

# SCIENTIFIC REPORTS

OPEN

## A Portion of the Apomixis Locus of *Paspalum Simplex* is Microsyntenic with an Unstable Chromosome Segment Highly Conserved Among Poaceae

Giulio Galla<sup>1</sup>, Lorena A. Siena<sup>2</sup>, Juan Pablo A. Ortiz<sup>2</sup>, Helmut Baumlein<sup>3</sup>, Gianni Barcaccia<sup>1</sup>, Silvana C. Pessino<sup>2</sup>, Michele Bellucci<sup>4</sup> & Fulvio Pupilli<sup>4</sup>

The introgression of apomixis in major seed crops, would guarantee self-seeding of superior heterotic seeds over generations. In the grass species *Paspalum simplex*, apomixis is controlled by a single locus in which recombination is blocked. In the perspective of isolating the genetic determinants of apomixis, we report data on sequencing, *in silico* mapping and expression analysis of some of the genes contained in two cloned genomic regions of the apomixis locus of *P. simplex*. *In silico* mapping allowed us to identify a conserved synteny group homoeologous to the apomixis locus, located on a telomeric position of chromosomes 12, 8, 3 and 4 of rice, *Sorghum bicolor*, *Setaria italica* and *Brachypodium distachyum*, respectively, and on a more centromeric position of maize chromosome 1. Selected genes of the apomixis locus expressed sense and antisense transcripts in reproductively committed cells of sexual and apomictic ovules. Some of the genes considered here expressed apomixis-specific allelic variants which showed partial non-overlapping expression patterns with alleles shared by sexual and apomictic reproductive phenotypes. Our findings open new routes for the isolation of the genetic determinants of apomixis and, in perspective, for its introgression in crop grasses.

The Poaceae angiosperm family, commonly referred to as grasses, accounts for about 70% of crops in the world. Although frequent events of genomic re-patterning and gene losses, especially occurring after whole genome duplication (WGD) events, caused considerable divergence in genome size and chromosome number<sup>1,2</sup>, genes tend to maintain their position in the same chromosome region with strict conservation of gene order (collinearity) or without it (synteny)<sup>3</sup>. Genomic collinearity between grasses is correlated with conservation of expression between orthologous gene pairs<sup>4</sup>. Therefore, comparative analysis of conservation of gene position makes it possible: (i) to get insights into the evolution of specific genes or entire metabolic networks and signalling pathways<sup>5</sup>; (ii) to identify niche specific genes important for adaptation<sup>6</sup>; (iii) to unveil large scale genomic events such as WGDs<sup>7</sup>; and, (iv) to link phenotypic traits with genotypic properties, thereby improving gene discovery<sup>8</sup>.

The genus *Paspalum*<sup>9</sup> Panicoideae is characterized by an extraordinary plasticity of the reproductive system<sup>10</sup>, including strict outcrossing, a variable extent of allowed self-compatibility, and nearly obligate apomixis, *i.e.* asexual reproduction by seeds<sup>11</sup>. Introgression of apomixis to crops would allow the fixation of heterosis and therefore the re-seeding of hybrid F1 seeds without vigour loss<sup>12</sup>. The three main biological components of apomixis in *Paspalum* are: (1) apospory (*i.e.* unreduced embryo sac development from nucellar cells of the ovule without meiosis); (2) parthenogenesis (*i.e.* fertilization-free embryo development); and (3) unbalanced endosperm formation (*i.e.* capacity to form endosperm deviating from the canonical 2 maternal: 1 paternal genome ratio,

<sup>1</sup>Department of Agriculture Food Natural resources Animals and Environment (DAFNAE), University of Padova, 35020, Legnaro (PD), Italy. <sup>2</sup>Instituto de Investigaciones en Ciencias Agrarias de Rosario (IICAR), CONICET-UNR, Laboratorio de Biología Molecular, Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, S2125ZAA, Zavalla, Argentina. <sup>3</sup>The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), D-06466, Gatersleben, Germany. <sup>4</sup>Institute of Biosciences and Bioresources (IBBR), National Research Council (CNR), 06128, Perugia, Italy. Correspondence and requests for materials should be addressed to F.P. (email: [fulvio.pupilli@ibbr.cnr.it](mailto:fulvio.pupilli@ibbr.cnr.it))

allowing a maternal contribution excess). These components are controlled by a single complex dominant super-locus (Apomixis Controlling Region or ACR)<sup>10</sup>. Comparative genetic mapping efforts in *Paspalum* evidenced various extents of synteny of the ACR with the subtelomeric part of rice chromosome 12 long arm<sup>13–16</sup>. The ACR of *P. simplex* revealed structural features of heterochromatin, namely presence of transposable elements (TEs), gene degeneration<sup>17</sup> and deregulation<sup>18</sup>. One ACR-specific pseudogene, *PsORC3*, constitutively expressed a non-coding RNA that silenced its functional counterpart via a sense-antisense mechanism<sup>19</sup>. As the rice homolog of this gene did not map on the distal end of chromosome 12, it was assumed that it migrated to the ACR from other locations after rice and *Paspalum* diverged from a common ancestor.

Although rice is considered the best-suited reference genome for comparative genomics in grasses<sup>20</sup>, this species is phylogenetically rather distant from *Paspalum*<sup>21</sup>. Nowadays, genome sequences representative of the most economically important grass clades, such as Panicoideae including *Sorghum*, maize and *Setaria* and Poideae with *Brachypodium* are publicly available<sup>22–25</sup>. Among these, the species of particular interest for comparisons with *Paspalum* are *Sorghum bicolor* (L.) and *Setaria italica* (L.) P. Beauv., as both of them diverged from the *Paspalum* clade ca. 10 mya<sup>26</sup>, thus representing the two most closely related genomes for which Whole Genome Sequencing (WGS) information is available. Although the apospory-specific genomic region (ASGR) of other grasses, such as *Pennisetum squamulatum* syn. *Cenchrus ciliaris*, shares many structural similarities with the *Paspalum* ACR, i.e. repression of recombination, accumulation of repetitive elements and gene degeneration, no relevant large-scale collinearity and/or synteny between the ASGR and reference genomes (rice, *Setaria* and *Sorghum*) were detected so far<sup>27</sup>. Although comparative mapping of apomixis within the *Paspalum* genus revealed a low extent of gene movement and divergence at the ACR even among very closely related species<sup>15,16</sup>, a portion of the same region, identified by markers of rice chromosome 12 was linked to apomixis in all apomictic species of *Paspalum* analysed to date<sup>15</sup>. We argued that markers not linked to the ACR in all species of *Paspalum* belong to genomic regions that are dispensable for the expression of the trait, whereas all those linked to apomixis in multiple *Paspalum* spp. delineate a genomic portion in which the essential genetic determinants of the trait are likely located. Bearing this in mind, we investigated on whether larger areas of synteny could be identified between the ACR of *P. simplex* and specific regions of grasses (i.e. *Sorghum*, maize, *Brachypodium* and *Setaria*) that are more closely related to *Paspalum* compared to rice.

