

A new method for the design of degenerate primers and its use to identify homologues of apomixis — associated genes in *Brachia*

Eduardo Gorrón^{1,2}, Fausto Rodríguez¹, Diana Bernal¹, Silvia Restrepo² and Joe Tohme^{1,*}

¹Centro Internacional de Agricultura Tropical (CIAT), A.A. 6713, Cali, Colombia

²Universidad de los Andes, Carrera 1 No 18A-10, Bogotá, Colombia

*Corresponding author, j.tohme@cgiar.org

Abstract: Apomixis is a reproductive phenomenon that occurs in flowering plants. It allows a plant to produce asexual seeds, with its same genetic constitution. The existence of a genetic basis for apomixis is crushing, but the molecular mechanisms are unclear. The search for the “master apomixis gene” had led to the isolation of various candidate transcripts, but neither of them could be confirmed in different plant species. Here we tried to isolate homologues to all those transcripts in one unique plant, *Brachiaria*. In order to achieve this, a new method for degenerate primer design was employed, since classical methods have proven to be unsuccessful. We used multiple local alignments, instead of global, with the Multiple Expectation – Maximization for Motif Elicitation (MEME) algorithm, to find conserved blocks and motifs. These alignments were followed by ePCR simulation and standard primer pair design programs. The method demonstrated to be useful to amplify fragments homologous to genes poorly molecular and biologically characterized, with which multiple global alignments showed no conserved regions. The obtained amplicons showed differential expression according to tissue in some cases. This technique can be used to design degenerate primers in cases where one sequence exhibits poor global similarity and has low biological characterization, and it is useful to amplify orthologous genes in an organism weakly described at the molecular level.

Keywords: Primer design, apospory, degenerate primers, MEME

1 INTRODUCTION

The development of degenerate primer pairs often involves the amplification of orthologous DNA fragments which have little conservation, even in the same taxonomic group. Therefore, typical techniques used to design them are unable to find the conserved blocks necessary to generate the oligonucleotides. One reason for this is the employment of multiple global alignment algorithms as the starting point to find the blocks [GMDK05, KCSW94]. The use of this kind of algorithms demands a relatively high degree of conservation along the entire sequence and between all the aligned sequences. If the analyzed

DNA region lacks it, we are forced to look for specific primers for each genera if we try to amplify it in different members of a taxa, or to find alternate ways to do that.

We faced this problem in our laboratory in the analysis of sequences related to a phenomenon called apomixis. Apomixis consists in the ability of some flowering plants to produce viable seeds by asexual ways. It has two basic characteristics: the avoiding of meiosis or the degeneration of meiotic-derived cells, a situation known as apomeiosis, and the generation of a megagametophyte genetically equal to the surrounding tissues [AJ92]. Many observations in different apomictic plants [AJ92, BBK00, KG03, OARH98, POL⁺97, Mat89] has led to the conclusion that apomixis has a very strong genetic component, and this phenomenon is regulated by one or a few genes [AHOA05, AMR⁺05, LBC⁺02, PMB⁺04, PEO⁺98]. Some previous works have found some candidate transcripts [CCA⁺99, GRR⁺00, LAP⁺97, PEM⁺01, RCD⁺03, VCNB⁺96], but neither could confirm clearly any of them. For these reasons, trying to analyze expression patterns of candidate transcripts in one plant species is very important, as a first step to obtain a general and clearer molecular model. To do this, we tried to design degenerate primers from many reported sequences related to apomixis in different plant species, in order to obtain all their orthologous counterparts in the genus *Brachiaria*. However, many of the reported transcripts have not obvious molecular functions and, in fact, are sequences with unknown function in many cases. There are not reference sequences with high similarities with them. Multiple global alignments did not show clear conserved regions. Hence, primer design, even degenerate primer design, was not possible with the standard bioinformatic techniques [GMDK05, KCSW94].

In this work, we propose the use of a new method to design primers, which could be used to create oligonucleotides for sequences with poor global similarity to their suspected homologues or when they don't show well conserved blocks in multiple global alignments. The technique is based in the use of multiple local alignments instead of global, process carried with the Multiple Expectation – Maximization for Motif Elicitation (MEME) algorithm. This algorithm was initially suited to look for conserved motifs in protein and DNA sequences, in order to identify possibly functional homologies between them. In our study we used it to search for conserved regions long enough to design an acceptable primer pair. The result of MEME is powered with a confirmation with electronic PCR (e-PCR) over the list of sequences used and a verification of annealing temperatures, secondary structure formation and primer dimers with standard programs. With it, amplification of homologues of one gene in a related species, which is poorly characterized at the molecular level, is possible, as we confirmed this with laboratory assays.

2 MATERIALS AND METHODS

Obtention of sequences related to apomixis. Sequences associated to the apomictic trait were retrieved from the GenBank database, after a literature review about expression analysis and candidate genes in different apomictic plant species. The original GenBank accession number was used to reference the obtained primer pairs and the respective results.

Construction of a database for BLAST search and BLAST analysis. In order to look for homologues of these transcripts in relative species of Brachiaria, we used the TIGR databases TIGR Gene Indices (<http://compbio.dfci.harvard.edu/tgi/>) and TIGR Plant Transcript Assemblies (<http://plantta.tigr.org>, both consulted on September 2007). The assemblies from all species included in the families Alliaceae and Poaceae were downloaded, to construct a database of EST assemblies specific for relatives of Brachiaria. The sequences obtained from GenBank as associated to apomixis were used as queries in searches against this database using the BLAST algorithm. The maximum accepted e-value in this test was 10⁻⁶; sequences without homology under this value were discarded.

