Comparative plastome genomics and phylogenomics of *Brachypodium*: flowering time signatures, introgression and recombination in recently diverged ecotypes

Rubén Sancho[1,2], Carlos P. Cantalapiedra[3], Diana López-Alvarez[1], Sean P. Gordon[4], John P. Vogel[4,5], Pilar Catalán[1,2] and Bruno Contreras-Moreira[2,3,6]

[1]Department of Agricultural and Environmental Sciences, High Polytechnic School of Huesca, University of Zaragoza, Huesca, Spain; [2]Grupo de Bioquímica, Biofísica y Biología Computacional (BIFI, UNIZAR), Unidad Asociada al CSIC, Saragossa, Spain; [3]Department of Genetics and Plant Breeding, Estación Experimental de Aula Dei-Consejo Superior de Investigaciones Científicas, Zaragoza, Spain; [4]DOE Joint Genome Institute, Walnut Creek, CA 94598, USA; [5]Department of Plant and Microbial Biology, University of California, Berkeley, CA 94720, USA; [6]Fundación ARAID, Zaragoza, Spain

Authors for correspondence:
Pilar Catalán
Tel: +34 974232465
Email: pcatalan@unizar.es

Bruno Contreras-Moreira
Tel: +34 976716089
Email: bcontreras@eead.csic.es

Received: *10 October 2016*
Accepted: *3 March 2017*

New Phytologist (2017)
doi: 10.1111/nph.14926

Key words: *Brachypodium distachyon–B. stacei–B. hybridum*, comparative ptDNA genomics, grass phylogenomics, intraspecific genealogy, nested dating analysis, plastid introgression and recombination.

Summary

- Few pan-genomic studies have been conducted in plants, and none of them have focused on the intraspecific diversity and evolution of their plastid genomes.
- We address this issue in *Brachypodium distachyon* and its close relatives *B. stacei* and *B. hybridum*, for which a large genomic data set has been compiled. We analyze inter- and intraspecific plastid comparative genomics and phylogenomic relationships within a family-wide framework.
- Major indel differences were detected between *Brachypodium* plastomes. Within *B. distachyon*, we detected two main lineages, a mostly Extremely Delayed Flowering (EDF+) clade and a mostly Spanish (S+) – Turkish (T+) clade, plus nine chloroplast capture and two plastid DNA (ptDNA) introgression and micro-recombination events. Early Oligocene (30.9 million yr ago (Ma)) and Late Miocene (10.1 Ma) divergence times were inferred for the respective stem and crown nodes of *Brachypodium* and a very recent Mid-Pleistocene (0.9 Ma) time for the *B. distachyon* split.
- Flowering time variation is a main factor driving rapid intraspecific divergence in *B. distachyon*, although it is counterbalanced by repeated introgression between previously isolated lineages. Swapping of plastomes between the three different genomic groups, EDF+, T+, S+, probably resulted from random backcrossing followed by stabilization through selection pressure.

Introduction

Plastid DNA (ptDNA) has been widely used in inter- and intraspecific phylogenetic analyses in multiple species and populations of plants (Waters *et al.*, 2012; Ma *et al.*, 2014; Middleton *et al.*, 2014; Wysocki *et al.*, 2015). Phylogenetic dating of monocots and eudicots has also been based on ptDNA (Chaw *et al.*, 2004). Comparative genomics of whole plastid genomes has provided a way to detect and investigate genetic variation across seed plants (Jansen & Ruhlman, 2012). The proliferation of whole genome sequencing (WGS), which typically includes a substantial amount of plastid sequence, has provided large data sets which can be utilized to assemble and analyze plastomes (Nock *et al.*, 2011).

Brachypodium is a small genus in the family Poaceae that contains *c.* 20 species (17 perennial and three annual) distributed worldwide (Schippmann, 1991; Catalán & Olmstead, 2000; Catalán *et al.*, 2012, 2016a,b). The three annuals include two diploids (*B. distachyon* ($2n = 2x = 10$; $x = 5$), *B. stacei* ($2n = 2x = 20$; $x = 10$)) and their derived allotetraploid (*B. hybridum* ($2n = 4x = 30$; $x = 5 + 10$)). These three species had previously been considered cytotypes of *B. distachyon* (Catalán *et al.*, 2012). In addition to the large, overlapping distribution in their native circum-Mediterranean region (Catalán *et al.*, 2012, 2016a; López-Alvarez *et al.*, 2012, 2015), *B. hybridum* has naturalized extensively around the world.

The evolutionary relationship between *Brachypodium* and other grasses has been thoroughly studied (Catalán *et al.*, 1997; Catalán & Olmstead, 2000; Döring *et al.*, 2007). Most recent phylogenetic analyses place *Brachypodium* in an intermediate position within the Pooideae clade (Minaya *et al.*, 2015; Soreng *et al.*, 2015; Catalán *et al.*, 2016a,b). By contrast, only a few studies of intraspecific variation have been conducted in the genus *Brachypodium*, primarily focusing on *B. distachyon* (e.g. Filiz *et al.*, 2009; Vogel *et al.*, 2009; Mur *et al.*, 2011; Tyler *et al.*, 2016).

Brachypodium distachyon has been selected as a model plant for temperate cereals and biofuel grasses (Vogel *et al.*, 2010; Mur

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

et al., 2011; Catalán *et al.*, 2014; Vogel, 2016). Additionally, the *B. distachyon* complex has been proposed as a model system for grass polyploid speciation (Catalán *et al.*, 2014; Dinh Thi *et al.*, 2016). Nuclear and plastid genomes of the Bd21 ecotype of *B. distachyon* have been sequenced, assembled and annotated. The nuclear genome is 272 Mbp in size (Vogel *et al.*, 2010) and contains 31 694 protein-coding loci. The current plastid genome reference (NC_011032.1) is 135 199 base pairs (bp) long and encodes 133 genes (Bortiri *et al.*, 2008).