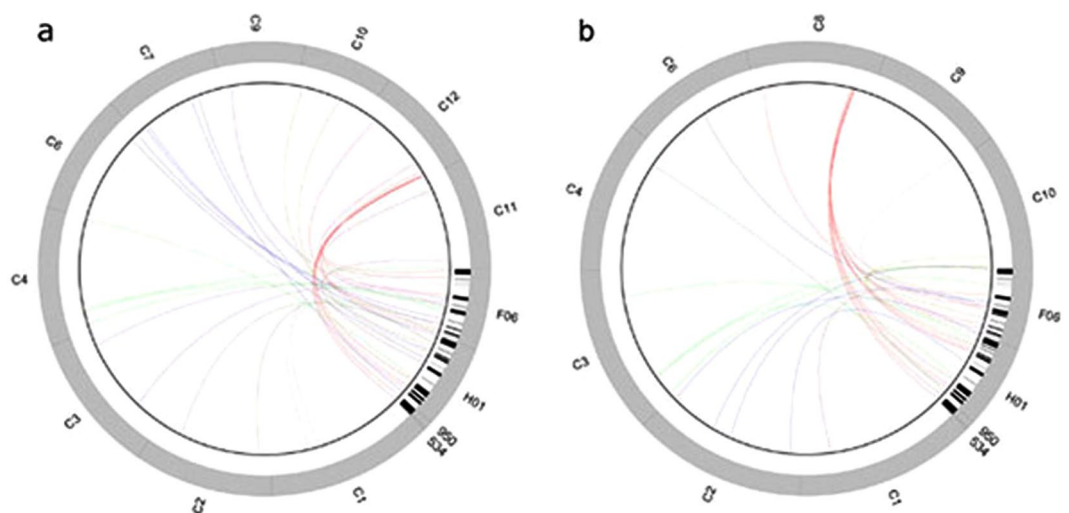
The aim of this research was to study the genomic arrangement of ACR portions in *P. simplex* at both structural and functional levels. The specific aspects we wanted to address were: *i*) to disclose more extensive areas of homology between the ACR of *P. simplex* and the closely related genomes of *S. italica*, *S. bicolor* and maize together with those of the more distantly related *Brachypodium* and rice as controls; and *ii*) to verify the existence of relationships between degeneration of the the apomixis-linked genes, and possible silencing effect on their homologues.

## Results

**Comparative mapping.** To identify BAC clones containing the genetic determinants of apomixis, 39 out of the 41 BACs positive to the SCAR markers co-segregating with apomixis<sup>28</sup> were hybridised with marker-derived probes delineating the sub-portion of the ACR linked to apomixis in several *Paspalum* species (c1069, c454 and c996)<sup>15</sup>. Of these, only c996 showed a clear signal in two overlapping clones (127F6 and 296A7; Supplementary Fig. S1), as these were selected with the same SCAR marker<sup>28</sup>. The other two BACs 333G1 and 312H12, included in the same contig<sup>28</sup>, showed a less intense signal likely due to either background signal (see below) or partial hybridisation with the labelled probe. Background signals in spots not related to 127F6, 296A7, 333G1 and 312H12 are due to unspecific hybridisation of template BAC vector DNA with residual labelled plasmid DNA vector still present in the probe mixture. Then, among these BACs only 127F6 was chosen for sequencing. The other BAC considered in this study (366H1) was isolated previously with the AFLP-derived SCAR marker EM 180<sup>29</sup>, which mapped in the same sub-portion of the ACR<sup>15</sup>. Both BAC clones were sequenced at 6× coverage. Finally, 6 and 5 contigs whose length ranged from 976 to 73,550 bp and from 9,403 to 75,967 bp were assembled for 127F6 and 366H1, respectively. Various transposon elements (TE)-related sequences (Supplementary Table S1) were identified in the considered contigs covering from 12% (PS127F6\_c1) to 20.40% (PsH10) (Table 1) of the total length analysed. No relevant differences were detected between apomixis-linked contigs reported in this study and those previously analysed (PsH10)<sup>17</sup> nor with the contig (Ps366H1\_c5) containing the hemizygous SCAR marker used to isolate the related BAC. Retrotransposons of the LTR gypsy and Copia subclasses were the most abundant, whereas among transposons, elements related to Helitron subclasses were the most frequent. Proportions of simple and low complexity repeats were highly similar in all 4 query sequences. Most of the apomixis-linked genes annotated on TAIR database are depicted to DNA/RNA binding molecular process (Supplementary Table S1). The largest contig assembled for each BAC was analyzed in detail for gene synteny and collinearity with five reference grass genomes. The contig PS127F6\_c1 of the BAC 127F6 contained 50 ORFs organized as 15 predicted genes and 2 pseudogenes (the latter so defined by the presence of one or more premature stop codons on their CDS) named PsACR|F.1–17 (Supplementary Table S2). However only 8 of them showed significant homology (*e* value ≤ 1.0e[−9]) with annotated genes in GRAMENE database and, of these, only 4 in TAIR repository. The gene PsACR|F.15 was homologous to the c996 EST rice marker (OS12G0616200) used to select the BAC. The position of this gene in the rice map marks the telomeric end of the rice region homologous to the ACR common to several *Paspalum* spp.<sup>15</sup>. However, the gene PsACR|F.7 was homologous to a more telomeric rice gene (OS12G0616500), suggesting that the ACR of *Paspalum* might be larger than that estimated previously. Thus, in this contig a region spanning 39,587 bp bracketed by the two genes PsACR|F.7 and PsACR|F.15 was syntenic with a portion of similar size of the rice genome located on chromosome 12 (39,515 bp, Supplementary Table S2 and Fig. 1a). However, as the rice gene OSG12G0616400, located between OSG12G0616200 and OSG12G0616500 (<http://www.gramene.org/>), has been replaced in *Paspalum* by seven genes whose rice homologs did not belong to a specific syntenic block (PsACR|F.8–14; Supplementary Table S2; Fig. 2), gene collinearity between rice and

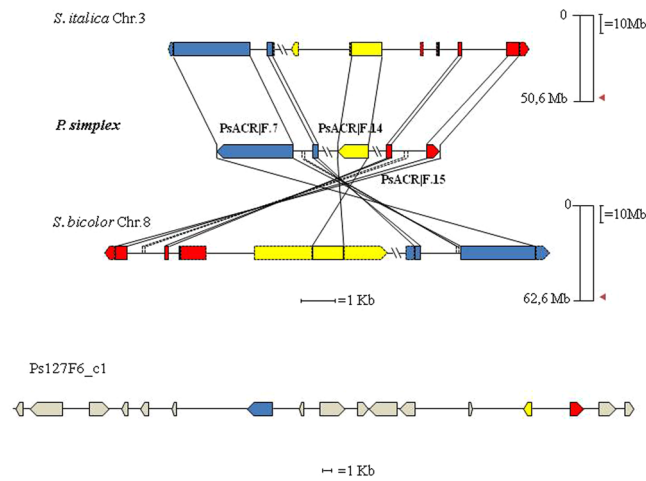
Type of repeats	Repeat class/family	Query sequence (length)							
		Ps127F06_c1 (73,550)		Ps366H1_c1 (75,957)		Ps366H1_c5 (11,754)		PsH10 (75,005)	
Retroelements		N	Length occupied (%)	N	Length occupied (%)	N	Length occupied (%)	N	Length occupied (%)
	LINE/L1			1	39 (0.05)	2	1,471 (12.51)	2	1,621 (2.16)
	LTR/Gypsy	9	4,355 (5.92)	16	3,696 (4.87)			6	1,260 (1.68)
	LTR/Copia	1	69 (0.09)	4	5,381 (7.08)			8	8,923 (11.90)
<b>Transposons</b>									
	DNA/TcMar-Stowaway			2	418 (0.55)			2	433 (0.58)
	RC/Helitron	4	2,736 (3.72)	2	173 (0.23)			1	143 (0.19)
	DNA/MuLE-MuDR	1	11 (0.01)						
	DNA/PIF-Harbinger	2	554 (0.75)			1	98 (0.83)		
	DNA/hAT-Ac	2	381 (0.52)						
	DNA/hAT-Tip100							2	1,580 (2.11)
	DNA/PIF-Harbinger	2	554 (0.75)					1	932 (1.24)
	DNA/CMC-EnSpm	1	166 (0.23)					3	409 (0.55)
Total interspersed repeats			8,826 (12)		9,707 (12.8)		1,569 (13.35)		15,301 (20.40)
Simple repeats		25	1,308 (1.78)	19	852 (1.12)	4	242 (2.06)	12	851 (1.13)
Low complexity		1	39 (0.05)	5	214 (0.28)	1	60 (0.51)	1	47 (0.06)

