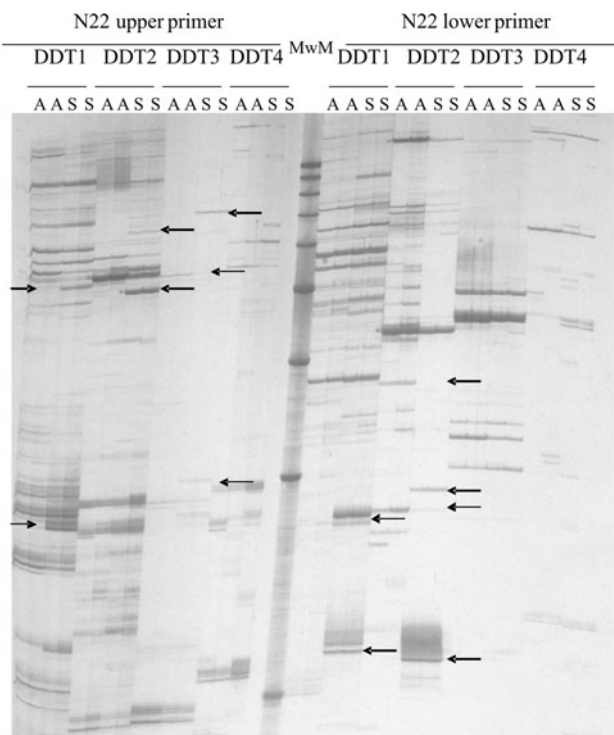


Fig. 8 Retrotransposon-specific differential display. Bands were produced by amplifying in duplicate cDNAs originated from inflorescences of Q4117 (A, apomictic) and Q4188 (S, sexual) plants with polyT-anchored oligonucleotides in combination with specific primers matching retrotransposon N22. Several bands differentially expressed were detected, cloned, and sequenced. The majority of differential bands were expressed in the sexual genotype, but a few bands expressed only in the apomictic plant were detected

toward the existence of a family of related TEs differentially expressed between in the sexual and the apomictic genotype, from which N22 is only a particular member. Family members predominate in the sexual genotype. However, differential display experiments do not allow assessment of the strand type expressed (sense or anti-sense), which should be further investigated through in situ hybridization.

Sequence fragments corresponding to repetitive elements were analyzed with Phylogeny (<http://www.phylogeny.fr/>) in order to produce a tree where sequence similarity among different members could be easily visualized. Two main clusters were detected (Fig. 9). However, the biological significance of this clustering pattern is unknown.

In silico classification of additional unknown sequences associated with apomixis

The remaining 17 unknown sequences originally isolated from differential display experiments by Laspina et al. (2008) were subjected to Blast searches at the TIGR Gramineae Repeats database. Out of the total 17 clones

analyzed, only two showed significant homology to plant repeats (Table 3). In one of them, the occurrence of a conserved short sequence which is part of a functional gene was detected.

Discussion

In the past few years, several studies have been published which identify transcripts that are differentially expressed in reproductive tissues of apomictic and sexual plants (Rodrigues et al. 2003; Albertini et al. 2004; Laspina et al. 2008; Cervigni et al. 2008; Yamada-Akiyama et al. 2009, Polegri et al. 2010). All these articles were focused on the characterization of protein-coding transcripts and presented valuable evidence indicating that central pathways (i.e., ERK signaling, protein degradation, RNA transcription and DNA modification) were differentially modulated in agamospermic plants with respect to sexual ones. Retrotransposons were occasionally listed among differentially expressed sequences (Albertini et al. 2004; Laspina et al. 2008; Cervigni et al. 2008).

In this work, full sequence analysis of two particular transcripts differentially expressed in inflorescences of apomictic and sexual *P. notatum* plants (N17 and N22) revealed that they might be retrotransposons carrying small transduplicated protein-coding gene segments. The first candidate (N17) keeps homology with the LTR region of retrotransposons. Mapping experiments revealed that at least one of its copies is located in a non-recombinant central region of a particular *P. notatum* linkage group (F3c), revealing a centromeric/pericentromeric location. Moreover, N17 was placed close to another *P. notatum* gene differentially expressed in inflorescences of apomictic and sexual plants (kinesin Mc10). The second candidate (N22) displayed low similarity to the gag-coding region of a Rire8 retrotransposon. RACE extension allowed the isolation of several related sequences homologous to *gypsy* retrotransposons, which curiously carried transduplicated segments of apomixis-related genes (*serk*, *cyt P450*).