In parallel with the creation of the nuclear pan-genome of *B. distachyon* from 53 diverse lines (Gordon *et al.*, 2017), and the genome sequencing of its close congeners *B. stacei* and *B. hybridum* (*Brachypodium stacei* v1.1 DOE-JGI, http://phytozome.jgi.doe.gov/ and *B. hybridum* early access available through Phytozome), we isolated ptDNA sequences from WGS paired-end reads to assemble the corresponding plastomes. Our aim was to compile a large plastome data set and investigate the evolutionary relationships of the annual *Brachypodium* species within the grass phylogenetic framework. The specific objectives of this study were to: assemble, annotate and compare 57 plastomes of *B. distachyon*, *B. stacei* and *B. hybridum*; reconstruct and date the divergences within the *Brachypodium* lineages and a family-wide plastome phylogeny; infer the genealogical relationships within the studied accessions of *B. distachyon* and compare them with the nuclear genome genealogy; and investigate the potential existence of plastid introgression and recombination in *B. distachyon* ecotypes known to hold nuclear introgressions.

Materials and Methods

Plant materials

Brachypodium distachyon, *B. stacei and B. hybridum* ecotypes used in this work are inbred lines derived from our own collections (Vogel *et al.*, 2009; Mur *et al.*, 2011; Catalán *et al.*, 2012) and from the National Plant Germplasm System (NPGS) and Brachyomics collections (USDA and ABER lines; Vogel *et al.*, 2006; Garvin, 2007; Garvin *et al.*, 2008). Most ecotypes were originally collected in Spain, Turkey and Iraq (Supporting Information Table S1; Fig. 1) (Vogel & Hill, 2008; Filiz *et al.*, 2009; Mur *et al.*, 2011). Available plastome data from the main grass lineages were retrieved from GenBank (Table S2). Flowering time data were obtained from Gordon *et al.* (2017). Briefly, flowering time was measured as the number of days elapsed from the end of vernalization to inflorescence heading, in the growth chamber, and assigned to flowering time classes following Ream *et al.* (2014; see Table S3).

Plastid DNA automated assembly, annotation and validation

Illumina paired-end and mate-pair libraries from 53 *B. distachyon*, one *B. stacei* and three *B. hybridum* accessions were produced from total genomic DNA, isolated as described previously (Peterson *et al.*, 2000), randomly sheared and filtered to target fragment sizes of 250 bp and 4 kbp, using Covaris LE220 (Covaris, Woburn, MA, USA) and HydroShear (Genomic Solutions, Ann Arbor, MI, USA), respectively. The KAPA-Illumina library creation (KAPA Biosystems) and TruSeq v3 paired-end cluster kits were used for library construction. Sequencing was performed at the Joint Genome Institute on the Illumina HiSeq2000 sequencer, yielding reads of 76, 100 and 150 bp length.

We developed a pipeline, available at https://github.com/eead-csic-compbio/chloroplast_assembly_protocol, for the assembly and annotation of plastid genomes (Methods S1; Table S4; Fig. S1). Briefly, plastid reads were extracted from WGS data using DUK (http://duk.sourceforge.net), followed by quality control and error correction, with FAStQC v.0.10.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc),

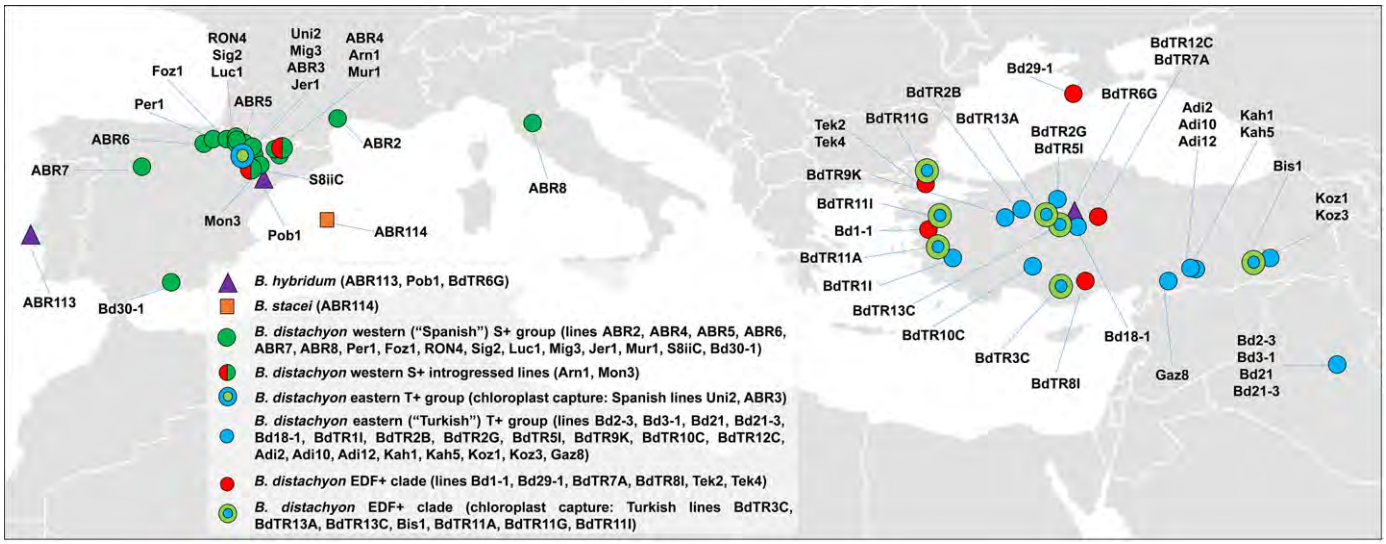

Fig. 1 Native circum-Mediterranean geographic distributions of the *Brachypodium distachyon*, *B. hybridum* and *B. stacei* ecotypes used in the plastome evolutionary and genomic analyses. Symbol and color codes for accessions are indicated in the chart. Accession numbers correspond to those indicated in Supporting Information Table S1.