**Table 1.** Analysis of repetitive elements in apomixis-linked BACs of *P. simplex*.

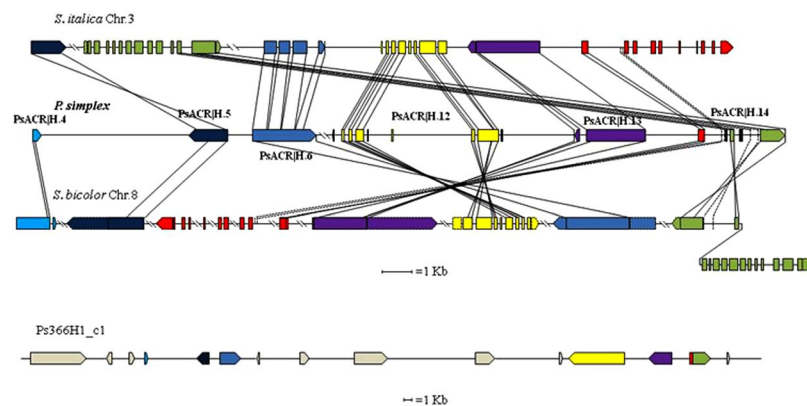


**Figure 1.** Mapping of genes contained in apomixis-linked BACs on chromosomes of (a) *Oryza sativa* and (b) *Sorghum bicolor*. F06, H01, 950 and 534 correspond to contigs PS127F6\_c1, PS366H1\_c1, H10\_950 and H10\_534, respectively. C1-C12 represent the chromosome number for each genome. Genes providing alignments with  $e\text{-value} < 1.0e^{-9}$  were plotted as black bars. Otherwise, gene regions were plotted as grey bars. Red lines link genes to the conserved chromosome area related to apomixis revealed in this study.

ACR has not been respected in this area. Looking at the other reference grass genomes we noticed that the ACR of *Paspalum* pointed specific areas of homology to a telomeric region of chromosomes 8 of *Sorghum* (Fig. 1b), 3 of *Setaria* and 4 of *Brachypodium*, and to a more centromeric region of chromosome 1 of maize (Supplementary Table S2). Similarly to rice, synteny but not collinearity was detected to the region of *Sorghum* delineated by the three genes SORBI\_008G172500, SORBI\_008G172200 and SORBI\_008G172100 (homologous of genes PsACR|F.7, PsACR|F.14 and PsACR|F.15 respectively; Supplementary Table S2). Conversely, the homologous area of *Setaria* maintained all the homologs (SETIT\_021254m.g., SETIT\_022811m.g. and SETIT\_023449m.g.) in perfect collinearity with the same genes of *P. simplex* (Supplementary Table S2; Fig. 2). Overall, homology inferred from statistically significant similarity was detected between predicted exons of *Paspalum* genes and their homologs in grass genomes, whereas no significant homology was detected for the intronic regions with the exception of last intron of the genes PsACR|F.7 and PsACR|F.15 and the related genes of *Sorghum* only (Fig. 2). Although the reciprocal orientation of genes was conserved between *Paspalum* and grasses, the whole region was inverted in *Sorghum*, with respect to *Paspalum* and *Setaria* (Fig. 2).



**Figure 2.** Diagram of microcollinearity of the genes contained in the contig PS127F6\_c1 with the conserved area of homology in *Setaria italica* and *Sorghum bicolor* genomes. Position of the apomixis-linked chromosome areas are highlighted on the right. Vertical continuous and dotted lines connect homologous coding and non-coding regions respectively, full boxes identify exons, dotted boxes UTRs, and horizontal lines refer to introns and/or intergenic regions. Each gene is represented by one colour with its orientation indicated by the arrow head. Ps127F6\_c1 identify the *P. simplex* whole contig; coloured boxes identify genes mapped to the apomixis related area of *Sorghum* and *Setaria* whereas those coloured in grey mapped elsewhere in the same genomes.



**Figure 3.** Diagram of microcollinearity of the genes contained in the contig PS366H1\_c1 with the conserved area of homology in *Setaria italica* and *Sorghum bicolor* genomes. Symbols are as in Fig. 2.

The analysis of the largest contig of the BAC 366H1 (PS366H1\_c1) highlighted the presence of 10 genes and 5 pseudogenes of which 8 and 4 showed highly significant homology to annotated genes in GRAMENE and TAIR repositories respectively (PsACR|H.1–15; Supplementary Table S2). Among the predicted genes, we found another group of genes/pseudogenes that pointed on the same chromosome regions that were located by the previous group of *Paspalum* genes (Supplementary Table S2; Figs 1 and 3). In *Sorghum* this group included the homologs of the genes PsACR|H.14, PsACR|H.13, PsACR|H.12, PsACR|H.6, PsACR|H.5 and PsACR|H.4. The homologues of the last two exons of SORBI\_008G171900 (PsACR|H.14b) were re-located immediately downstream of the first exon of the gene SORBI\_008G171400 (PsACR|H.14a) indicating the occurrence of a translocation within the ACR (Supplementary Table S2; Fig. 3). In *Setaria* the homology group was delineated by the homologs of the genes PsACR|H.5, PsACR|H.6, PsACR|H.12, PsACR|H.13 and PsACR|H.14. The translocation that generated the hybrid gene PsACR|H.14 of *Paspalum* was also detected for the related homologs of *Setaria* (Fig. 3) as well as rice, maize and *Brachypodium* (Supplementary Table S2). Similarly to what is reported for the contig PS127F6\_c1, the homology regions of genes contained in the contig PS366H1\_c1, are mainly confined to the ORFs and in small intronic areas in *Setaria* and *Sorghum* (Fig. 3). The large-scale inversion that affected *Setaria* and *Paspalum* on one side and *Sorghum*, on the other, was also evident for the latter contig (Fig. 3). Of the 3 predicted genes contained in the contig PS366H1\_c5, only one showed relevant homology with related genes of *Brachypodium*, maize and rice and none of them showed matching sequences in TAIR database. Scarce homology detected in this contig could be related to sequence divergence that generated areas of hemizygoty on which apomixis-specific SCARs could be developed. The homologs of the two genes (H10\_950 and H10\_534), belonging to the apomixis-linked BAC H10 previously sequenced<sup>17</sup>, are also located in the same chromosome areas related to apomixis of each of the 5 reference genomes even though they are positioned between 2.2 (*Sorghum*)