Real-time PCR quantification showed that both retroelements were represented at an increased number in the genome of the sexual genotype. Interestingly, results obtained from reproductive tissues in situ hybridization indicate that the sense strand of both candidates is up-regulated in immature ovules of sexual plants with respect to apomictic ones. The positive correlation between an increased gene copy number in the genome and a higher concentration of the retroelement sense mRNA strand suggest that de novo insertion might be occurring in the sexual genotype. Recently, the mobilization of various families of endogenous LTR retrotransposons was reported in *Arabidopsis thaliana ddm1* mutants (Tsukahara et al.

Table 2 List of N22-related repetitive elements identified by either differential display (DD) or RACE experiments from the sexual (Q4188) and the apomictic (Q4117) genotypes

Clone	Genotype of origin	Identified by	Element Id	Element description	e-V	Transduplicated gene fragment ID ^a	Transduplicated gene fragment description	e-V
N22	Apo	DD	ORSiTERT00200152	nire8 Ty3-gypsy retrotransposon	0.97	–	–	–
A43	Apo	RACE	gblEU558521.11	Gypsy20 retrotransposon	1e ⁻⁶⁵	CAH56436.1	serk2	3e ⁻²⁸
A44	Apo	RACE	gblEU558521.11	Gypsy20 retrotransposon	8e ⁻³⁵	100191539	LOC100191539 cyt P450	8e ⁻¹⁰⁰
A45	Apo	RACE	gblEU558521.11	Gypsy 20 retrotransposon	1e ⁻¹⁶	–	–	–
A46	Apo	RACE	gblEU558521.11	Gypsy 20 retrotransposon	2e ⁻⁶⁰	CAH56436.1	serk2	2e ⁻⁵⁴
A47	Apo	RACE	ZRSiTERTO00146	dagaf_576C20-1 retrotransposon	3e ⁻⁸	–	–	–
ADD12-1	Sex	TE-specific DD	ORSiTERTO00318	m_128-73 retrotransposon	7e ⁻⁴	8260283	RCOM_01189501	–
ADD13-2	Sex	TE-specific DD	ZRSiTERTO00278	opie_af090447-2 retrotransposon	0.9	NM_001054388.1	Hypothetical protein	2e ⁻⁶¹
ADD16-1	Sex	TE-specific DD	ZRSiTERTO00248	grande_af546188-1 retrotransposon	0.4	NM_001112535.1	LTP-protein	2e ⁻¹¹
ADD17-2	Sex	TE-specific DD	gblAF187823.1	transposase DOPD and transposase DOPA mRNA	0.02	BT086830.1	LTP-protein	2e ⁻¹¹
ADD2-2	Sex	TE-specific DD	gil68445404	transposon nDart1-201 DNA,	0.75	embIIFP097948.11	<i>Phyllostachys edulis</i> mRNA clone	8e ⁻⁹
ADD5-2	Sex	TE-specific DD	ORSiTERT00200157	Ty3-gypsy-like retrotransposon	0.72	–	–	–
ADD6-1	Sex	TE-specific DD	ORSiTERTO00412	retrotransposon	0.37	NM_001149478.1	Hypothetical protein	0.003
ADD7-1	Sex	TE-specific DD	ORSiTERT00200004	Copia type retrotransposon RIRE 7	0.99	AK319168.1	<i>Arabidopsis thaliana</i> mRNA	0.32
ADD7-2	Sex	TE-specific DD	ORSiCMM00100019	LTR-retrotransposon in the centromeric region of rice chromosome 8	0.55	AY108819.2	Hypothetical protein	0.026
ADD8-2	Sex	TE-specific DD	ZRSiTERTO00235	Ji5bl2 retrotransposon	0.99	EU976081.1	Hypothetical protein	2e ⁻⁴
ADD9-1	Sex	TE-specific DD	ORSiTETNO00120	Transposon AnacB1	0.86	NM_001112535.1	LTP-protein, anther-specific protein	2e ⁻¹⁷
ADD10-1	Sex	TE-specific DD	ORSiTERT00200080	Ty3 Gypsy-like retrotransposon	0.99	U90944.1	PDI-like protein	0.99
ADD12-2-2	Sex	TE-specific DD	ORSiTERT00200080	Ty3 Gypsy-like retrotransposon	0.99	–	–	–