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

Trimmomatic v.0.32 (Bolger *et al.*, 2014) and Musket v.1.0.6 (Liu *et al.*, 2013). Pass-filtered reads were then assembled with Velvet v.1.2.07 (Zerbino, 2010), SSpace Basic v.2.0 (Boetzer *et al.*, 2011) and GapFiller v.1.11 (Boetzer & Pirovano, 2012; Nadalin *et al.*, 2012).

This pipeline can be used to perform both *de novo* and reference-guided assemblies. Both strategies were performed with 55 out of 57 accessions; in most cases (46, see Table S5) the reference-guided approach produced fewer and longer contigs than *de novo* assemblies. Other parameters affecting assembly outcome were optimized, such as k-mer size or the number of input reads. Assembly errors were corrected with SEQuel v.1.0.2 (Ronen *et al.*, 2012), and by visual inspection of read mappings using IGV v.2.3.8 (Thorvaldsdóttir *et al.*, 2013).

Gene annotation was performed exhaustively for a single plastome of each species, and then transferred with custom scripts to the remaining plastid assemblies. The ptDNA genomes were compared with Organellar-Genome DRAW web version (Lohse *et al.*, 2013) and Circos v.0.69 (Krzywinski *et al.*, 2009). Typical plant plastomes show four main regions: large single-copy (LSC), first inverted-repeat (IRa), short single-copy (SSC) and second inverted-repeat (IRb), as sorted in the current Bd21 accession (NC_011032.1). Junctions between IR–LSC, LSC–IR, IR–SSC and SSC–IR regions, as well as main structural variations of *B. stacei* and *B. hybridum* plastomes, were confirmed by PCR amplification and Sanger sequencing (Table S6). The annotated plastomes of *B. distachyon*, *B. stacei* and *B. hybridum* ecotypes were deposited at ENA (European Nucleotide Archive) with accession numbers LT222229–LT222230 and LT558582–LT558636.

Intraspecific genealogy, haplotypic network, and genomic diversity and structure analyses

Plastomes from the 53 *B. distachyon* accessions (Table S1) were aligned using Mafft v.7.031b (Katoh & Standley, 2013); poorly aligned regions were removed with trimAl v.1.2rev59 (Capella-Gutiérrez *et al.*, 2009) using option automated1, which excludes columns after heuristically computing appropriate gap and similarity thresholds. However, most robust gaps were included in the final aligned data set and used in the phylogenetic maximum-likelihood (ML), Bayesian inference (BI) and dating Bayesian evolutionary analysis (Beast) approaches. The second inverted repeat region (IRb) accumulated most ambiguous nucleotides in our assemblies, probably due to biases in the pipeline (see histogram in Fig. 2). Considering that both repeats are essentially redundant in plastids, only IRa was included in subsequent phylogenetic analyses (Nock *et al.*, 2011; Middleton *et al.*, 2014; Saarela *et al.*, 2015). Alignments were revised and manually curated using Geneious v.8.1.4 (Kearse *et al.*, 2012).

ML and BI phylogenomic analyses were performed with RAxML v.8.1.17 (Stamatakis, 2014) and MrBayes v.3.2.4 (Ronquist & Huelsenbeck, 2003; Ronquist *et al.*, 2011), respectively. The generalized time-reversible plus gamma distribution plus proportion of invariant sites substitution model (GTR + G + I), selected by JModelTest v.2.1.7 based on the Akaike information criterion (Guindon & Gascuel, 2003; Darriba *et al.*, 2012), was imposed in the searches. In the ML search we computed 20 starting trees from 20 distinct randomized maximum parsimony (MP) trees and 1000 bootstrap replicates. In the BI search, two sets of four chains were run for 2 million generations, sampling trees and parameters every 100th generation. A 50% majority rule consensus tree was computed discarding the first 25% saved trees as 'burn-in'. All trees were midpoint rooted.

Haplotypic network analysis was conducted with the 53 *B. distachyon* plastome alignment after removing IRb and columns with missing data (*Ns*), both including and excluding indels. Statistical parsimony analysis was performed with TCS v.1.21 (Clement *et al.*, 2000), setting a maximum connection of 1000 steps. Haplotype polymorphism and genetic diversity statistics of the plastome data set, such as the number of segregating sites (S) and haplotypes (h), the haplotype diversity index (Hd), and the number of shared mutations (*shm*) and the average number of nucleotide differences (d) among the three intraspecific genetic groups retrieved from the phylogenomic analysis (see the 'Results' section) were calculated with DnaSP v.5 (Librado & Rozas, 2009).

Bayesian genomic clustering analysis was performed to infer the structure of the data, using a *B. distachyon* ptDNA data matrix of 298 mapped polymorphic positions, and to assign accessions' plastomes to the inferred groups using Structure v.2.3.4 (Pritchard *et al.*, 2000). The program was run for a number of potential genomic groups (K) from 1 to 6, imposing ancestral admixture and correlated allele frequencies priors. Ten independent runs with 100 000 burn-in steps followed by 1000 000 generations were computed for each K. The number of genetic clusters was estimated using Structure Harvester (Earl & vonHoldt, 2012), which identifies the optimal K based both on the posterior probability of the data for a given K and the ΔK (Evanno *et al.*, 2005). The potential existence of interplastome recombination in two introgressed ecotypes (see the 'Results' section) was further assessed through visual inspection of the mapped polymorphic alignments and through the recombination detection methods implemented in Rdp4 v.4.56 (Rdp, Geneconv, BootScan, MaxChi, Chimaera, SiScan, Lard, 3Seq; Martin *et al.*, 2015) and in OrgConv v.1.1 (Hao, 2010), using default settings in all cases.