MEME alignment. Each sequence and the group of sequences similar to it according to BLAST were aligned, initially using T-COFFEE (<http://tcoffee.vital-it.ch/cgi-bin/Tcoffee/tcoffee.cgi/index.cgi?stage1=1&daction=TCOFFEE::Regular&referer0=embnet>) to see global similarity, and then with MEME (http://meme.sdsc.edu/meme4_1/cgi-bin/meme.cgi) [BWML06]. The conditions of this last alignment included a minimum length of 18 nt and a maximum of 25, and minsites = 2/3 of the sum of sequences, conditions that was considered optimal for primer design. In order to maintain a relatively equal representation of different species, when some plant species predominated in BLAST results (for example rice), some of its associated sequences were eliminated. Sequences with low similarity in MEME alignment were also deleted. With the remaining sequences, a second round of MEME was made.

Construction of motif pairs and e-PCR simulation. The 20 first motifs obtained in each MEME run were extracted and ordered by their score given in the program. All possible pairs between them were virtually assembled, creating all the combinations of motifs, in order to test each pair as primer. Every pair was run over the group of corresponding sequences (the sequences used in the second MEME alignment) in the e-PCR program [Sch97] under the condition that amplicon size must be at least 200 bp in length.

Final selection of primer pairs and synthesis. Motif pairs, which show an adequate amplicon size in e-PCR, were manually verified in their corresponding MEME alignment. First, the motif pair with the highest score was evaluated. The presence of the original query sequence and the number of degenerations was verified. If the query sequence was absent in motif alignments and/or the number of degenerations was greater than 12 between both motifs, the pair was discarded and the test continued with the next pair. If we saw some pattern associated with taxa (for example an A shared between all members of Panicoideae) in a degenerate position, the nucleotide present in Panicoideae was left, and no degeneration was considered. When one motif pair was acceptable, it was tested in NetPrimer (<http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html>), in order to verify annealing temperatures, second structure formation and possible primer dimers. If the pair had undesirable T_m, a T_m difference greater than 5°C, or formed dimers or secondary structures with high ΔG (over

-8KJ/mol), bases of one or both motifs were eliminated, until the conditions were acceptable. The minimum size allowed for each motif was 17 nt. If the motif pair, even with these adjustments, did not show to be adequate, the pair was discarded and the test continued with the next pair. Finally, when one pair was accepted, the degenerations were put in their positions and it was synthesized as degenerate primers (Integrated DNA Technologies, Coralville, Iowa). If neither motif pair had the required conditions, the analysis for this particular sequence stopped.

Plant material. *Brachiaria decumbens* accession 16494 (CIAT code) was employed. Pistils of 1,7-2,2 mm in length were collected with all RNA extraction cautions, and they were deposited in 600 μ l of RNAlater (Ambion Inc, Austin, Texas). A total of 200 pistils were taken. This sample was maintained in -80°C until RNA extraction. Leaves, roots and stems of plants were also taken.

RNA extraction and cDNA synthesis. The pistils sample was centrifuged at 14000rpm by at least 20 minutes to allow deposit of pistils at the bottom. RNAlater solution was discarded and pistils were macerated in liquid nitrogen. Total RNA was extracted using Picopure RNA Isolation Kit (Molecular Devices, Sunnyvale, CA), following manufacturer's instructions. Leaves, roots and stems were also macerated in liquid nitrogen and total RNA was extracted using Trizol (GibcoBRL, Carlsbad, CA). cDNA of all type of samples were synthesized using the Creator SMART cDNA Library Construction Kit (Clontech, Mountain View, CA).

Amplification with the designed degenerate primers. PCR reactions were carried with cDNA from pistils (reproductive tissue), in one hand, and a bulk of cDNA from leaves, roots and stems (vegetative tissue), on the other hand, both diluted at 5ng/ μ l. PCR reactions were initially tested with the next mix: Tris-HCl 20mM, KCl 50mM, MgCl₂ 1,4mM, dNTPs 0,2mM each one, 1 U Taq Polimerase, each primer at 0,6 μ M, 5 μ l of diluted DNA, final volume 25 μ l. PCR cycling program was 94°C 2 minutes, followed by 40 cycles of 94°C 30 seconds, the theoretical annealing temperature of each primer pair by 30 seconds, and 72°C 1 minute 30 seconds, and finally 72°C 5 minutes. If amplification was not observed under these conditions for a particular primer pair, the DNA concentration was increased, making a lower DNA dilution. When more DNA did not improve results, the annealing temperature in the program was reduced another 5°C. The primer pairs that did not amplify with these changes were not assayed again.

Sequencing of resulting amplicons. Amplicons obtained in the PCR reactions were cloned using pGEM-T Easy Cloning Kit (Promega, Madison, WI) and sequenced with ABI PRISM Big Dye kit (Applied Biosystems, Foster City, CA). Every different amplicon was sequenced three times. Sequences were edited and analyzed using Sequencher 4,8 (Gene Codes Corporation, Ann Arbor, MI).

Bioinformatic analysis of obtained sequences. The obtained sequences were analyzed using BLAST against the GenBank database. Each sequence was compared against its original query, obtained initially from GenBank, using BLAST2Seqs (<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>) (blastn algorithm) and LALIGN (http://www.ch.embnet.org/software/LALIGN_form.html) (global alignment, the other parameters by default). This comparison was also made when we obtained, with the same primer pair, amplicons of different sizes when the PCR was made in reproductive or vegetative tissue.

3 RESULTS

Sequences obtained from GenBank database. After the literature review, only 27 sequences associated to apomixis in GenBank were obtained (Table 1). Many of the sequences are short sequences (200-400 bp) and are not highly similar to known sequences. Two sequences from *Saccharomyces cerevisiae* were also included; their corresponding two genes were reported as linked with a yeast meiotic phenomenon also called apomixis [BMG83]. When BLAST analysis was made, only 11 of the 27 sequences showed results with similarities with e-value under 10-25. For this reason, the parameters of algorithm were changed and similarities with e-value under 10-20 were accepted. This allowed the inclusion of another 5 sequences in the analysis. With a final change in the limits of BLAST search to tolerate results with similarities with e-value less than 10-6, nine more sequences were included, and three sequences (u65386, u65387 and yscspoa) did not show any acceptable results and were no more analyzed. For the two sequences of *S. cerevisiae*, BLAST was made in GenBank database too and plant results were also included in further examination.