N22 was validated to be differentially expressed in the apomictic genotype by real-time PCR

^a Some of the repetitive elements carry transduplicated fragments of plant genes

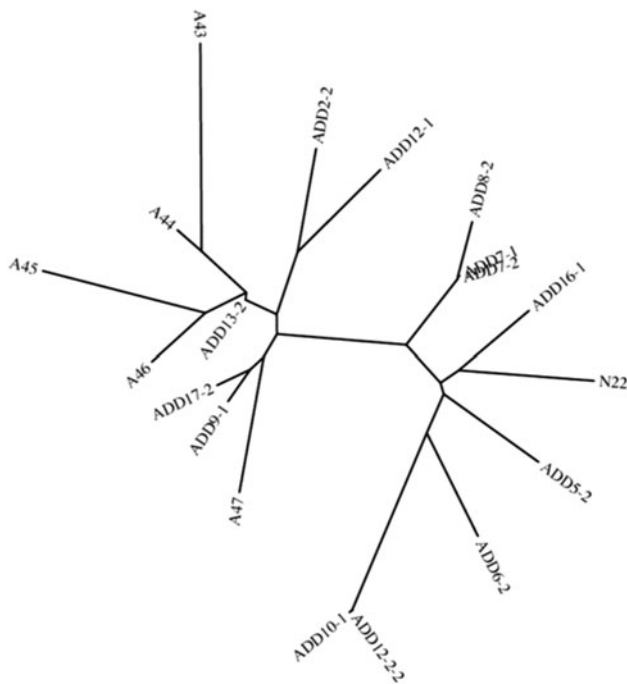


Fig. 9 Phylogeny tree showing clustering of reproduction-associated retrotransposon. Two main groups of differentially expressed retrotransposons, related to *ty3-gypsy* and *gypsy-20*, were detected

2009). A burst of retrotransposition occurred stochastically and independently for each element, suggesting an additional autocatalytic process. Furthermore, comparison of the identified LTR retrotransposons in related *Arabidopsis* species revealed that a lineage-specific burst of retrotransposition of these elements did indeed occur in natural *Arabidopsis* populations. The cDNA methylation state of N17 and N22 should be investigated in order to establish a correlation among epigenetic modification, transcription level, and de novo insertion in the genome of apomictic and sexual plants.

The original N17 and N22 polymorphic bands had been isolated from the apomictic genotype in differential display experiments. However, the allegedly contrasting results between differential display and in situ hybridization experiments could be justified in terms of balancing concentrations of the sense and antisense strand. While in situ hybridization experiments allow discrimination between sense and antisense expression, in differential display experiments a band would be classified as polymorphic without actual knowledge of its origin (sense or antisense).

Moreover, the hybridization detection reactions are usually finished synchronically when at least one of the specimens displays high levels of expression, allowing preferential detection of the most represented strand (sense or antisense). Meanwhile, in differential display polymorphic bands could be originated even from rare sense or antisense transcripts. The evidence presented here indicates: (a) In the sexual genotype, the sense strands of putative retroelements N17 and N22 are strongly up-regulated; (b) In the apomictic plant, a strand of unknown origin (sense or antisense) was detected overexpressed in the differential display experiments. These observations might support the hypothesis that the sense strands are being transcribed at high levels in the sexual genotype, while the antisense strands are being transcribed at moderate or low levels in the apomictic one. If it were so, it would be simple to explain why in the in situ hybridization experiments the sense probe (detecting the antisense strand) never revealed a signal: probably the incubation was concurrently finished for the four specimens (sexual antisense, sexual sense, apomictic antisense, apomictic sense) as soon as the detection of the sense strand produced a considerable signal in the sexual plant.

A retrotransposon-specific differential display aimed at the isolation of N22-related sequences allowed the identification of several repetitive family members, most of them represented in the sexual genotype but some of them in the apomictic one. These results suggest that families of related retrotransposons are differentially expressed in apomictic and sexual genotypes. Interestingly, many of these particular retroelements carry transduplicated gene fragments, some of which had already been associated with apomictic development in previous reports (Albertini et al. 2004; Laspina et al. 2008; Cervigni et al. 2008).