Phylogenetic and molecular dating analyses

A grass plastome alignment was built including all *B. distachyon* ecotypes, one *B. stacei* ecotype and one *B. hybridum* ecotype (55 accessions; Table S1) plus the plastomes of 90 grasses (Table S2). ML analysis was performed with RAxML following the same steps indicated above. Pairwise Tamura–Nei (TN) raw genetic distances and pairwise TN patristic (RAxML-tree) distances were computed between all pairs of grass entries using Mega v.7.0.14 (Kumar *et al.*, 2016) and Geneious (Kearse *et al.*, 2012), respectively.

Divergence time estimations of the *Brachypodium* lineages were calculated within a family-wide dated phylogeny using a Bayesian nested dating partitioned approach (Pokorny *et al.*,

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

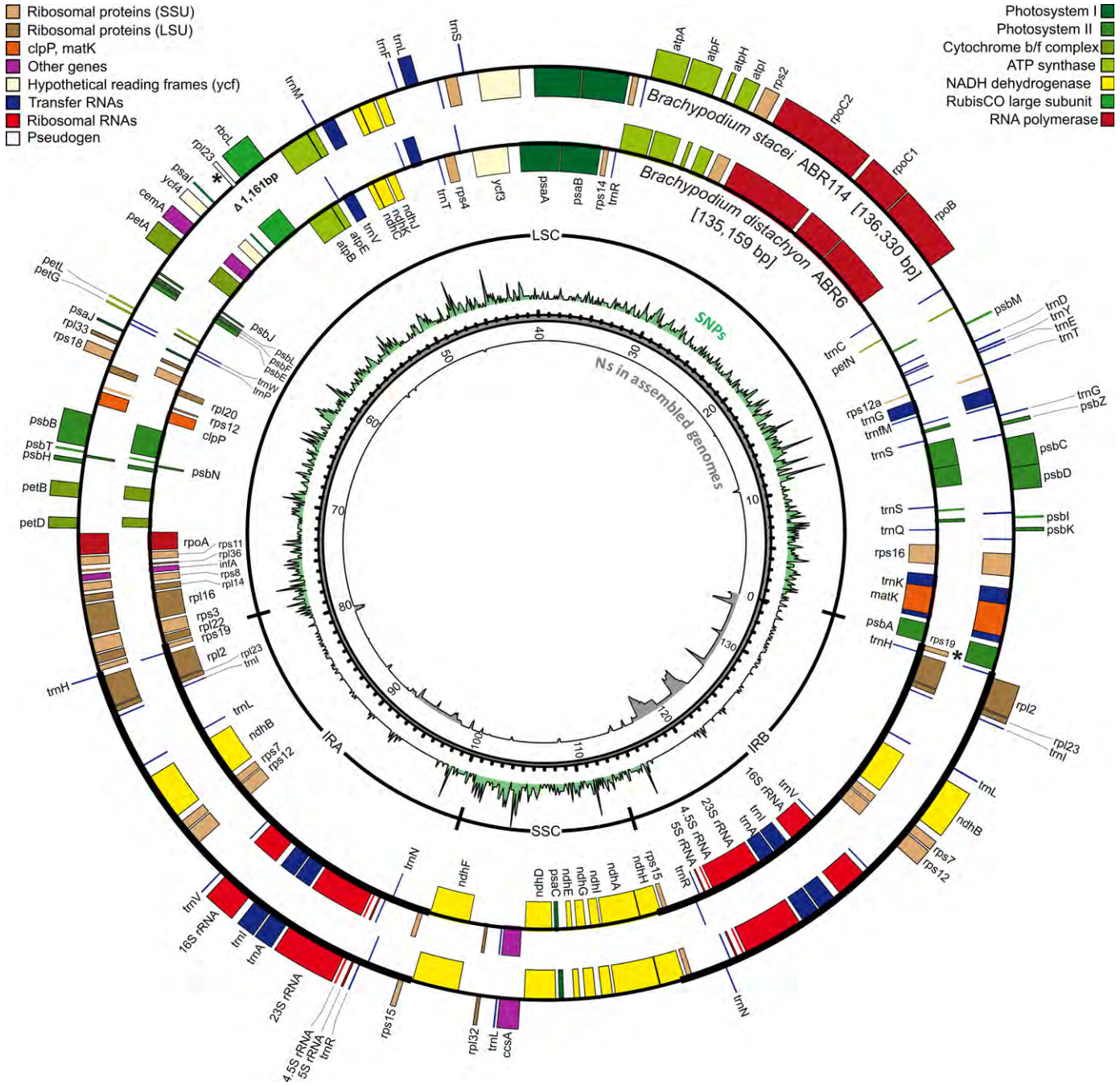

Fig. 2 Plastome maps of *Brachypodium distachyon* ABR6 (inner circle) and *B. stacei* ABR114 (outer circle). A 1161 bp insertion is shown in the *B. stacei* map (Δ, see upper-left quadrant), as well as a deletion of *rps*19 locus (*, see lower-right quadrant). Smaller inner circles and tracks correspond, respectively, to a map of plastome regions (LSC, SSC, IRA and IRB), a histogram of observed single nucleotide polymorphisms (SNPs) across all 57 aligned plastomes and a histogram of undetermined nucleotides, marked as N characters in the alignments.

2011; Mairal *et al.*, 2015) in Beast v.1.8.2 (Drummond *et al.*, 2012). Because there are no known fossil records of *Brachypodium*, a high-level more inclusive grass data set (93 samples = 90 grass species + 1 *B. distachyon* + 1 *B. stacei* + 1 *B. hybridum* accession, 110 370 bp in length, 22 489 polymorphic positions) was used to estimate divergence times within the *B. distachyon* ingroup (53 samples, 110 370 bp in length, 415 polymorphic positions). The grass tree was rooted with the ancestral species *Anomochloa marantoidea*. The estimated ages were drawn from deep-time calibrations imposed in the Poaceae partition and were used to constrain the molecular clock rate of the linked *B. distachyon* population-level data set and to calibrate the divergence time of its crown node. We estimated divergence times among the Poaceae lineage imposing GTR + G + I, lognormal relaxed clock and Yule tree models, a broad uniform distribution prior for the uncorrelated lognormal distribution (ucld)