Multiple alignments. Many of the sequences did not show very high similarity with their respective BLAST results. So, as we expected, T-COFFEE alignments showed no satisfactory global similarity in any case (data not shown). MEME alignments, however, were capable to reveal motifs with good levels of conservation. More than half of the results exhibited a common arrangement of motifs between the analyzed sequences, indicating a global similarity that could not be well resolved with common algorithms of multiple global alignments (Figure 1a). This was a good precedent for primer design, because it proved the existence of well-conserved regions and similarity in these cases. Nevertheless, 5 of the remaining 24 query sequences did not have a satisfactory similarity in its alignment, especially because the query did not share motifs with the others (Figure 1b). The first eleven motifs have an acceptable level of degeneracy, and they were preferred to make subsequent scrutiny.

e-PCR and primer design. The MEME motifs were virtually combined in pairs and assayed in electronic PCR against every set of sequences (the query plus their BLAST results). An average between 30 and 50 pairs of motifs had an amplicon of at least 200 bp,

| GenBank code | Name |
|--------------|--|
| AB000809 | <i>Panicum maximum</i> A2-134 mRNA, complete cds |
| AF242539 | <i>Paspalum notatum</i> clone apo417 apomixis-related protein 1 mRNA, partial cds |
| AF475105 | <i>Triticum aestivum</i> apomixis-associated protein mRNA, partial cds |
| AJ786393 | <i>Poa pratensis</i> APOSTART2 gene for START domain-containing protein, exons 1-22 |
| AJ810708 | <i>Poa pratensis</i> mRNA for RAB1-like (rab1-2 gene) |
| AJ810709 | <i>Poa pratensis</i> mRNA for Ankyrin protein kinase-like (apk gene) |
| AJ810710 | <i>Poa pratensis</i> mRNA for Armadillo-like (arm gene) |
| AJ841698 | <i>Poa pratensis</i> serk1 gene for somatic embryogenesis receptor-like kinase 1, exons 1-11, allele 1. |
| AY375366 | <i>Pennisetum squamulatum</i> Opie-2-like retrotransposon, partial sequence |
| D37938 | <i>Pennisetum ciliare</i> apomixis-associated mRNA, clone:psb C |
| D37939 | <i>Pennisetum ciliare</i> apomixis-associated mRNA, clone:psb3-1a |
| D37940 | <i>Pennisetum ciliare</i> apomixis-associated mRNA, clone:psb H. |
| EF517497 | <i>Cenchrus ciliaris</i> apomixis-related protein Pca21 (Pca21) mRNA, complete cds |
| EF517498 | <i>Cenchrus ciliaris</i> apomixis-related protein Pca24 (Pca24) mRNA, partial cds |
| EF530198 | <i>Hieracium caespitosum</i> DMC1 gene, complete cds |
| EF530199 | <i>Hieracium piloselloides</i> putative ubiquitin associated/T5-N domain-containing protein (UBA) mRNA, complete cds |
| M32653 | <i>S. cerevisiae</i> sporulation protein (SPO16 and SPO12) genes, complete cds |
| M38357 | <i>Saccharomyces cerevisiae</i> meiosis-specific protein (SPO13) gene, complete cds, clone p(SPO13)1 |
| U40219 | <i>Pennisetum ciliare</i> possible apospory-associated protein mRNA, complete cds |
| U65082 | PCU65382 Buffelgrass obligatory sexual ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65383 | PCU65383 Buffelgrass obligatory sexual ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65384 | PCU65384 Buffelgrass obligatory sexual ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65385 | PCU65385 Buffelgrass obligatory sexual ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65386 | PCU65386 Buffelgrass obligatory apomictic ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65387 | PCU65387 Buffelgrass obligatory apomictic ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65388 | PCU65388 Buffelgrass obligatory apomictic ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |
| U65389 | PCU65389 Buffelgrass obligatory apomictic ovary (M.A.Hussey) <i>Cenchrus ciliaris</i> cDNA, mRNA sequence |

Table 1: Sequences retrieved from GenBank and used for degenerate primer design

which were organized by their score. The MEME alignment of every pair, starting with the pair with the highest score, was manually revised. We saw some cases in which one tribe or genus has one nucleotide and the others had another nucleotide, indicating a taxonomic pattern of the degeneracy (Figure 2). In that case, that degeneracy was not considered and the nucleotide shared by members of the tribe Panicoideae, to which Brachiaria belong, is putted in that position.

Many motif pairs, however, were discarded in the verification step with NetPrimer, because their Tms were very different and/or the predicted dimers had ΔG over 8KJ/mol. So, the finding of the adequate motif pair took a long time. This manual process of verification was particularly very laborious and time consuming.

Finally, primer pairs were designed for 22 sequences (Table 2), and for another two this was impossible. The degeneracies given by MEME were put in their corresponding posi-