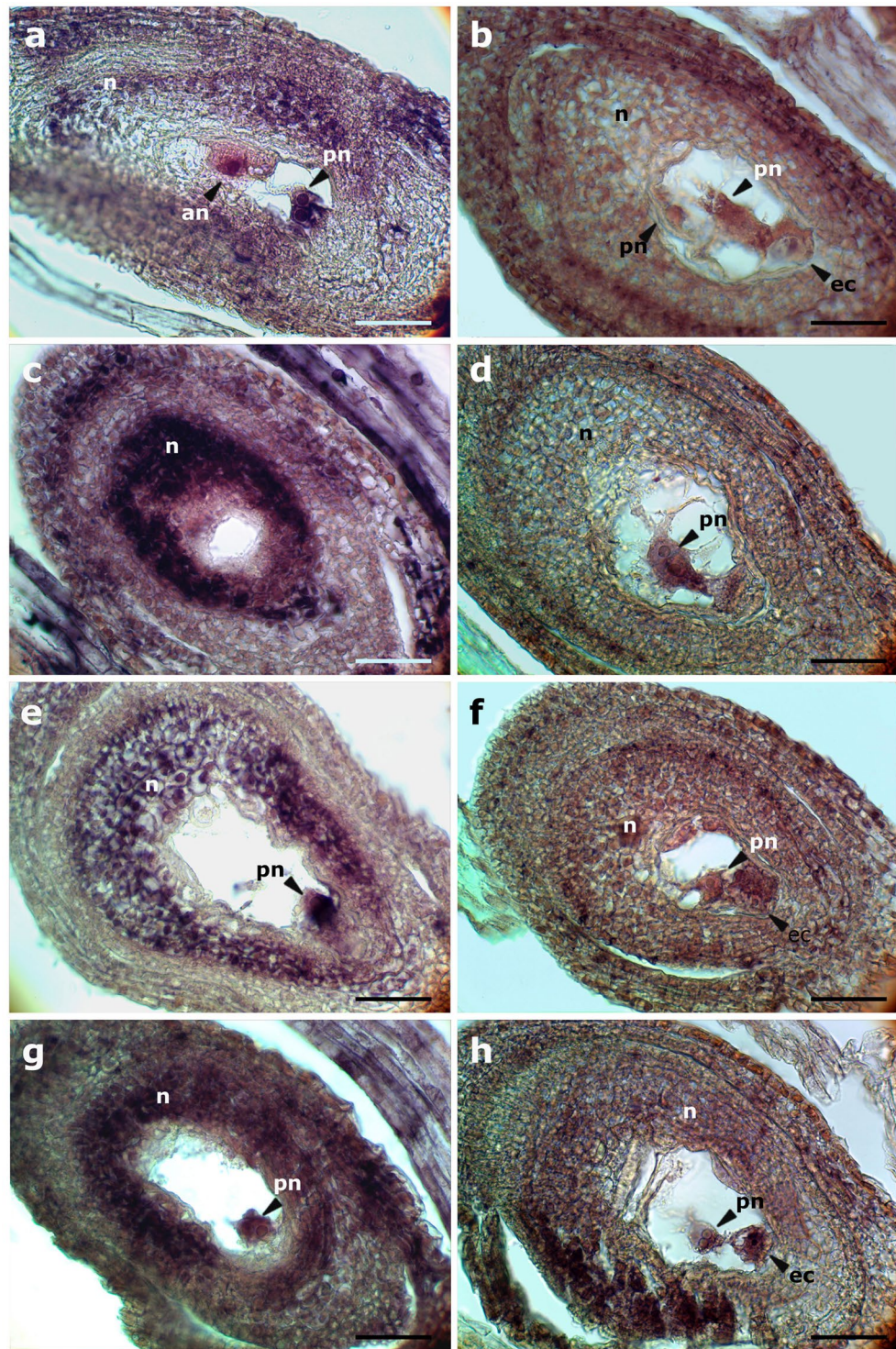
and 7.8 (maize) Mb respectively away from the conserved main homology areas (Supplementary Table S2). Based on the distance of the homologs at the extremities of the ACR-related area in *Sorghum* and *Setaria*, we estimate the total length of the ACR be around 2.2 Mb and 1.6 Mb respectively. Then, the contigs sequenced and analysed in this study (0.16 Mb) correspond approximately to between 7% and 10% of the total length of the apo locus. To sum up, the two apomixis-linked BACs described here (366H1 and 127F6), together with the previously reported H10, show strong synteny with a chromosome area in the telomeric positions of chromosomes 12, 8, 3, and 4 of rice, *Sorghum*, *Setaria* and *Brachypodium*, respectively, together with a more centromeric position of chromosome 1 in maize. Both large- and small-scale rearrangements affected the ACR of *Paspalum* compared to the related areas of the reference genomes considered. While the large-scale inversions detected likely derived from grass speciation (see GRAMENE database at <http://www.gramene.org/>), the small-scale translocations, such as those affecting genes PsACR|F.7 and PsACR|H.14, are likely *Paspalum*-specific and could have arisen either at the *Paspalum* divergence or because of polyploidization.

**In situ hybridisation analysis.** Based on *in silico* mapping, we noticed that some of the genes contained in the apomixis-linked BACs of *P. simplex* belong to a kind of “ancestral” chromosome segment homologous to the subtelomeric region of rice chromosome 12 (genes PsACR|F.7, PsACR|F.15, PsACR|H.12, PsACR|H.13, PsACR|H.14 and both genes of BAC H10), while others (for example genes PsACR|F.14, PsACR|H.5, and PsACR|H.6) migrated from different parts of the grass genome and contributed to form a more recent syntenic group homologous to the telomeres of the newly formed chromosomes 8 and 3 of *Sorghum* and *Setaria*, respectively. Furthermore, as both sense and antisense transcripts were detected in the female reproductive cell lineages for the pseudogene PsACR|H.6 (*PsORC3*<sup>19</sup>), we wondered whether antisense-mediated regulation of gene expression could be related to the condition of pseudogene. To address this hypothesis, *in situ* hybridisation of the following genes located on the portion of the ACR represented by the contig PS366H1\_c1 was undertaken on apomictic and sexual flowers of *P. simplex* at anthesis: *i*) gene PsACR|H.5 (encoding for a F-box domain containing protein) as representative of newly migrating functional genes and, *ii*), gene PsACR|H.13 (similar to a PPR repeat containing gene) as an “ancestor” non-functional gene (Supplementary Table S2). The antisense probe for gene PsACR|H.5 (detecting the sense transcript), revealed a signal in polar nuclei and antipodals in sexual (Fig. 4a) and polar nuclei in apomictic ovules (Fig. 4b). The sense probe (detecting the antisense transcript) showed an intense hybridisation signal in the nucellus (Fig. 4c), and polar nuclei (Supplementary Fig. S2a) of sexual ovules. Conversely, the same probe showed no signal in nucellus and a strong one in polar nuclei of apomictic ovules (Fig. 4d). The gene PsACR|H.13 was expressed as sense transcripts in the nucellus, polar nuclei (Fig. 4e), egg cell (Supplementary Fig. S2b) and antipodals (Supplementary Fig. S2c) of sexual ovules, and in polar nuclei, egg cell (Fig. 4f) and antipodals (Supplementary Fig. S2d) of apomictic ovules. A strong hybridising signal related to antisense transcripts was detected in the nucellus, polar nuclei (Fig. 4g), and antipodals (Supplementary Fig. S2e) of sexual ovules, and in polar nuclei, egg cell (Fig. 4h), and antipodals of the apomictic ones (Supplementary Fig. S2f). Although multiple aposporic embryo sacs are a distinctive character of apomictic reproduction in *P. simplex*; these are normally detected at early stages of development. Usually at the stage of anthesis, fewer or more often a single embryo sac is contained in mature apomictic ovules<sup>30</sup>.

In conclusion, both genes are expressed mainly in nucellus and polar nuclei indicating that they are probably subjected to an extent of coordinate expression. Furthermore, as these genes expressed both sense and antisense transcripts in apomictic and sexual phenotypes, no relationships could be evidenced between gene migration and expression mode for the same genes.

**Real Time RT-PCR.** To investigate whether quantitative differences on gene expression between apomictic and sexual flowers could be related to the condition of pseudogene, Real-Time qPCR assays were performed on 4 genes of the contig PS36601\_c1, including PsACR|H.5 and PsACR|H.13 used for *in situ* analyses and the two additional genes PsACR|H.7 and PsACR|H.9. The two genes PsACR|H.5 and PsACR|H.7 are predicted to encode for functional proteins, whereas PsACR|H.9 and PsACR|H.13, likely expressed non-coding transcripts. On the basis of the presence of phenotype-specific SNPs on the cloned alleles (see M&M), we amplified 2 apomixis-specific alleles for the gene PsACR|H.5 (#5\_1 and #5\_2; Fig. 5a) and one apomixis-specific splicing variant for PsACR|H.13 (#13\_1; Fig. 5d). No sex-specific alleles could be amplified, as sexual genomes are shared between apomictic and sexual genotypes of *P. simplex*. As a consequence of this, a single allele common to both genotypes was detected for the genes PsACR|H.5, PsACR|H.7 and PsACR|H.13 (#5\_3, #7\_1 and #13\_2 respectively; Fig. 5a,b,d) and 2 for the gene PsACR|H.9 (#9\_1, #9\_2; Fig. 5c). Specificity was confirmed by partial non-overlapping expression patterns. More in detail, for the gene PsACR|H.5, while the allele variant # H.5\_3 displayed uniform expression in all investigated samples, #H.5\_1 and #H.5\_2, were highly expressed in apomictic florets and showed no expression in the sexual ones (Fig. 5a). The expression of the single allele #H.7\_1 was identical in both reproductive phenotypes at pre-anthesis, increased at anthesis in sexual florets and decreased in post-anthesis (Fig. 5b). Regarding the gene PsACR|H.9, while no difference in transcript abundance was detected between sexual and apomictic flowers at pre-anthesis, both alleles were dramatically down regulated in apomictic florets at anthesis and in both phenotypes at post-anthesis stages (Fig. 5c). For the gene PsACR|H.13, our investigations allowed to identify a splicing variant that forms a transcript with a longer coding sequence compared to the expected one. The expression of this splicing variant (#13\_1; Fig. 5d) was clearly detectable in all apomictic samples, with little or no expression variation among developmental stages, whereas it was undetectable in all sexual samples. The relative abundance of the common allele # 13\_2 was similar in sexual and apomictic samples in pre-anthesis stages, while it was up regulated in apomictic flowers at both anthesis and post anthesis stages (Fig. 5d).