Both class-1 (retroelements) and class-2 (DNA mobile elements) TEs were previously associated with transduplication or readout synthesis of gene fragments (reviewed in Bennetzen 2005). Regarding retroelements, early studies in maize had shown that LTR retrotransposons could acquire specific genic sequences and amplify them across the genome (Jin and Bennetzen 1994). Numerous instances of the acquisition of genes by a closely related mobile DNA, the retroviruses of animals, have also been described (Malik et al. 2000). The transcriptional activity of chimeric LTR retrotransposons including genes fragments was investigated in polyploid wheat and rice (Kashkush et al.

Table 3 Additional sequences differentially expressed in apomictic and sexual genotypes showing homology to repetitive sequences

Clone	Homologous repetitive sequence	eV	Transduplicated gene fragment	eV	Transduplicated gene fragment description
N43	dbj AB014755.1	$6e^{-12}$	–	–	–
N82	ZRSiTERTO00248	$6e^{-5}$	gblEU940865.1	$5e^{-4}$	Zea mays clone 1165244 mRNA sequence

2003, Kashkush and Khasdan 2007). It was proposed that during activation, long terminal repeats drive the readout synthesis of new transcripts from adjacent sequences including the antisense or sense strands of known genes (Kashkush et al. 2003, Kashkush and Khasdan 2007).

Regarding class-2 transposons (DNA mobile elements), transduplication events have been frequently reported for several families. More than 3000 Pack-MULEs (Mutator-like elements containing fragments of genes) were found in the rice genome (Jiang et al. 2004), and at least 5% of them were found to be expressed. More than 90% of these expressed Pack-MULEs appeared to have been transcriptionally initiated within the element itself (Jiang et al. 2004). A further work aimed at the characterization of 8274 rice MULEs showed that 1337 of them contained duplicated host gene fragments (Juretic et al. 2005). Detailed examination of the 5% of duplicated gene fragments that were effectively transcribed demonstrated that virtually all of them contained pseudogenic features, such as fragmented conserved protein domains, frameshifts, and premature stop codons. In addition, the distribution of the ratio of non-synonymous to synonymous amino acid substitution rates for the duplications agreed with the expected distribution for pseudogenes. These results suggested that MULE-mediated host gene duplication results in the formation of pseudogenes, not novel functional protein-coding genes (Juretic et al. 2005). However, the transcribed duplications possessed characteristics consistent with a potential role in the regulation of host gene expression (Juretic et al. 2005). Also helitrons, a class of broadly distributed DNA mobile elements known to transpose via a rolling circle process, often capture gene fragments (reviewed in Bennetzen 2005). Sixty percent of maize helitrons were found to have taken fragments of nuclear genes (Yang and Bennetzen 2009). Most acquired gene fragments are undergoing random drift, but 4% were calculated to be under purifying selection, whereas another 4% exhibit apparent adaptive selection, suggesting beneficial effects for the host or *Helitron* transposition/retention (Yang and Bennetzen 2009).

It has been proposed that chimeric retroelement transcripts including transduplicated gene segments could lead to the creation of novel genes (Jiang et al. 2004, Bennetzen 2005). However, a proven case of gene creation by any transposable element activity in plants remains to be demonstrated (Bennetzen 2005). On the other hand, Juretic et al. (2005) found that all expressed transduplicates with conserved domains have disablements preventing them from encoding functional proteins, and their non-synonymous/synonymous amino acid substitution rates ratios do not indicate that they have been evolving under purifying selection. However, transduplicates constitute a large source of pseudogenes with characteristics suggestive of a

role in regulating host gene expression (Juretic et al. 2005). It was proposed that many chimeric Helitron or Pack-MULE RNAs will not only induce their own epigenetic regulation (e.g., silencing) but also contribute to the epigenetic regulation of the intact genes that have donated gene fragments to the element (Morgante et al. 2005).