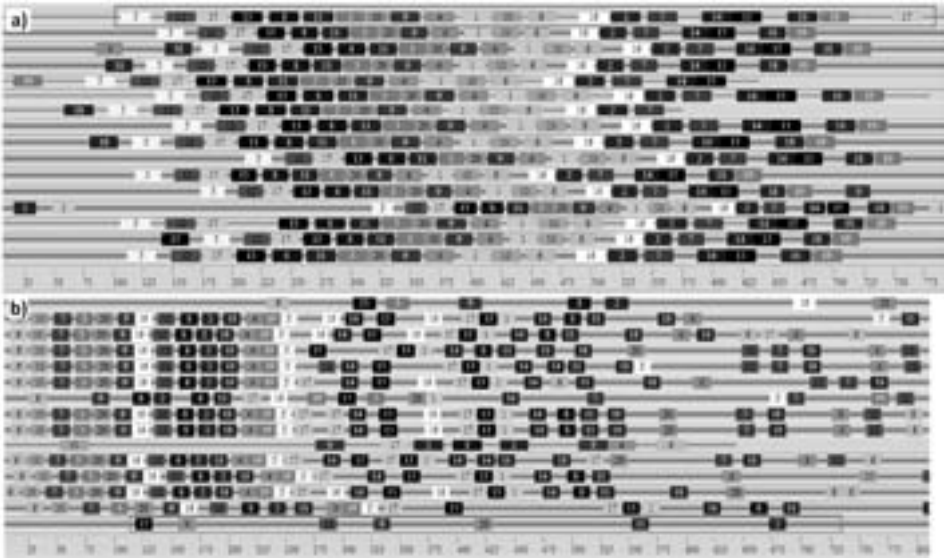


Figure 1: MEME general results (motif general arrangement) for a) the sequence aj810708.1, which serves as an example of MEME results, and b) for the sequence yscspo13. In the red rectangle, the original query sequence is shown; the other sequences, which were taken from its corresponding BLAST results, are represented below it. Every numbered color box represents a specific motif. MEME general motif arrangement results for the sequence yscspo13. In the red rectangle, the original query sequence is shown; the other sequences, which were taken from its corresponding BLAST results, are represented over it. Every numbered color box represents a specific motif.

tions. The strict process of selection followed in this analysis allowed primers to have 120 of degeneracy as maximum. For each primer pair, the accession number of the original query sequence was left as name.

PCR assays in vegetative and reproductive tissues. PCR reactions were assayed using the cDNA of pistils or vegetative tissue (leaves, roots and stems). The length of the pistils was chosen in basis of previous observations [AMP⁺00, DW99]; in them, the embryo sac formation at this specific pistil size is reported. So, if the query sequences are specific of a process related to reproductive development, there is more chance to obtain them if we assay this specific part and moment of maturity.

Initially, we tried with 25 ng of starting cDNA as template, but neither of the primers showed positive results. It could imply that all amplified transcripts have low expression levels. So, we had to work with more concentrated DNA. A ten-dilution fold of the original cDNA gave amplification of 12 primer pairs in vegetative tissues (ab000809, af475105, aj810709, aj841698, penpsbca, penps31ab, ef517497, ef517498, ef530198, yscspo13, pcu40219, u65082) and only six in pistils (ab000809, aj810709, aj841698, ef517497, ef517498, yscspo13). As we expected always the presence of the product in reproductive tissue, we assayed again with a five-dilution fold. In this case, one primer (penps-

MOTIF 6 width = 25 sites = 8 llr = 236 E-value = 1.9e-039

| | | | | <u>Multilevel</u> <u>consensus</u> <u>sequence</u> | | |
|--------------------|-------|----------|------------|--|------------|---|
| | | | | GAGGCACGGAAATTCATCAAAGGGT | | |
| | | | | T | C | A |
| NAME | START | P-VALUE | SITES | | | |
| gnl zmgj TC339781 | 453 | 1.29e-15 | CACACCAGAA | GAGGCACGGAAATTCATCAAAGGGT | ACTCTGAAAG | |
| gnl zmgj TC345176 | 423 | 1.29e-15 | CACACCAGAA | GAGGCACGGAAATTCATCAAAGGGT | ACTCTGAAAG | |
| gnl sbgi TC99813 | 434 | 1.29e-15 | CACTCCAGAA | GAGGCACGGAAATTCATCAAAGGGT | ACTCTGAAAG | |
| gnl hrvgj TC141784 | 364 | 9.42e-15 | CACTCCAGAG | GAGGCACGCTAAATTCATCCAAGGAT | ATTCCCAAAG | |
| gnl tagj TC238321 | 341 | 9.42e-15 | CACTCCAGAG | GAGGCACGCTAAATTCATCCAAGGAT | ATTCCCAAAG | |
| gnl tagj TC238322 | 155 | 9.42e-15 | CACTCCAGAG | GAGGCACGCTAAATTCATCCAAGGAT | ATTCCCAAAG | |
| U65082 | 15 | 1.90e-14 | CACTCCAGAA | GAAGCACGGAAATTCATCAAAGGGT | ACTCTGAAAG | |
| gnl soji CA271065 | 290 | 1.79e-11 | CAACTCAAAA | GGGCCCGGGGAATCCTCCAAGGGT | TCITTTAAAA | |

Figure 2: MEME alignment for motif #6 for the sequence u65082, as an example of alignments with taxonomic patterns. In green rectangles, the sequences of Panicoideae are shown; in the red rectangle there are sequences of Pooideae. Two letters followed by gi are employed to annotate the species associated with a particular sequence: zm = Zea mays; sb = Sorghum bicolor; so = Saccharum officinarum, hv = Hordeum vulgare; ta = Triticum aestivum. In this example, in the three degeneracies, the 9th base is always G in Panicoideae and T in Pooideae; the 19th is A in Panicoideae and C in Pooideae, and the 24th is G in Panicoideae and A in Pooideae; so, the consensus considered in this case don't include the T, A and A of Pooideae and the considered motif was GAGGCACGGAAATTCATCAAAGGGT.

bca) showed amplification in reproductive tissue and another primer pair in both samples (aj810708). Finally, we employed cDNA directly for the remaining ones, allowing positive results in pistils for the other five primer pairs that amplified only in vegetative tissue in 1:10 cDNA dilutions and another 2 primer pairs (aj786393, u65384) in both tissues. This may mean that some sequences have greater expression levels in vegetative tissues than in reproductive ones. Seven primer pairs (aj810710, penpsbhc, ef530199, u65383, u65385, u65388 and u65389) never gave amplicons under neither of the assayed conditions.

Many of the obtained products had no more than 300bp in size (Figure 3). This is in concordance with the amplicon predicted by ePCR. In many cases, the amplicons obtained from vegetative and reproductive tissue were undistinguishable. However, 4 primer pairs (aj786393, aj810708, penpsbca and ef517498) gave products of different size according to the tissue assayed; the fragment in these situations was always of larger molecular weight in pistils. Events of specific tissue expression and differential splicing are plausible.