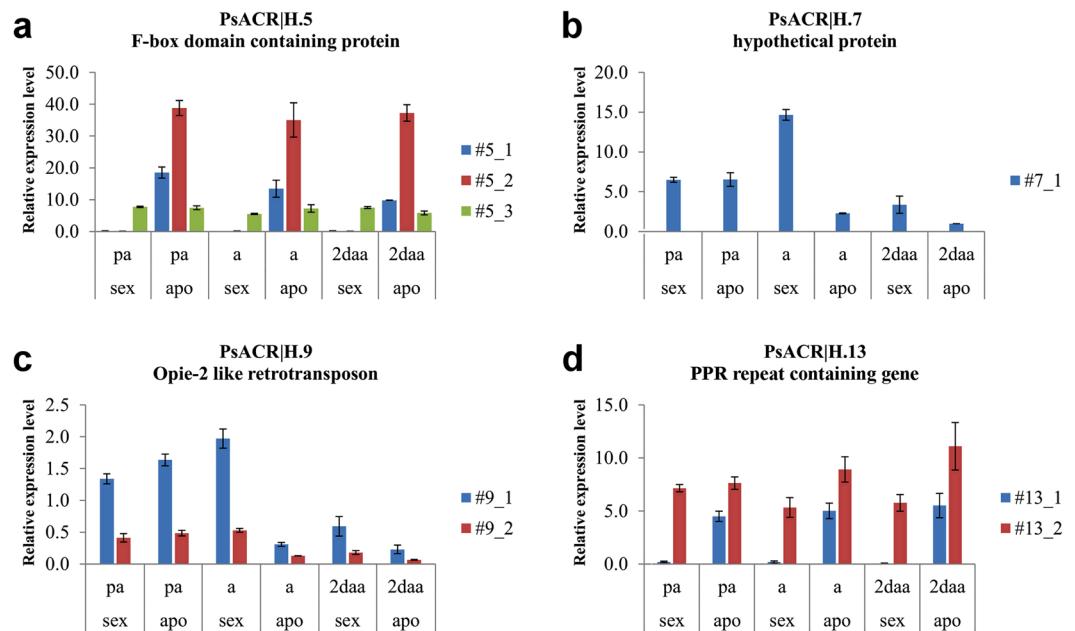
To sum up, apomixis-specific allelic variants were characterized by constitutive expression across developmental stages considered. Furthermore, no obvious relationship was found between expression pattern of the selected apomixis-linked genes and their degeneration.



**Figure 4.** *In situ* hybridisation analysis of PsACR[H5 (a–d) and PsACR[H.13 (e–h) transcripts in female reproductive tissues at anthesis of sexual and apomictic *P. simplex*. Sections of sexual (a,e) and apomictic (b,f) ovules hybridised with antisense probes (detecting the sense transcripts); sexual (c,g) and apomictic (d,h) ovules hybridised with the sense probes. an = antipodal cells, ec = egg cell, n = nucellus and pn = polar nuclei. Bar = 50  $\mu$ m.

***In silico* gene expression analysis.** To investigate whether the apomixis-linked genes of *P. simplex* analyzed here are expressed in reproductive tissues and validate the expression pattern of those analysed by RT-qPCR, their sequences were queried onto the available 454/Roche reference flower transcriptome database of the related species *P. notatum*<sup>31</sup>. In doing so, expressed sequences were found for 20 out of the 34 apomixis-linked





**Figure 5.** Determination of the transcriptional profile of 4 apomixis-linked genes through qRT-PCR assay carried out at three development stages. pa = pre-anthesis, a = anthesis, 2daa = 2 days after anthesis, sex = sexual plant, apo = apomictic plant. Error bars indicate the SEs.

<i>P. notatum</i> Roche-454 transcriptome expression analysis						
Isotig ID	13832*	20690	27671	10887	9764*	9721
Reads Apo <sup>a</sup>	0.987 (1)	9.873 (10)	86.856 (88)	8.883 (9)	12.831 (13)	67.116 (68)
Reads Sex <sup>a</sup>	17.204 (17)	2.024 (2)	49.588 (49)	1.012 (1)	0	148.764 (147)
logFC	3.894	-2.287	-0.879	-3.054	-6.732	1.075
logConc	3.974	3.522	6.663	3.329	3.611	3.611
p-value	1.46e <sup>-04</sup>	3.86e <sup>-02</sup>	6.06e <sup>-04</sup>	2.15e <sup>-02</sup>	2.45e <sup>-04</sup>	2.57e <sup>-07</sup>
FDR	5.57e <sup>-03</sup>	3.83e <sup>-01</sup>	1.80e <sup>-02</sup>	2.70e <sup>-01</sup>	8.49e <sup>-03</sup>	2.07e <sup>-05</sup>
<i>P. simplex</i> (in RT-qPCR assays)						
Gene ID	PsACR F.3	PsACR H.13	H10_950	PsACR F.14	PsACR H.5	PsACR H.7
RUE <sup>b</sup> Apo	nd	5.01	nd	nd	37.04	3.27
RUE <sup>b</sup> Sex	nd	0.16	nd	nd	0.13	8.18
p-value	nd	3.68e <sup>-05</sup>	nd	nd	1.81e <sup>-06</sup>	2.76e <sup>-02</sup>
FDR	nd	7.36e <sup>-05</sup>	nd	nd	7.25e <sup>-06</sup>	3.68e <sup>-02</sup>

**Table 2.** Expression analysis of the apomixis-linked genes of *P. simplex* and their *P. notatum* homologs.

<sup>a</sup>Normalized number of reads in apomictic (A) and sexual (Sex) 454/Roche libraries. In brackets are the non-normalized read count values; \*Transcripts showing significant differences in abundance between apomictic and sexual libraries ( $\logFC > |2|$  and  $FDR < 0.05$ ); <sup>b</sup>Relative Units of Expression.

genes present in the BACs (Supplementary Table S3). A total of 34 isogroups were detected. Most isogroups (82.35%) showed 1–4 transcript variants (isotigs) that could represent specific allelic forms or splicing variants expressed during the reproductive development. Six genes (17.65%) showed 5 or more (up to 19) isotigs. Of the 36 isotigs, 19 corresponded to *Arabidopsis* orthologs. The detection of multiple isotigs for the same gene could be associated with the heterozygous and polyploid nature of both *P. simplex* and *P. notatum*.

Out of the 36 isotigs detected, six showed differential representation between the apomictic and sexual databases (Table 2): two of them (isotigs 9721 and 13832, homologous to genes PsACR|H.7 and PsACR|F.3, respectively) were less represented in the apomictic database and four (isotigs 9764, 20690, 10887 and 27671, homologous to genes PsACR|H.5, PsACR|H.13, PsACR|F.14, and H10\_950, respectively) were less represented in the sexual database (Table 2). Isotigs 13832 and 9764 are significantly overexpressed in sexual and apomictic flowers, respectively ( $\logFC > |2|$  and  $FDR < 0.05$ ). Both transcripts deserve further analysis to prove their association with the corresponding phenotype. The case of gene PsACR|H.13 deserves a separate description. Three isotigs, namely 20690, 28985 and 43473 were identified for this gene (Supplementary Table S3). Of these, isotig 20690 is up regulated in sexual sample compared to the apomictic one (Table 2), whereas the other two did not differ for their expression between the two phenotypes. This expression pattern mirrored that of the RT-qPCR of *P. simplex*

according to which there is an allele specifically expressed in apomictics (PsACR|H.13\_1, Fig. 5d) and the other expressed in both phenotypes. This scenario is consistent with the presence in both species of apomixis-specific alleles together with others that are shared between the sexual and apomictic phenotypes.