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

mean (lower = 1.0E-6; upper = 0.1) and a default exponential prior for ucld standard deviation (SD). Calibrations were drawn from the compilation of grass fossils of Strömberg (2011) and from fossil-rich dating analyses of the grass family (Bouchenak-Khelladi *et al.*, 2010; Christin *et al.*, 2014). To accommodate uncertainties in the fossil records and fossil-based calibrations, we incorporated in the divergence time analysis normal distribution priors with mean and SD values of the normal distribution set for upper and lower dates of the geological period of the fossil, or the estimated divergence ages of the calibrated tree node, representing 5 and 95% quantiles of the distribution. We used two calibration points, imposing secondary age constrains for the crown nodes of Poaceae (normal prior mean = 90.0 Ma, SD = 1.0) and of the BOP (Bambusoideae, Oryzoideae, Pooideae) + PACMAD (Panicoideae, Arundinoideae, Chloridoideae, Micraioideae, Aristidoideae, Danthonioideae) clade (normal prior mean = 55.0 Ma, SD = 0.5), covering the age ranges of their respective fossil records and nodal age estimates. For the intraspecific *B. distachyon* data set we imposed a coalescent constant-size tree model. We ran 1000 000 000 Markov chain Monte Carlo (MCMC) generations in Beast with a sampling frequency of 1000 generations after a burn-in period of 1%. The adequacy of parameters was checked using Tracer v.1.6 (http://beast.bio.ed.ac.uk/Tracer), noting effective sample size (ESS) values > 200. Maximum clade credibility (MCC) trees were computed for the Poaceae and for the *B. distachyon* data sets after discarding 1% of the respective saved trees as burn-in.

Results

Structure, gene content and sequence in *B. distachyon*, *B. stacei* and *B. hybridum* plastomes

Assemblies were obtained for 57 plastomes. Forty-one contained ≤ 10 contigs, with an average longest contig length of 84 kbp and 176× depth coverage (Table S5). After scaffolding, 45 assemblies had ≤ 4 scaffolds with a mean plastome length of 124.5 kbp. Missing data ranged from 0 to 6%, with most plastomes (38) showing ≤ 0.1%. Most of the missing sequence was located in the IRb region, which was difficult to assemble because of its redundancy. The resulting *Brachypodium* plastomes were highly conserved in terms of synteny and gene number. Plastome lengths varied from 134 991 to 135 214 bp in *B. distachyon*, and between 136 326 and 136 330 bp in *B. stacei* and *B. hybridum* (Table S5).

Reference accession *B. distachyon* Bd21 (NC_011032.1; Bortiri *et al.*, 2008) and the *B. distachyon* Bd21 control (Bd21C, assembled and annotated in the current study) showed some differences (10 single nucleotide polymorphisms (SNPs) and 19 indels; Table S7a). These polymorphisms had read depth coverage ranging from 219× to 16 750× and were also confirmed in several of the other *B. distachyon* accessions (see Table S7a). While most of these polymorphisms lay in intergenic regions, some were located in protein coding genes such as *psb*A (one synonymous (Syn) mutation), *psb*K (one nonsynonymous (NSyn) mutation), *rpo*C2 (one Syn and one NSyn), *psa*A (one Syn), and also in one copy of the rRNA 16S locus.

Brachypodium distachyon plastomes showed the same gene arrangement and number (133) as Bd21C (Table S7a,b). In particular, they contained 76 protein coding genes, seven of which were duplicated genes, 20 nonredundant tRNAs (out of a total 38), four rRNAs in both inverted repeats, four pseudogenes (*trn*I, *rps*12a, *trn*T and *trn*I) and two hypothetical open reading frames (*ycf*). Several polymorphisms, mostly nonsynonymous, were detected in comparison to several grass plastomes. The most polymorphic loci were *rpo*C2 (70 SNPs), *ndh*F (59 SNPs), *rpo*B (31 SNPs) and *mat*K (30 SNPs), suggesting a significant correlation between SNP frequency and gene length ($R^2 = 0.68$, $P < 2.2e-16$; Table S7b).

Brachypodium stacei and *B. hybridum* accessions showed the same overall plastid genomic features as the *B. distachyon* accessions, with two exceptions (Fig. 2). They both contained a 1161 bp insertion between *psa*I and *rbc*L in the LSC region. This insertion was confirmed by read mapping (Fig. S2a,b), and it was also detected in homologous regions of several grasses (Table S7c). It corresponds to a coding sequence (CDS) fragment annotated as pseudogene *rpl*23 (Table S7d). The *B. stacei* and *B. hybridum* plastomes also contained a deletion of an *rps*19 copy between *psb*A and *trn*H in the IRb repeat, which was confirmed through PCR amplification and Sanger sequencing (Fig. S2c; Methods S1). The presence of these indels in the plastid genomes of the three *B. hybridum* accessions suggests that they were inherited from *B. stacei*-type maternal parents. Six polymorphisms were detected between the *B. hybridum* and *B. stacei* plastomes (Table S7e). These polymorphisms were located in intergenic regions, except for a Syn substitution in *psb*T (ecotype BdTR6G, *B. hybridum*) and an NSyn mutation in one copy of *rpl*23 (ecotype ABR113, *B. hybridum*).

Furthermore, a conceptual RNA-edited translation (U to C) was inferred in the *ndh*B gene of all the *B. hybridum* accessions and *B. stacei*, as well as in the *ndh*K gene of the *B. distachyon* Gaz8 accession.