One very interesting result about these reactions is the fact that almost neither primer pair gave multiple bands, as could happen if the primer is highly degenerate. In most cases, a unique band was observed. The primer pair aj786393 is the only one that had multiple bands. But even that, a predominant band is observed, not only in the vegetative bulk but also in pistils. These bands could be obtained apart in the cloning tests (data not shown).

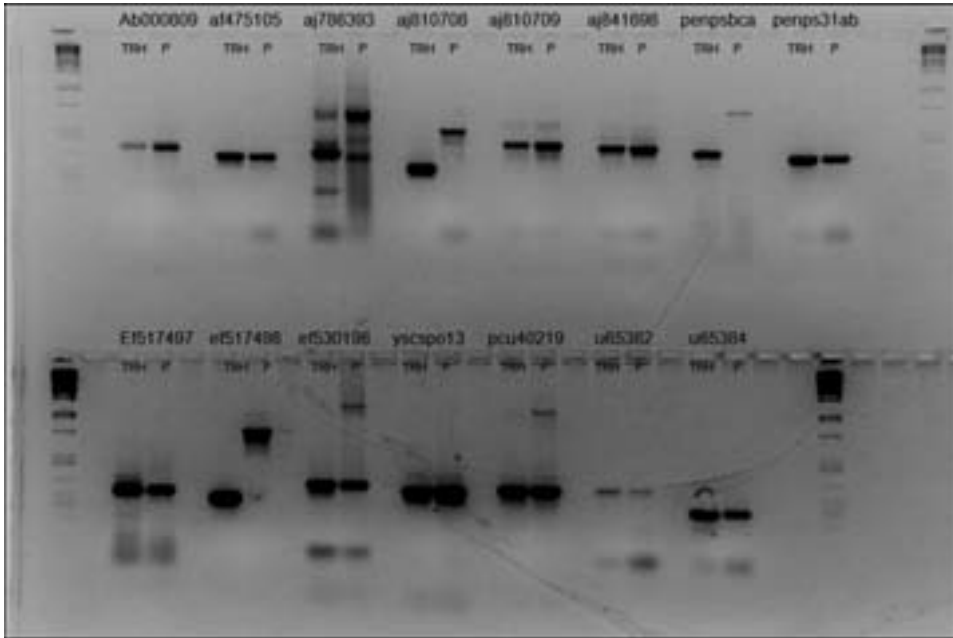


Figure 3: PCR products obtained for the 5 primer pairs that gave positive results. The name of each primer pair (equal to the accession number of the original query sequence) is shown above its corresponding result. The TRH title over one well indicates that the product there is from vegetative tissue, and a P, from pistils. Ladder: 1 kb DNA ladder (Invitrogen).

Bioinformatic analysis of sequenced amplicons. The PCR products were cloned and sequenced. Comparisons between the obtained sequences and the original query sequences using BLAST2SEQS and LALIGN showed high similarity including the entire acquired product and covering at least part of the query in all cases, except three. The query also appeared as heading result when BLAST search was made in these cases.

The three cases that did not show the expected similarity with queries were yscspo13, ab000809 and ef517498. The BLAST results of yscspo13 were overall sequences of the family of N-acetyltransferases, having similarities above 95% over the entire query. Ab000809 sequence matched with a group of functionally uncharacterized transcripts with high resemblance with threonin aldolase gene. Ef517498 is 88% similar to diverse transcripts of *Oryza sativa* related to the DMC1 gene in *S. cerevisiae*, a gene specific to meiotic events in that yeast.

Comparisons between the amplicons of different tissues were made. In the situations where both amplicons had the same molecular weight, the sequences were also the same. In the four cases when the PCR product showed different size, the sequence from vegetative tissue was always longer than from pistils. This supports the idea of differential splicing according to tissue, indicating a high spatial and, possibly, functional specificity.

4 DISCUSSION

In the present study, we developed a new method to design degenerate primers to target specific genes in *Brachiaria*. In contrast to earlier approaches, the method employed here does not require to start from large conserved blocks, extracted from multiple global alignments. Instead, it employs directly conserved motifs obtained from multiple local alignments. This allows us to work with poorly characterized sequences at the biological level and with few or neither homologues previously verified. The MEME algorithm [BWML06] was directly employed for this purpose. To the best of our knowledge, this is the first time that this program is used in this way. Also and for the first time, MEME was combined with ePCR (electronic PCR) [Sch97], which served as an initial filter for the motif pairs. On whole, the method developed in this study combines the use of bioinformatic programs which were not integrated before, and with a goal for which they were not originally made.

Our approach amplified finally and correctly 12 sequences (44% with respect to the 27 original sequences and 54% with respect to the 22 sequences for which primers could be made). These results are considered very positive, not only because the low knowledge about them, but also because many times we started from few similar sequences.

The three false positive results and, in general terms, the cases where an amplicon cannot be obtained, were associated to low similarity patterns in MEME analysis. As a consequence, the respective genetic family could not be detected or one family, different from the target group, was amplified, producing unspecific results. Another consequence of using these patterns is the obtention of high degeneracy levels (ej. *penpsbhc*, see Table 1). In some situations (*u65082*, *u65383* and *u65384*), the small size of the fragments made difficult to find similar sequences, and this caused low similarity arrangements. Patterns like *yscspo13* (Figure 1b), and in general where some conservation in the motif arrangement cannot be seen, should be avoided as much as possible, in order to increase the successfulness of the technique.

Excluding these cases, the method reported here showed good results. This is reflected in a low difference in the T_m of the primer pairs and between theoretical and experimental T_m , and in low degeneracy levels. The careful selection process of the degeneracies, which took into account taxonomic relationships, contributed to the favorable results obtained. It is desirable to look for ways to automatize and optimize this phase. Bioinformatic analysis could verify the amplification of the correct homologues. Thus, in general terms, the method has a good efficiency, which could be improved in further studies being more rigorous with the differences in T_m , the length of the sequences and the local and global similarity seen.