Most of the genes contained in the apomixis-linked BACs of *P. simplex* are expressed in flowers of *P. notatum* regardless of their nature of pseudogene. Considering the limited number of genes analysed, there is a good agreement between the expression pattern of the apomixis-linked genes of *P. simplex* and their homologs of *P. notatum*.

## Discussion

As grass genomes are largely collinear and apomixis is spread across the family, several authors have hypothesized it might be controlled by the same set of genes wherever it occurs<sup>32</sup>. However, while the genomic regions controlling apomixis appear to be similar within genera<sup>15,33</sup>, the apomixis-controlling loci seem to be highly divergent between them<sup>34</sup>. Such interpretative framework is consistent with the observation that apomixis has originated many times independently in the grass family<sup>35</sup>, whereas it spreads among the several species of the same genus by intra or inter-specific hybridisation<sup>36</sup>. In any cases, these findings render the identification of the genetic determinants of apomixis difficult. In *P. simplex*, comparative sequence analysis of genes contained in apomixis-linked BACs clearly points to a genomic region that is syntenic among the five main reference grass species<sup>3,23</sup> and is located in a telomeric position on chromosome 12, 8, 3 and 4 of rice, *Sorghum*, *Setaria* and *Brachypodium*, respectively, and in centromeric area of maize chromosome 1. In particular, the *Setaria* chromosome 3, together with chromosome 7, originated through a series of translocations and inversions involving ancestral chromosomes similarly to what concerned the actual chromosomes 4, 5, 12 of rice and 6, 8, 9 of *Sorghum*<sup>23</sup>. An 840 Kb inversion was reported in this genomic area of *Setaria*, with respect to its orthologous regions of rice and *Sorghum*<sup>37</sup> and, similarly, a large-scale inversion was documented in the same region of *Brachypodium* (<http://www.gramene.org/>). Although the limited portion of the ACR analysed does not allow a generalization, we argue that the origin and structure of the apomixis locus of *Paspalum* shares several commonalities with other multi-gene complexes, such as that related to the Y-chromosome of dioecious plants. Above all, among them are repression of recombination, presence of TE and gene degeneration<sup>38</sup>. The Y-chromosome originated from autosomal chromosomes by initial suppression of recombination in the regions containing the sex controlling genes. Further expansion and rearrangement of the non-recombining Y-locus together with migration of male determining genes caused chromosome heteromorphism and dioecism<sup>39</sup>. Such expansion of recombinational suppression occurred stepwise across chromosomes producing a kind of stratification of the strength of recombination suppression<sup>40</sup>. Such structural stratification is recognizable in the chromosomes 12 of rice and 8 of *Sorghum* harbouring the syntenic group to the apomixis locus of *Paspalum*<sup>41</sup>. From a functional point of view, the evolution of Y-chromosome induces both the silencing of the female genes through the action of degenerated genes and, at the same time, the development of male-specific organs and function by the action of master functional genes<sup>39</sup>. In this sense, the female phenotype is recognized as the default state in some dioecious systems<sup>42</sup>, as well as the sexual phenotype is considered the default state in apomictic systems<sup>43</sup>. In a more general view, the ACR of *Paspalum* shares commonalities with operon-like gene clusters that in plants control complex traits such as those designated to the production of secondary metabolites mainly involved in plant defense<sup>44</sup>. Gene clusters are commonly defined as a set of two or more non homologous functionally related genes that share a close genomic neighbourhood<sup>45</sup>. Genes contained in a cluster can be transcribed independently or subjected to various extents of co-expression<sup>45</sup>. If clustered genes are organized as a single transcriptional unit they are defined as operon<sup>46</sup>. While operons are specific of prokaryotic genomes and likely originated by horizontal gene transfer<sup>47</sup>, genes clusters are reported for many eukaryotes including plants<sup>48</sup> and evolved *de novo* by initial gene duplication followed by neo- or sub-functionalization and genome rearrangements of various nature<sup>49</sup>. Among the several features common to gene clusters is their origin from subtelomeric dynamic regions characterized by high rates of gene rearrangements<sup>44</sup>. Furthermore, a segmental duplication followed by gene inversion and recruitment has been proposed as origin of gene clusters required for the synthesis of triterpenes in *Arabidopsis*<sup>50</sup>. As a consequence of these rearrangements, the genes within these clusters are coordinately expressed at the chromatin level and associated with repressive marks<sup>44</sup>.

Another point of similarity between the ACR of *Paspalum* and multi gene complexes is related to gene expression. Two of the apomixis-linked genes considered here (genes PsACR|H.5 and PsACR|H.13), together with the previously analysed *PsORC3*<sup>19</sup>, are expressed as sense and antisense transcripts in reproductively committed cell lineage indicating they might be regulated by the same promoter. Furthermore, at least for the apomixis-linked *PsORC3* allele, its repressive role on the expression of its sexual counterpart is evidenced<sup>19</sup>. Finally, recent studies showed that the parthenogenetic development of the embryo in *Paspalum* is superimposed over the sexual one by a mechanism mediated by DNA methylation indicating that the ACR might be subjected to a chromatin-mediated gene silencing mechanism<sup>51</sup>.

Conversely, TE accumulation is not a common feature between apomixis loci and multi gene complexes. First sequencing efforts of apomixis loci led to the observation that there was an unusual accumulation of TEs in these loci<sup>52</sup>. This fact suggested that TEs may act as a sink to sequester factors involved in sexual reproductive pathway and possibly causing apomixis<sup>53</sup>. Nowadays, as WGS of the most representatives of the grass clades become available, we can conclude that in *Paspalum* as well as in other natural apomictic systems<sup>54</sup>, the proportion of repeated elements between apomixis-linked BACs and that detected in the whole genome were not different<sup>24</sup>. Furthermore, an extensive repetitive structure associated with apomixis was confirmed to be dispensable to express the apomictic phenotype in *Hieracium*<sup>55</sup>.

To sum up, as in several multi gene complexes, the ACR of *Paspalum* originated from chromosomically unstable subtelomeric regions that experienced both large and small scale inversions. Hence, the rise of apomixis in *Paspalum* could be the consequence of: (i) a casual grouping of a series of sexual genes in the same genomic context during speciation and (ii) a polyploidization event that generated a divergent and recombinationally blocked