Genealogy, haplotypic groups and diversity of *B. distachyon* plastomes

Beast (Fig. 3a), ML (Fig. S3a) and BI (Fig. S3b) analyses detected two main diverging lineages within *B. distachyon* that were structured phenotypically (Fig. 3a, Plastome tree; Table S3). One of them corresponded to an EDF+ clade, and the second to an S+T+ clade of remaining accessions, which showed a mixture of flowering phenotypes (Fig. 3a, Plastome tree; Table S3). The second clade was divided by further geographical substructure into a paraphyletic Western group ('Spanish' group – S+), including almost all ecotypes from Spain, France and Italy, and a monophyletic Eastern group ('Turkish' group – T+), including ecotypes from Turkey and Iraq, plus two Spanish accessions (ABR3, Uni2). While the divergences of the main lineages and sublineages had high bootstrap support (BS) and posterior probability support (PPS), the support of some internal branches of the S+ group was low (Figs 3a, Plastome tree, S3a,b).

Haplotypic network analyses detected 36 or 32 distinct ptDNA haplotypes, including or excluding indels, respectively (Table S8). A set of 298 nucleotide polymorphic sites extracted

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

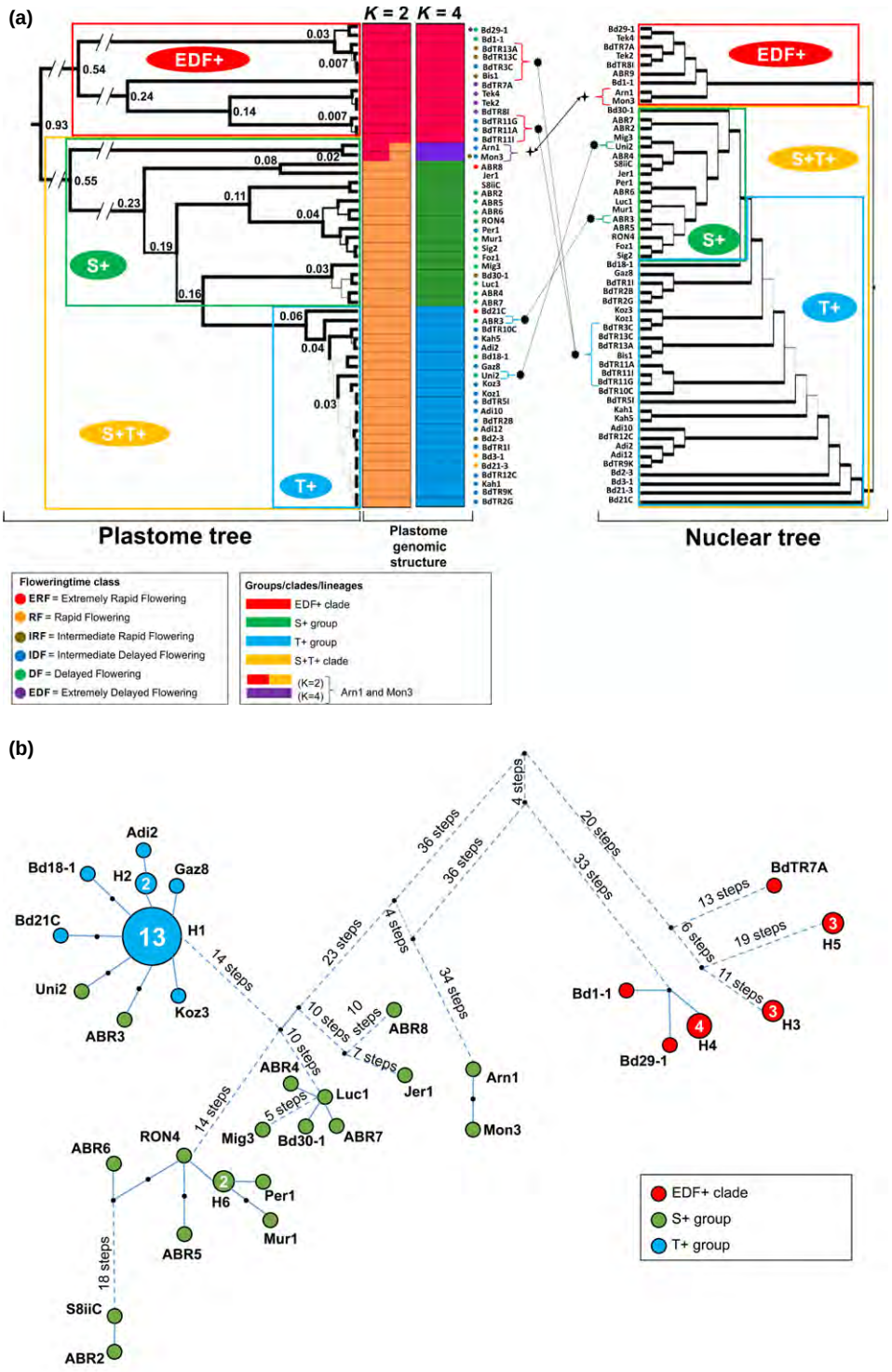

Fig. 3 Intraspecific evolutionary analysis of *Brachypodium distachyon* plastomes, including dated plastome genealogy, haplotypic network and genomic structure plots compared against the *B. distachyon* nuclear genealogical tree. (a) Beast nested dated chronogram of 53 *B. distachyon* plastomes showing estimated divergence times for below-species-level lineages (left). Datings (Ma) were inferred from calibrations obtained from above-species-level estimations. Thickness of branches indicates posterior probability support (thick branches, 0.95–1; intermediate branches, 0.90–0.94; thin branches, < 0.90). Genomic structure plots showing percentages of membership of plastomes' profiles to $K = 2$ and $K = 4$ genomic groups (center). Chloroplast capture and introgression events detected through topological contrast of the plastome and the nuclear trees (nuclear DNA (nDNA) tree from Gordon *et al.*, 2017) (right). Discontinuous and continuous lines mark potential chloroplast capture events and introgression events, respectively. Color codes for flowering time class groups and phylogenetic groups are indicated in the respective charts. Flowering time class groups are classified according to Ream *et al.* (2014) (see Supporting Information Table S3). (b) Haplotypic statistical parsimony network constructed with the *B. distachyon* plastomes using TCS. Dots represent mutation steps; numbers of mutation steps are indicated on branches. Color codes for clusters are indicated in the chart.