Qualitative differences in the expression of some of the evaluated transcripts could be observed, when sequences from somatic and reproductive tissues were compared. The loss of DNA segments could suggest specific differential splicing events. This suggests the existence of factors which alter mRNA in a tissue – specific manner, a possibility that should be deeply explored in *Brachiaria*. In events like this and in differential splicing, part of the explanation to apomixis could be found. So, the long sequence (from pistils)

could have a specific role in the reproductive development, and it must be analyzed in major detail.

In conclusion, we could develop a new bioinformatic method which allows us to amplify homologous genes associated to a very poorly characterized phenomenon at the molecular level like apomixis. The obtention of these sequences by this method is a very important step in order to establish a clear and concrete molecular model of apospory, in this case in the plant genus *Brachiaria*. This method could help to make amplifications in other poorly understood biological events, in which there are few related sequences. In general terms, the technique can be used for sequences that have very few known homologues, or to confirm them, and to design degenerate primers when the classical methods do not work. The method, however, needs to be improved, for example with the automatization of some time consuming steps and the avoiding of patterns of low similarity. The results obtained here point to possible events related to differential splicing that could help to explain this very interesting trait. Our results will allow the analysis of all those proposed candidate genes in a unique plant species. In additional studies, the differences in expression and functional characteristics of the obtained sequences must be evaluated, like microarray analysis and real-time PCR. These sequences are being included in microarray analysis in our laboratory.

References

- [ACA01] E. R. Alves, V. T. Carneiro, and A. C. Araujo. Direct evidence of pseudogamy in apomictic *Brachiaria brizantha* (Poaceae). *Sexual Plant Reproduction*, 14(4):207–212, 2001.
- [AHOA05] Yukio Akiyama, Wayne W. Hanna, and Peggy Ozias-Akins. High-resolution physical mapping reveals that the apospory-specific genomic region (ASGR) in *Cenchrus ciliaris* is located on a heterochromatic and hemizygous region of a single chromosome. *Theoretical and Applied Genetics*, 111(6):1042–1051, 2005.
- [AJ92] Sven Asker and Lenn Jerling. *Apomixis in Plants*. CRC Press Inc. Boca Raton, Florida, USA, 1 edition, 1992.
- [AMB⁺04] Emidio Albertini, Gianpiero Marconi, Gianni Barcaccia, Lorenzo Raggi, and Mario Falcinelli. Isolation of candidate genes for apomixis in *Poa pratensis* L. *Plant Molecular Biology*, 56(6):879–894, 2004.
- [AMP⁺00] A. C. Araújo, S. Mukhambetzhonov, M. T. Pozzobon, E. F. Santana, and V. T. Carneiro. Female gametophyte development in apomictic and sexual *Brachiaria brizantha* (Poaceae). *Rev. Cytol. Biol. Vég. Bot.*, 23:13–28, 2000.
- [AMR⁺05] Emidio Albertini, Gianpiero Marconi, Lara Reale, Gianni Barcaccia, Andrea Porceddu, Francesco Ferranti, and Mario Falcinelli. SERK and APOSTART. Candidate Genes for Apomixis in *Poa pratensis*. *Plant Physiology*, 138(4):2185–2199, 2005.
- [BBK00] R. A. Bicknell, N. K. Borst, and A. M. Koltunow. Monogenic inheritance of apomixis in two *Hieracium* species with distinct developmental mechanisms. *Heredity*, 84(2):228–237, 2000.

- [BMG83] C. A. Bilinski, J. J. Miller, and S. C. Girvitz. Events associated with restoration by zinc of meiosis in apomictic *Saccharomyces cerevisiae*. *Journal of Bacteriology*, 155(3):1178–1184, 1983.
- [BWML06] Timothy L. Bailey, Nadya Williams, Chris Misleh, and Wilfred W. Li. MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Research*, 34(suppl_2):W369–373, 2006.
- [CCA⁺99] L. Chen, M. Chikara, K. Akio, S. Akira, and A. Taiji. Isolation and Characterization of a Gene Expressed during Early Embryo Sac Development in Apomictic Guinea Grass (*Panicum maximum*). *Journal of Plant Physiology*, 154:55–62, 1999.
- [CGS⁺05] Lanzhuang Chen, Liming Guan, Misuk Seo, Franz Hoffmann, and Taiji Adachi. Developmental expression of ASG-1 during gametogenesis in apomictic guinea grass (*Panicum maximum*). *Journal of Plant Physiology*, 162(10):1141–1148, 2005.
- [CHZ⁺07] Kevin L. Childs, John P. Hamilton, Wei Zhu, Eugene Ly, Foo Cheung, Hank Wu, Pablo D. Rabinowicz, Chris D. Town, C. Robin Buell, and Agnes P. Chan. The TIGR Plant Transcript Assemblies database. *Nucleic Acids Research*, 35(suppl_1):D846–851, 2007.
- [dVSJ89] C. B. do Valle, Y. H. Savidan, and L. Jank. Apomixis and sexuality in *Brachiaria decumbens* Stapf. In *Proceedings of the XVI International Grassland Congress*, pages 407–408, 1989.
- [DW99] D. M. A. Dusi and M. T. M. Willemse. Apomixis in *Brachiaria decumbens* Stapf.: Gametophytic Development and Reproductive Calendar. *Acta Biologica Cracoviensia Series Botanica*, 41:151–162, 1999.
- [DWCB01] Zhao-Jun Ding, T. Wang, Kang Chong, and Shunong Bai. Isolation and characterization of OsDMC1, the rice homologue of the yeast DMC1 gene essential for meiosis. *Sexual Plant Reproduction*, 13(5):285–288, 2001.
- [GDY⁺05] P. J. Gulick, S. Drouin, Z. Yu, J. Danyluk, G. Poisson, A. F. Monroy, and F. Sarhan. Transcriptome comparison of winter and spring wheat responding to low temperature. *Genome*, 48(5):913–23, 2005.
- [GMDK05] Michael D. Gadberry, Simon T. Malcomber, Andrew N. Doust, and Elizabeth A. Kellogg. Primaclade—a flexible tool to find conserved PCR primers across multiple species. *Bioinformatics*, 21(7):1263–1264, 2005.
- [GNvD01] Ueli Grossniklaus, Gian A. Nogler, and Peter J. van Dijk. How to Avoid Sex: The Genetic Control of Gametophytic Apomixis. *Plant Cell*, 13(7):1491–1498, 2001.
- [GRR⁺00] Jenny Guerin, Jan Bart Rossel, Stanley Robert, Tohru Tsuchiya, and Anna Koltunow. A DEFICIENS homologue is down-regulated during apomictic initiation in ovules of *Hieracium*. *Planta*, 210(6):914–920, 2000.
- [GSH93] D. L. Gustine, R. T. Sherwood, and D. A. Hulce. A strategy for cloning apomixis-associated cDNA markers from buffelgrass. In *Proceedings of the XVII International Grassland Congress*, pages 1033–1034, 1993.
- [KCSW94] S. Kwok, S. Y. Chang, J. J. Sninsky, and A. Wang. A guide to the design and use of mismatched and degenerate primers. *Genome Research*, 3(4):S39–S47, 1994.
- [KG03] Anna M. Koltunow and Ueli Grossniklaus. APOMIXIS: A Developmental Perspective. *Annual Review of Plant Biology*, 54(1):547–574, 2003.