a minimum of three genotypes (biological replicates) for each phenotype which were processed individually. Total RNA was extracted from collected samples using the Spectrum™ Plant Total RNA Kit (Sigma-Aldrich) following the protocol provided by the manufacturer. The contamination of genomic DNA was avoided by a DNase I treatment using the On-Column DNase I Digestion Set (Sigma-Aldrich) Kit. The abundance and purity of RNAs were assessed using a NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific, Pittsburgh, PA). The integrity of extracted RNA samples was estimated by electrophoresis on a 0.8% agarose/1 × TAE gel containing 1 × SYBR Safe DNA stain (Life Technologies, Carlsbad, CA, USA). cDNA synthesis was performed starting from 400 ng of total RNA, by using the RevertAid First Strand cDNA Synthesis Kit (Thermo Scientific) following the supplier's instructions. For each of the 4 selected genes for Realtime PCR, a single primer combination (PsACR|H.gene number\_gfor/rev; Supplementary Table S4) was designed on CDS and used for PCR amplification on two apomictic and two sexual DNA samples. The relative amplicons were cloned in pGEM-T Easy vectors and sequenced bidirectionally using SP6/T7 primers. By aligning the resulting sequences with those of the related genes on the BACs, several SNPs were identified. On the basis of these SNPs, primers were designed to amplify each specific allelic variants. The expression of each gene was analysed by using up to three different primer combinations (Hgene number.1–3\_for/rev; Supplementary Table S4) designed to assay the expression of multiple alleles and/or splicing variants detected in apomictic and sexual cDNAs. Amplification reactions were performed using StepOne thermal cycler (Applied Biosystems), equipped with 96-well plate systems, and FAST SYBR green Master Mix reagent (Applied Biosystems). Three technical replicates were adopted for each amplification reaction. The amplification efficiency was calculated from raw data using OneStep Analysis software (Life Technologies). Relative amplification performance, expressed as fold change, was calculated with the  $\Delta\Delta C_t$  method<sup>63</sup> using cytidine deaminase gene (*PsCDA*; GeneBank accession no. AM400871)<sup>18</sup> as the internal control (housekeeping). Error bars indicate the standard error observed among the three biological replicates (Fig. 5). Statistical analysis of expression counts in Table 2 was carried out according to the False Discovery Rate method<sup>64</sup>. For *in situ* analysis approximately 20–30 ng of cDNA from flowers of an apomictic genotype were amplified using specific primer pairs for each of the genes PsACR|H.5 and PsACR|H.13 (PsACR|H.gene number\_ishfor/ishrev; Supplementary Table S4). The derived 792 bp (PsACR|H.13) and 916 bp (PsACR|H.5) amplicons were cloned into a pGEM-T Easy vector (Promega) and sequenced bidirectionally with SP6 and T7 primers to establish the direction of the insertion. Both probes were labeled using a Roche DIG RNA labeling kit (SP6/T7) and hydrolyzed into 150–200 bp fragments. Spikelets of sexual and apomictic *P. simplex* genotypes were collected at anthesis stage, fixed, and embedded in paraffin and used for *in situ* hybridisation, as reported by Siena *et al.*<sup>65</sup>. Detection was performed following the Roche DIG detection kit instructions using anti-DIG AP and NBT/BCIP as substrates. Fourteen sexual and 11 apomictic ovules were observed as hybridised with the antisense probe of the gene PsACR|H.5 and 10 sexual and 4 apomictic ovules with the sense probe of the same gene; 8 sexual and 11 apomictic ovules were observed as hybridized with the antisense probe of PsACR|H.13 and 8 sexual and 11 apomictic ovules with the sense probe of the same gene.

## Data Availability

Supplementary Table S2 provides information on the primers used for the RT-qPCR and *in situ* experiments. The sequence data of the contigs reported in this paper have been deposited in the GeneBank database (<https://www.ncbi.nlm.nih.gov/genbank>) [accession nos MH106546, MH106547, MH106548, MH106549].

## References

- Gaut, B. S. Evolutionary dynamics of grass genomes. *New Phytol.* **154**, 15–28 (2002).
- Bennezen, J. L. Patterns in grass genome evolution. *Curr. Opin. Plant Biol.* **10**, 176–181 (2007).
- Wang, X. *et al.* Genome alignment spanning major poaceae lineages reveals heterogeneous evolutionary rates and alters inferred dates for key evolutionary events. *Mol. Plant* **8**, 885–898 (2015).
- Davidson, R. M. *et al.* Comparative transcriptomics of three Poaceae species reveals patterns of gene expression evolution. *Plant J.* **71**, 492–502 (2012).
- Hamel, L. P., Sheen, J. & Séguin, A. Ancient signals: comparative genomics of green plant CDPKs. *Trends Plant Sci.* **19**, 79–89 (2014).
- Tian, C. F. *et al.* Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *P. Natl. Acad. Sci. USA* **109**, 8629–8634 (2012).
- Kim, C. *et al.* Comparative analysis of Miscanthus and Saccharum reveals a shared whole-genome duplication but different evolutionary fates. *Plant Cell* **26**, 2420–2429 (2014).
- Brady, S. M. & Provart, N. J. Web-queryable large-scale data sets for hypothesis generation in plant biology. *Plant cell* **31**, 1034–1051 (2009).
- Linnaeus, C. *Systema naturae*, 10th edn. Stockholm (1759).
- Ortiz, J. P. A. *et al.* Harnessing apomictic reproduction in grasses: what we have learned from Paspalum. *Ann. Bot.* **112**, 767–787 (2013).
- Nogler, G. A. Gametophytic apomixis in *Embryology of Angiosperms*, (Ed. Johri, B. M.) 475–518 (Springer-Verlag, 1984).
- Vielle-Calzada, J. P., Crane, C. F. & Stelly, D. M. Apomixis: the asexual revolution. *Science* **274**, 1322–1323 (1996).
- Martínez, E. J., Hopp, H. E., Stein, J., Ortiz, J. P. A. & Quarin, C. L. Genetic characterization of apospory in tetraploid Paspalum notatum based on the identification of linked molecular markers. *Mol. Breeding* **12**, 312–327 (2003).
- Pupilli, F., Labombarda, P., Cáceres, M. E., Quarin, C. L. & Arcioni, S. The chromosome segment related to apomixis in Paspalum simplex is homoeologous to the telomeric region of the long arm of rice chromosome 12. *Mol. Breeding* **8**, 53–61 (2001).
- Pupilli, F. *et al.* Comparative mapping reveals partial conservation of synteny at the apomixis locus in Paspalum spp. *Mol. Genet. Genomics* **270**, 539–548 (2004).
- Hojsgaard, D. H., Martínez, E. J., Acuña, C., Quarin, C. L. & Pupilli, F. A molecular map of the apomixis-control locus in Paspalum procurrens and its comparative analysis with other species of Paspalum. *Theor. Appl. Genet.* **123**, 959–971 (2011).
- Calderini, O. *et al.* Molecular cytogenetics and DNA sequence analysis of an apomixis-linked BAC in Paspalum simplex reveal a non pericentromere location and partial microcolinearity with rice. *Theor. Appl. Genet.* **112**, 1179–1191 (2006).
- Polegri, L., Calderini, O., Arcioni, S. & Pupilli, F. 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 (2010).