A recent article reporting the central role of silencing mechanisms in the determination of the gametic fate in *Arabidopsis thaliana* revealed that retrotransposons might be involved in the regulation of both sexual and asexual reproductive development (Olmedo-Monfil et al. 2010). Inactivation of the *A. thaliana* PAZ-piwi domain-encoding gene *ago9* forced a decrease in the generation of sRNAs in the nucellus, and the concomitant differentiation of several non-reduced enlarged cells expressing gametic markers. These enlarged companion cells often gave origin to additional non-reduced embryo sacs that coexisted with the meiotic megagametophyte within the same ovule. Regarding the generation of multiple non-reduced embryo sacs, *ago9* mutants present a phenotype that resembles apomeiosis of aposporous plant (Olmedo-Monfil et al. 2010). Surprisingly, the sRNA fraction associated with the protein AGO9 was related mainly with retrotransposons and, at a lesser extent, with miRNAs and protein-coding genes. Among the different retrotransposon families processed by *ago9*, the best represented was Gypsy (Olmedo-Monfil et al. 2010). It also was shown that AGO9 can bind in vitro to 24 nt sRNAs corresponding to *Athila* retrotransposons expressed in the ovule prior to pollination (Durán-Figueroa and Vielle-Calzada 2010). AGO9 is necessary to inactivate a significant proportion of long terminal repeat retrotransposons (LTRs) in the ovule. Predominant TE targets are located in the pericentromeric regions of all 5 chromosomes, suggesting a link between the AGO9-dependent sRNA pathway and the heterochromatin formation (Durán-Figueroa and Vielle-Calzada 2010). The results presented here are in agreement with those reported by Olmedo-Monfil et al. (2010) and Durán-Figueroa and Vielle-Calzada (2010). In *ago9* mutants, an absence of retrotransposon processing due to a failure of gene *ago9* originated a phenotype resembling apospory. In natural aposporous plants analyzed here, a down-regulation of N17 and N22-related retrotransposon expression is observed. Both in artificial *Arabidopsis ago9* mutants as well as in aposporous natural plants, the capacity for retrotransposon processing appears to be affected, in the first case because the interference machinery is altered, in the latter because retrotransposons themselves are not being expressed. The effect of these differential regulation on the mechanisms of heterochromatin formation and interference regulation in aposporous and sexual plants should be further investigated.

In this article, we present a preliminary characterization of the activity of a group of retrotransposons during sexual

and apomictic development. Many of these elements include transduplicated copies of plant gene fragments. Our results are consistent with a potential role in the regulation of host gene expression via interference and/or heterochromatin formation, but further experiments are needed to investigate this possibility. Efforts should be made to perform a detailed structural examination of the elements identified herein, as well as at characterizing their chronological expression at different developmental stages in a significant number of plants. Functional analyses of the putative target genes with regard to reproductive modes should also be carried out. Likewise, the generation of plants with altered expression levels of each particular retrotransposon could be used to define what roles they may play during sexual or apomictic reproduction.

Acknowledgments This work was supported by: Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), Ministerio de Ciencia y Tecnología, Argentina, PICT 2007 00476 and PME-2006-03083; Centro Brasil-Argentina de Biotecnología, Ministério da Ciência e Tecnologia, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CBAB-MCT-CNPq), Brasil, projeto 400753/2004-9; Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, PIP 11220090100613; Universidad Nacional de Rosario, proyecto 19G/253 and ProSul (Programa Sudamericano de Apoyo a las Actividades de Cooperación en Ciencia y Tecnología).