- [LAP⁺97] O. Leblanc, I. Armstead, S. Pessino, J. P. A. Ortiz, C. Evans, C. do Valle, and M. D. Hayward. Non-radioactive mRNA fingerprinting to visualise gene expression in mature ovaries of *Brachiaria* hybrids derived from *B. brizantha*, an apomictic tropical forage. *Plant Science*, 126(1):49–58, 1997.
- [LBC⁺02] P. Labombarda, A. Busti, M. E. Caceres, F. Pupilli, and S. Arcioni. An AFLP marker tightly linked to apomixis reveals hemizygoty in a portion of the apomixis-controlling locus in *Paspalum simplex*. *Genome*, 45(3):513–519, 2002.
- [Mat89] F. Matzk. Genetic studies on parthenogenesis in *Poa pratensis* L. *Apomixis Newsletter*, 1:32–34, 1989.
- [ME90] M. J. Malavasic and R. T. Elder. Complementary transcripts from two genes necessary for normal meiosis in the yeast *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 10(6):2809–2819, 1990.
- [OARH98] Peggy Ozias-Akins, Dominique Roche, and Wayne W. Hanna. Tight clustering and hemizygoty of apomixis-linked molecular markers in *Pennisetum squamulatum* implies genetic control of apospory by a divergent locus that may have no allelic form in sexual genotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 95(9):5127–5132, 1998.
- [OCJ⁺07] Takashi Okada, Andrew S. Catanach, Susan D. Johnson, Ross A. Bicknell, and Anna M. Koltunow. An *Hieracium* mutant, loss of *apomeiosis 1* (*loa1*) is defective in the initiation of apomixis. *Sexual Plant Reproduction*, 20(4):199–211, 2007.
- [OPL⁺97] J. P. A. Ortíz, S. C. Pessino, O. Leblanc, M. D. Hayward, and C. L. Quarín. Genetic fingerprinting for determining the mode of reproduction in *Paspalum notatum*, a subtropical apomictic forage grass. *Theoretical and Applied Genetics*, 95(5-6):850–856, 1997.
- [PEM⁺01] Silvina C. Pessino, Francisco Espinoza, Eric J. Martínez, Juan Pablo A. Ortiz, Estela M. Valle, and Camilo L. Quarín. Isolation of cDNA Clones Differentially Expressed in Flowers of Apomictic and Sexual *Paspalum Notatum*. *Hereditas*, 134(1):35–42, 2001.
- [PEO⁺98] Silvina C. Pessino, Clive Evans, Juan Pablo A. Ortiz, Ian Armstead, Cacilda B. do Valle, and Michael D. Hayward. A Genetic Map of the Apospory-Region in *Brachiaria* Hybrids: Identification of two Markers Closely Associated with the Trait. *Hereditas*, 128(2):153–158, 1998.
- [PMB⁺04] F. Pupilli, E. J. Martinez, A. Busti, O. Calderini, C. L. Quarin, and S. Arcioni. Comparative mapping reveals partial conservation of synteny at the apomixis locus in *Paspalum* spp. *Molecular Genetics and Genomics*, 270(6):539–548, 2004.
- [PNT⁺99] Bogdan Polevoda, Joakim Norbeck, Hikaru Takakura, Anders Blomberg, and Fred Sherman. Identification and specificities of N-terminal acetyltransferases from *Saccharomyces cerevisiae*. *The EMBO Journal*, 18(21):6155–6168, 1999.
- [POL⁺97] S. C. Pessino, J. P. A. Ortiz, O. Leblanc, C. B. do Valle, C. Evans, and M. D. Hayward. Identification of a maize linkage group related to apomixis in *Brachiaria*. *Theoretical and Applied Genetics*, 94(3-4):439–444, 1997.
- [QLH⁺00] John Quackenbush, Feng Liang, Ingeborg Holt, Geo Perlea, and Jonathan Upton. The TIGR Gene Indices: reconstruction and representation of expressed gene sequences. *Nucleic Acids Research*, 28(1):141–145, 2000.