19. Siena, L. A. *et al.* An apomixis-linked ORC3-like pseudogene is associated with silencing of its functional homolog in apomictic *Paspalum simplex*. *J. Exp. Bot.* **67**, 1965–1978 (2016).
20. Devos, K. M. & Gale, M. D. Genome relationships: the grass model in current research. *Plant Cell* **12**, 637–646 (2000).
21. Soreng, R. J. *et al.* A worldwide classification of the Poaceae (Gramineae). *J. Syst. Evol.* **2**, 117–137 (2015).
22. Paterson, A. H. *et al.* The Sorghum bicolor genome and the diversification of grasses. *Nature* **457**, 551–556 (2009).
23. Portwood J. L. *et al.* MaizeGDB 2018: the maize multi-genome genetics and genomics database. *Nucleic Acids Res.* gky1046, <https://doi.org/10.1093/nar/gky1046> (2018)
24. Zhang, G. *et al.* Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat. Biotechnol.* **30**, 549–554 (2012).
25. International Brachypodium Initiative. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**, 763–768 (2010).
26. Giordano, A. *et al.* Gene discovery and molecular marker development, based on high-throughput transcript sequencing of *Paspalum dilatatum* poir. *PLoS ONE* **9**, e85050 (2014).
27. Conner, J. A. *et al.* Sequence analysis of bacterial artificial chromosome clones from the apospory-specific genomic region of *Pennisetum* and *Cenchrus*. *Plant Physiol.* **147**, 1396–1411 (2008).
28. Calderini, O. *et al.* Partial isolation of the genomic region linked with apomixis in *Paspalum simplex*. *Mol. Breeding* **28**, 265–276 (2011).
29. Labombarda, P. *et al.* An AFLP marker tightly linked to apomixis reveals hemizygoty in a portion of the apomixis-controlling locus in *Paspalum simplex*. *Genome* **45**, 513–519 (2002).
30. Caponio, I. & Quarini, C. L. El sistema genético de *Paspalum simplex* y de un híbrido interspecífico con *P. dilatatum*. *Kurtziana* **19**, 35–45 (1987).
31. Ortiz, J. P. A. *et al.* A reference floral transcriptome of sexual and apomictic *Paspalum notatum*. *BMC Genomics* **18**, 318 (2017).
32. Kellogg, E. A. Reproductive Systems in *The Families and Genera of Vascular Plants. Flowering Plants Monocots 13*, (ed. K. Kubitzki, K.) 93–101 (Springer International Publishing, 2015).
33. Akiyama, Y. Evolution of the apomixis transmitting chromosome in *Pennisetum*. *BMC Evol. Biol.* **11**, 289 (2011).
34. Ozias-Akins, P. Apomixis: developmental characteristics and genetics. *Crit. Rev. Plant Sci.* **25**, 199–214 (2006).
35. Ozias-Akins, P., Akiyama, Y. & Hanna, W. W. Molecular characterization of the genomic region linked with apomixis in *Pennisetum/Cenchrus*. *Funct. Integr. Genomics* **3**, 94–104 (2003).
36. Worthington, M. *et al.* A parthenogenesis gene candidate and evidence for segmental allopolyploidy in apomictic *Brachiaria decumbens*. *Genetics* **203**, 1117–1132 (2016).
37. Bennetzen, J. L. *et al.* Reference genome sequence of the model plant *Setaria*. *Nat. Biotech.* **30**, 555–561 (2012).
38. Pupilli, F. & Barcaccia, G. Cloning plants by seeds: inheritance models and candidate genes to increase fundamental knowledge for engineering apomixis in sexual crops. *J. Biotechnol.* **159**, 291–311 (2012).
39. Bergero, R. & Charlesworth, D. The evolution of restricted recombination in sex chromosomes. *Trends Ecol. Evol.* **24**, 94–102 (2008).
40. Charlesworth, D. Plant sex determination and sex chromosomes. *Heredity* **88**, 94–101 (2002).
41. Wang, X. Y., Tang, H. & Paterson, A. H. Seventy million years of concerted evolution of a homoeologous chromosome pair, in parallel, in major poaceae lineages. *Plant Cell* **23**, 27–37 (2011).
42. Janousek, B., Siroky, J. & Vyskot, B. Epigenetic control of sexual phenotype in a dioecious plant, *Melandrium album*. *Mol. Genet. Genomics* **250**, 483–490 (1996).
43. Koltunow, A. M. *et al.* Sexual reproduction is the default mode in apomictic *Hieracium* subgenus *Pilosella*, in which two dominant loci function to enable apomixis. *Plant J.* **66**, 890–902 (2011).
44. Boycheva, S., Daviet, L., Wolfender, J. L. & Fitzpatrick, T. B. The rise of operon-like gene clusters in plants. *Trends in Plant Sci.* **19**, 447–459 (2014).
45. Nützmann, H. W., Scaccocchio, C. & Osbourn, A. Metabolic gene clusters in eukaryotes. *Annu. Rev. Genet.* **52**, 159–183 (2018).
46. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* **3**, 318–356 (1961).
47. Rocha, E. P. C. The organization of the bacterial genome. *Annu. Rev. Genet.* **42**, 211–233 (2008).
48. Hurst, L. D., Pál, C. & Lercherer, M. J. The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **5**, 299–310 (2004).
49. Chu, H. Y., Wegel, E. & Osbourn, A. From hormones to secondary metabolism: the emergence of metabolic gene clusters in plants. *Plant J.* **66**, 66–79 (2011).
50. Field, B. *et al.* Formation of plant metabolic gene clusters within dynamic chromosomal regions. *Proc. Natl. Acad. Sci. USA* **108**, 16116–16121 (2011).
51. Podio, M. *et al.* A methylation status analysis of the apomixis-specific region in *Paspalum* spp. suggests an epigenetic control of parthenogenesis. *J. Ex. Bot.* **65**, 6411–6424 (2014).
52. Hand, M. L. & Koltunow, A. M. The genetic control of apomixis: Asexual seed formation. *Genetics* **197**, 441–450 (2014).
53. Koltunow, A. M. & Grossniklaus, U. Apomixis: a developmental perspective. *Annu. Rev. Plant Biol.* **54**, 547–574 (2003).
54. Okada, T. *et al.* Chromosomes carrying meiotic avoidance loci in three apomictic *Hieracium* subgenus *Pilosella* species share structural features with two monocot apomicts. *Plant Physiol.* **157**, 1327–1341 (2011).
55. Kotani, Y., Henderson, S. T., Suzuki, G., Johnson, S. D. & Okada, T. The LOSS OF APOMEIOSIS (LOA) locus in *Hieracium praealtum* can function independently of the associated can function independently of the associated large-scale repetitive chromosomal structure. *New Phytol.* **201**, 973–981 (2013).
56. Conner, J. A., Mookkan, M., Huo, H., Chae, K. & Ozias-Akins, P. A parthenogenesis gene of apomict origin elicits embryo formation from unfertilized eggs in a sexual plant. *Proc. Natl. Acad. Sci. USA* **112**, 11205–11210 (2015).
57. Garcia-Aguilar, M., Michaud, C., Leblanc, O. & Grimanelli, D. Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes. *Plant Cell* **22**, 3249–3267 (2010).
58. Ohad, N. *et al.* Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**, 407–415 (1999).
59. Parra, G., Blanco, E. & Guigo, R. GeneID in *Drosophila*. *Genome Res.* **10**, 511–515 (2000).
60. Supek, B. E., Huang, H., McGarvey, P., Mazumder, R. & Wu, C. H. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282–1288 (2007).
61. Youens-Clark, K. *et al.* Gramene database in 2010: updates and extensions. *Nucleic Acids Res.* **39**, 1085–1094 (2010).
62. Rhee, S. Y. *et al.* The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic acids research* **31**, 224–228 (2003).
63. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup> $\Delta\Delta$ CT method. *Methods* **25**, 402–408 (2001).
64. Benjamini, Y. & Hochodol, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B Met.* **57**, 289–300 (1995).
65. Siena, L. A., Ortiz, J. P. A., Leblanc, O. & Pessino, S. C. PnTgs1 expression during reproductive development supports a role for RNA methyltransferases in the aposporous pathway. *BMC Plant Biology* **14**, 297–307 (2014).



## Acknowledgements

This work was funded by the Ufficio Relazioni Internazionali del Consiglio Nazionale delle Ricerche, Italy (Laboratori Congiunti Bilaterali Internazionali CNR, Prot. 0005651) and the Ministero degli Affari Esteri, e della Cooperazione Internazionale, Direzione Generale per la Promozione del Sistema Paese; J.P.A. Ortiz received a Short Term Mobility grant for a visit from the Ufficio Relazioni Internazionali del Consiglio Nazionale delle Ricerche, Italy. *Paspalum notatum* 454/Roche transcriptome database was funded by Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT), Argentina, PICT-2014-1080; Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, Project: PIP 2015-2017 11220150100702CO and Universidad Nacional de Rosario (UNR), Argentina, Project: 1AGR271. G. Galla received a grant as a principal investigator from the “Scientific Independence of Young Researchers (SIR)” program of the MIUR entitled “Transcriptomic analysis of ovule-specific cell lineages to unveil the genetic and molecular bases of apomictic seed production in model species (code: RBS114K1ON)”. L.A. Siena, J.P.A. Ortiz, and S.C. Pessino are research staff members of CONICET, Argentina.

## Author Contributions

G.G. carried out the qPCR experimental work and elaborated the graphic artworks; H.B. supervised the BAC sequencing and performed the annotation; J.P.A.O. and S.C.P. collaborated to the *P. notatum* 454 library bioinformatic studies; F.P. and M.B. performed the comparative *in silico* mapping; L.A.S. performed the *in situ* analyses; G.B., J.P.A.O., M.B. and F.P. designed the experimental approach and wrote the paper. All authors critically revised and approved the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-39649-6>.

**Competing Interests:** The authors declare no competing interests.

**Publisher’s note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019