References

- Albertini E, Marconi G, Barcaccia G, Raggi L, Falcinelli M (2004) Isolation of candidate genes for apomixis in *Poa pratensis*. *Plant Mol Biol* 56:879–894
- Albertini E, Marconi G, Reale L, Barcaccia G, Procceddu A, Ferranti F, Falcinelli M (2005) SERK and APOSTART. Candidates genes for apomixis in *Poa pratensis*. *Plant Physiol* 138: 2185–2199
- Bashaw EC, Hovin AW, Holt EC (1970) Apomixis, its evolutionary significance and utilization in plant breeding. In: Norman MJT (ed) Proceedings of 11th international grasslands congress. University of Queensland Press, St. Lucia, pp 245–248
- Bennetzen JL (2005) Transposable elements, gene creation and genome rearrangement in flowering plants. *Curr Opin Genet Dev* 15:621–627
- Burton GW (1946) Bahiagrass types. *Agron J* 38:273–281
- Burton GW (1948) The method of reproduction in common Bahiagrass, *Paspalum notatum*. *Agron J* 40:443–452
- Burton GW (1955) Breeding Pensacola bahiagrass *Paspalum notatum*: I. Method of reproduction. *Agron J* 47:311–314
- Burton GW (1967) A search for the origin of Pensacola bahiagrass. *Econ Bot* 21:379–382
- Capron A, Gourgues M, Neiva LS, Faure J-E, Berger F, Pagnussat G, Krishnan A, Alvarez-Mejia C, Vielle-Calzada J-P, Lee Y-R, Liu B, Sundaresan V (2008) Maternal control of male-gamete delivery in *Arabidopsis* involves a putative GPI-anchored protein encoded by the *LORELEI* gene. *Plant Cell* 20:3038–3049
- Cervigni GD, Paniego N, Díaz M, Selva JP, Zappacosta D, Zanazzi D, Landerreche I, Martelotto L, Felitti S, Pessino S, Spangenberg G, Echenique V (2008) Expressed sequence tag analysis and development of associated markers in a near-isogenic plant system of *Eragrostis curvula*. *Plant Mol Biol* 67:7–10
- Chenchic A, Diachenko L, Moqadam F, Tarabykin V, Lukyanov S, Siebert PD (1996) Full-length cDNA cloning determination of mRNA 5' and 3' ends by amplification of adaptor-ligand cDNA. *BioTechniques* 21:526–534
- Crane C (2002) Classification of apomictic mechanisms. In: Savidan Y, Carman JG, Dresselhaus T (eds) The flowering of apomixis: from mechanisms to genetic engineering. CIMMYT, IRD, European Commission, DG VI, (FAIR), pp 24–34
- Daurelio LD, Espinoza F, Quarin CL, Pessino SC (2004) Genetic diversity in sexual diploid and apomictic tetraploid populations of *Paspalum notatum* situated in sympatry or allopatry. *Plant Syst Evol* 244:189–199
- Duarte Silveira E, Alves-Ferreira M, Arrais Guimarães L, Rodrigues da Silva F, Tavares de Campos Carneiro V (2009) Selection of reference genes for quantitative real-time PCR expression studies in the apomictic and sexual grass *Brachiaria brizantha*. *BMC Plant Biol* 9:84. doi:10.1186/1471-2229-9-84
- Durán-Figueroa N, Vielle-Calzada JP (2010) ARGONAUTE9-dependent silencing of transposable elements in pericentromeric regions of *Arabidopsis*. *Plant Signal Behav* 6(5):11
- Jarret RL, Ozias-Akins P, Phatak S, Nadimpalli R, Duncan R, Hiliard S (1995) DNA contents in *Paspalum* spp. determined by flow cytometry. *Genet Resour Crop Evol* 42:242–273
- Jiang N, Bao Z, Zhang X, Eddy SR, Wessler SR (2004) Pack-MULE transposable elements mediate gene evolution in plants. *Nature* 431:569–573
- Jin Y-K, Bennetzen JL (1994) Integration and nonrandom mutation of a plasma membrane proton ATPase gene fragment within the Bsl retroelement of maize. *Plant Cell* 6:1177–1186
- Juretic N, Hoen DR, Huynh ML, Harrison PM, Bureau TE (2005) The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res* 15:1292–1297
- Kashkush K, Khasdan V (2007) Large-scale survey of cytosine methylation of retrotransposons and the impact of readout transcription from long terminal repeats on expression of adjacent rice genes. *Genetics* 177:1975–1985
- Kashkush K, Feldman M, Levy AA (2003) Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet* 33:102–106
- Kosambi DD (1944) The estimation of map distance from recombination values. *Ann Eugen* 12:172–175
- Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1:174–181
- Laspina NV, Vega T, Martelotto L, Stein J, Podio M, Ortiz JP, Echenique V, Quarin C, Pessino SC (2008) Gene expression analysis at the onset of aposporous apomixes in immature inflorescences of *Paspalum notatum*. *Plant Mol Biol* 67:615–628
- Liang P, Pardee AB (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science* 257:967–971
- Malik HS, Henikoff S, Eickbush TH (2000) Poised for contagion: evolutionary origins of the infectious abilities of invertebrate retroviruses. *Genome Res* 10:1307–1318
- Martínez EJ, Hopp E, Stein J, Ortiz JPA, Quarin CL (2003) Genetic characterization of apospory in tetraploid *Paspalum notatum* based on the identification of linked molecular markers. *Mol Breed* 12(4):319–327
- Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat Genet* 37:997–1002