- [RCB⁺02] D. Roche, A. Conner, A. Budiman, D. Frisch, R. Wing, W. Hanna, and P. Ozias-Akins. Construction of BAC libraries from two apomictic grasses to study the microcolinearity of their apospory-specific genomic regions. *Theoretical and Applied Genetics*, 104(5):804–812, 2002.
- [RCD⁺03] Júlio C. M. Rodrigues, Gláucia B. Cabral, Diva M. A. Dusi, Luciane V. de Mello, Daniel J. Rigden, and Vera T. C. Carneiro. Identification of differentially expressed cDNA sequences in ovaries of sexual and apomictic plants of *Brachiaria brizantha*. *Plant Molecular Biology*, 53(6):745–757, 2003.
- [SBF07] Manjit Singh, Byron L. Burson, and Scott A. Finlayson. Isolation of candidate genes for apomictic development in buffelgrass (*Pennisetum ciliare*). *Plant Molecular Biology*, 64(6):673–682, 2007.
- [Sch97] Gregory D. Schuler. Sequence Mapping by Electronic PCR. *Genome Research*, 7(5):541–550, 1997.
- [VCNB⁺96] J.-Philippe Vielle-Calzada, Michael L. Nuccio, Muhammad A. Budiman, Terry L. Thomas, Byron L. Burson, Mark A. Hussey, and Rod A. Wing. Comparative gene expression in sexual and apomictic ovaries of *Pennisetum ciliare* (L.) Link. *Plant Molecular Biology*, 32(6):1085–1092, 1996.

| Accession number (GenBank) | Theoretical amplicon length (bp) ^a | Sequence Primer Fwd (5'-3') | Tm Fwd (°C) | Degen- eration Fwd ^b | Sequence Primer Rev (5'-3') | Tm Rev (°C) | Degen- eration Rev | Total degen- eration ^c |
|-------------------------------|---|-----------------------------|----------------|---------------------------------------|-----------------------------|----------------|--------------------------|---|
| AB000809 | 380 | TGCSCACCTCGCTGGAG | 63.7 | 2 | CATACGGCAKARGTGGCAG | 58.5 | 4 | 8 |
| AF475105 | 371 | GAAGTACTGGAAYGACCCCTG | 53.7 | 2 | GGGGTGTTCTTGTTCCTTRTC | 52.2 | 2 | 4 |
| AJ786393 | 380 | GTTGGACAGTACCTGAYAGCA | 55.5 | 2 | CGCACTATCCATGAKCCCT | 55.6 | 2 | 4 |
| AJ810708 | 301 | CAAATYTGGAAYACTGCTGG | 53.9 | 4 | TTCCTKGCACCTGGTCTCCATG | 56.3 | 2 | 8 |
| AJ810709 | 431 | GGCATTGCTGCTTCTTACAT | 57.7 | 1 | GCTTKCTGCACYGTCAAAGCT | 57.8 | 4 | 4 |
| AJ810710 | 326 | TACACAAGTATTACAGACAGATGA | 55.2 | 1 | ATRAITTTCTTGGTCCCG | 51.3 | 2 | 2 |
| AJ841698 | 407 | ATGGGTTTGGTATTCTCTCTTGGG | 58.1 | 1 | GGACCAGATAGCTCAATTGGCTC | 57.5 | 1 | 1 |
| D37938 (pempsbca) | 416 | AYACATTTGCATATCACACATACTT | 52.1 | 2 | TCTTCRCAGGYTTCAGAG | 54.3 | 4 | 8 |
| D37939 (pempsb3lab) | 375 | TGAAGTAYTGGAAATGACCCCTG | 53.1 | 2 | CARYGGRTGTTCTTGTCT | 54 | 4 | 8 |
| D37940 (pempsbhc) | 214 | AATGTTCAAGTTACACYNNGNCTATGC | 58 | 32 | ATTGARRAATTTTCAGGGAGGA | 51.5 | 4 | 128 |
| EF517497 | 423 | GGCAAAWTTAGAAAWGAGG | 48 | 8 | ACTTRCTYRGTTTTACATACCATAA | 51.3 | 8 | 64 |
| EF517498 | 256 | GAATTGGGTCCACNATTKCTC | 54.8 | 8 | TCGTTCAGCATGTCTWGGTAGAACT | 57.6 | 2 | 16 |
| EF530198 | 330 | GAYGTGAAGAAGCTGMARGAT | 53.9 | 8 | AAGCTGAGTKGARACACAWAGAGT | 56.2 | 8 | 64 |
| EF530199 | 336 | ATTCCATAACGCKCCGG | 53.9 | 2 | GCASTGATGTAGTTTGTATCTTMA | 52.3 | 4 | 8 |
| M38357 (yscsp013) | 305 | ATGCAGGGCTGCAACCTGAT | 60.3 | 1 | GGACGTGGAGGGASACGTACT | 60.5 | 8 | 8 |
| U40219 (pcu40219) | 342 | GATTTCTCACCAAGGAGGT | 54.3 | 1 | CATGCTTYGCCTTYTCT | 52.3 | 2 | 2 |
| U65082 | 355 | GARGCAGKAAATTCATCAA | 51.5 | 4 | GGATTCCWWTGATSAGCTTCTCC | 55.3 | 4 | 16 |
| U65383 | 248 | CRRCGCTACATTGATGATGATG | 54.1 | 4 | CATCAGAGRARGTAAGCSATCA | 54.2 | 8 | 32 |
| U65384 | 205 | ATTCGGCTTYGGTCSAGA | 55.1 | 4 | TTCAACAGCTTTCACACACAT | 54.8 | 2 | 8 |
| U65385 | 243 | TTGCMAGCTTCTGCTGCTCTT | 60.8 | 2 | CACYGCRGCTCCAGCTGG | 61.9 | 4 | 8 |
| U65388 | 246 | ATAYGAKACTCTGCAATTGATAA | 48.9 | 4 | GCTATGTAATGTACTGCTACTACTA | 51.5 | 1 | 4 |
| U65389 | 261 | ATCAGGGCTGCAACCTGAT | 63.1 | 1 | AGCTCAAGCTTGTAGCTTGCCGTAG | 60.3 | 1 | 1 |

Table 2: Sequences obtained for the degenerate primers in this work, theoretical amplicon, Tm and level of degeneration. a) Average length calculated by ePCR in the group of sequences with which alignments were made. b) The degeneracy of every primer is determined according to the existent degeneracies. Every two-base degeneracy (S, Y, W, R, K, M) sums 2 to total degeneracy. Every three-base degeneracy (B, D, H, V) sums 3 to total degeneracy. Each N sums 4 to total degeneracy. Total primer degeneracy is calculated multiplying the contribution of each base position. c) Combined degeneracy of the primer set, calculated by multiplying the degeneracy of forward and reverse primers.