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

from the full *B. distachyon* plastome alignment confirmed the occurrence of 32 distinct ptDNA haplotypes; six haplotypes were shared by different accessions (H1, 13; H2, two; H3, three; H4, four; H5, three; H6, two) and 26 haplotypes were unique (Table S8). The TCS analysis clustered the 32 haplotypes into six groups (Fig. 3b), matching the structure observed in the genealogical ptDNA tree (Fig. 3a, Plastome tree). The haplotypic network was fully resolved except for one internal loop. The EDF+ haplotypes were separated from the cluster of S+ group and T+ group haplotypes by 59 and 74 step mutations, respectively. Within the EDF+ group there were two highly isolated clusters separated by 57 steps, one including only Turkish accessions (BdTR7A, H3, H5) and the second including Turkish and eastern European accessions (H4, Bd1-1, Bd29-1). The isolated Spanish Arn1+ Mon3 accessions of the S+T+ group showed an internal loop connecting its haplotypes with those of the EDF+ group (70 steps) and those of the remaining accessions of the S+T+ group (61 steps). Within the core S+T+ group, haplotypes clustered into four relatively close clusters, three of them including only accessions from the West (Spain, France and Italy), and the fourth cluster including mostly accessions from the East (Turkey, Iraq, plus Uni2 and ABR3) (Fig. 3b).

Plastome genomic diversity was variable within *B. distachyon* accessions (number of segregating sites (S) = 298, haplotypes (h) = 32, haplotype diversity index (Hd) = 0.933), and especially within the S+ (S = 137, h = 17, Hd = 0.993) and EDF+ (S = 107, h = 6, Hd = 0.846) groups (Table 1a). Our analyses indicated that the T+ group was less variable (S = 12, h = 9, Hd = 0.658) than the others. Diversity $\theta\pi$ values were not significantly different among groups. The S+ and T+ groups showed the lowest average number of nucleotide differences (d = 33.970), reflecting their close genomic affinities. By contrast, the EDF+ group showed the highest nucleotide differences to any of them (EDF+ – S+, d = 112.632; EDF+ – T+, d = 112.790), although it also shared six polymorphisms with the S+ group (EDF+ – S+, shm = 6) (Table 1b).

When the *B. distachyon* plastome genealogy was compared to an SNP-based nuclear pan-genome genealogy generated in our parallel study (Fig. 3a, Nuclear tree; Gordon *et al.*, 2017), the plastome tree revealed 11 cases of potential chloroplast capture and introgression. Seven cases (BdTR11A, BdTR11I, BdTR11G, BdTR13A, BdTR13C, BdTR3C, Bis1) corresponded to nuclear T+ ecotypes nested within the plastid EDF+ clade, two cases (ABR3, Uni2) to nuclear S+ ecotypes nested within the plastid T+ group, and two cases (Arn1, Mon3) to introgressed nuclear EDF+ ecotypes nested (and introgressed) within the plastid S+T+ clade (Fig. 3). All these cases suggest the existence of gene flow between the most diverged *B. distachyon* lineages. The STRUCTURE search further confirmed the potential 'admixed' nature of the Arn1 and Mon3 plastomes. The Bayesian structure analysis selected two optimal plastome groups with respect to second-order rate of change of the log probability of data between successive K values for a particular K (ΔK), the best ΔK = 2 corresponded to the EDF+ and S+T+ clades, with individual haplotypes showing high percentages of membership (> 95%) to their respective groups except the Arn1 and Mon3 haplotypes which showed similar percentages (40–60%) to both groups (Fig. 3a, plastome structure; Table S9). The next optimal grouping was for ΔK = 4; in this partition EDF+, S+ and T+ haplotypes clustered separately and the Arn1 and Mon3 haplotypes formed an independent group (all memberships > 95%). None of the recombination methods assayed in RDP4 and ORGCONV detected significant recombination in our data set; however, visual inspection of the polymorphic data matrix detected potential micro-recombination events in Arn1 and Mon3 (Fig. S4). Both haplotypes showed a large part of their sequences (polymorphic positions 1–225) as being similar to S+T+ sequences, and a small part of them (polymorphic positions 226–230) similar to EDF+ sequences. Polymorphic positions 1–237, 238–245 and 246–298 were located in the LSC, IR and SSC regions, respectively (Figs 2, S4).

Table 1 (a) Chloroplast haplotype diversity analysis of *Brachypodium distachyon* ecotypes and genomic groups (EDF+, S+, T+); group size and chloroplast haplotype diversity parameters: (b) pairwise estimates of the number of shared mutations (above diagonal) and the average number of nucleotide differences (below diagonal) between genomic groups

(a) Genomic groups	*N*	*S*	*h*	*Hd*	θπ
EDF+	13	107	6	0.846	12.780 (3.872–31.128)
S+	18	137	17	0.993	12.388 (3.804–30.837)
T+	22	12	9	0.658	12.683 (3.784–28.087)
B. distachyon (all ecotypes)	53	298	32	0.933	12.442 (4.218–28.245)

(b)	*shm*		
d	EDF+	S+	T+
EDF+	—	6	0
S+	112.632	—	0
T+	112.790	33.970	—

Plastid phylogenomics and divergence time estimations of Poaceae and *B. distachyon* lineages

ML (Fig. S5a,b) and BI (Fig. S5c,d) phylogenomic analysis of the grass plastome data set (Table S2) placed the monophyletic *Brachypodium* lineage in an intermediate and strongly supported diverging position within the Pooideae clade. *Brachypodium* was resolved as sister to the recently evolved core pooid clade, whereas the close Diarrheneae (*Diarrhena*) lineage was sister to the *Brachypodium* + core clade. Relationships among successively diverging basal Pooideae (Brachyelytreae, Phaenospematae, Meliceae, Stipeae) and BOP (Bambusoideae, Oryzoideae) and PACMAD (six Panicoideae species) lineages were congruent with previous studies; most bifurcations in the topology showed strong BS and PPS values. Within *Brachypodium*, the *B. stacei* clade (formed by *B. stacei* and the stacei-like *B. hybridum* plastomes) was resolved as sister to the *B. distachyon* clade. The latter lineage showed the divergence of the strongly supported EDF+ and S+T+ clades (Fig. S5a,c).