- Nogler GA (1984) Gametophytic apomixis. In: Johri BM (ed) Embriology of angiosperms. Springer, New York, pp 475–518
- Olmedo-Monfil V, Durán-Figueroa N, Arteaga-Vázquez M, Demesa-Arévalo E, Aufran D, Grimaneli D, Slotkin RK, Martienssen RA, Vielle-Calzada JP (2010) Control of female gamete formation by a small RNA pathway in *Arabidopsis*. *Nature* 464:628–632. doi:10.1038/nature08828
- Ortiz JP, Pessino SC, Leblanc B, Hayward MD, Quarin CL (1997) Genetic fingerprint for determinig the mode of reproduction in *Paspalum notatum*, a subtropical apomictic forage grass. *Theor Appl Genet* 95:850–856
- Ouyang S, Buell CR (2004) The TIGR plant repeat databases: a collective resource for the identification of repetitive sequences in plants. *Nucleic Acids Res* 32:D360–D363
- Pessino S, Martelotto L (2006) Genome structure and gene expression in polyploid plants. In: Teixeira da Silva JA (ed) Floriculture, ornamentals and plant biotechnology: advances and topical issues, 1st edn. Global Science Books, London, UK
- 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. *J Exp Bot* 61:1869–1883
- Quarin CL, Norrmann GA, Urbani MH (1989) Polyploidization in aposporous *Paspalum* species. *Apomixis Newsletter* 1:28–29
- Quarin CL, Espinoza F, Martinez EJ, Pessino SC, Novo OA (2001) A rice of ploidy level induces the expression of apomixis in *Paspalum notatum*. *Sex Plant Reprod* 13:243–249
- Quarin CL, Urbani MH, Blount AR, Martinez EJ, Hack CM, Burton GW, Quesenberry KH (2003) Registration of Q4188 and Q4205, sexual tetraploid germplasm lines of Bahiagrass. *Crop Sci* 43:745–746
- Rodrigues JC, Cabral GB, Dusi DMA, Mello LV, Rinden D, Carneiro VTC (2003) Identification of differentially expressed cDNA sequences in ovaries of sexual and apomictic plants of *Brachiaria brizantha*. *Plant Mol Biol* 53:745–757
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning. A laboratory manual, 2nd edn. Cold Spring Harbour Press, New York
- Shagai-Marooof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci USA* 81:8014–8018
- Spillane C, Curtis MD, Grossniklaus U (2004) Apomixis technology development—virgin births in farmers’ fields? *Nat Biotechnol* 22:687–691
- Stein J, Pessino SC, Martinez E, Quarin CL, Ortiz JP (2004) Tetraploid races of *Paspalum notatum* showed polysomic inheritance and preferential chromosome pairing around the apospory-controlling locus. *Theor Appl Genet* 109:186–191
- Stein J, Pessino S, Martinez E, Rodríguez MP, Siena L, Quarin C, Ortiz JP (2007) A genetic map of tetraploid *Paspalum notatum* Flügge (bahiagrass) based on single-dose molecular markers. *Mol Breeding* 20:153–166
- Tischler CR, Burson BL (1995) Evaluating different bahiagrass cytotypes for heat tolerance and leaf epicuticular wax content. *Euphytica* 84:229–235
- Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T (2009) Burst of retrotransposition reproduced in *Arabidopsis*. *Nature* 461:423–426
- Yamada-Akiyama H, Akiyama Y, Ebina M, Xu Q, Tsuruta S, Yazaki J, Kishimoto N, Kikuchi S, Takahara M, Takamizo T, Sugita S, Nakagawa H (2009) Analysis of expressed sequence tags in apomictic Guinea grass (*Panicum maximum*). *J Plant Physiol* 166:750–761
- Yang L, Bennetzen JL (2009) Distribution, diversity, evolution, and survival of Helitrons in the maize genome. *PNAS* 106(47): 19922–19927