Both plastome raw pairwise genetic distances and pairwise patristic (RAXML tree) distances (Table S10; Fig. 4) supported the intermediate evolutionary position of *Brachypodium* within

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

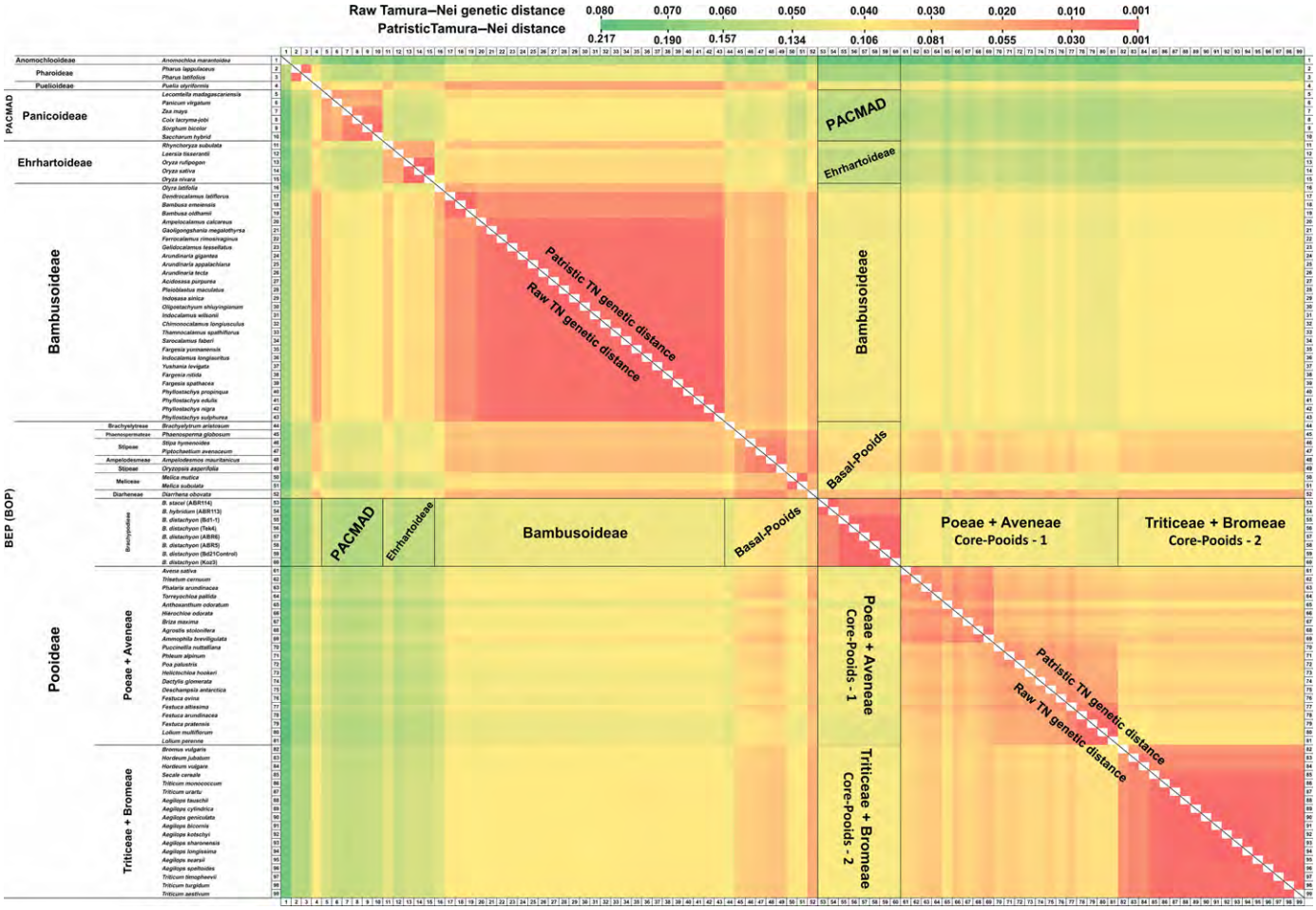

Fig. 4 Color-coded matrices of pairwise Tamura–Nei (TN) genetic distances between the plastome sequences of 99 Poaceae species and three *Brachypodium* (*B. distachyon*, *B. stacei*, *B. hybridum*) species. Below diagonal, pairwise raw TN genetic distances; above diagonal, pairwise phylogenetically based patristic TN genetic distances (computed on the RAxML tree, see Supporting Information Fig. S5b). Color-associated distance values are indicated in the chart.

the Pooideae clade (Fig. S5a–d). Moreover, Tamura–Nei (raw) genetic and patristic distances indicated a closer relationship of Brachypodieae to more ancestral basal pooid lineages (e.g. smaller genetic/patristic distances to Stipeae and Phaenospermatae than to recently evolved core pooid lineages (Triticodae, Poodae)) (Table S10; Fig. 4). They also revealed its closest relatedness to its evolutionarily nearest relative Diarrheneae. Distances of Brachypodieae to some Poodae lineages (e.g. Loliinae, Anthoxanthiinae) were similar to those observed to less closely related (e.g. Bambusoideae, Oryzeae (*Rhynchorhiza*)), or even much less closely related Puelioideae (*Puelia*) lineages (Table S10; Fig. 4).

The BEAST ptDNA MCC tree yielded the same topology of Poaceae (Figs 5, S6a) as that of the ML and BI trees (Fig. S5a–d). The dating analysis inferred intermediate Early Oligocene divergence times for the stem nodes of the Diarrheneae (31.9 Ma) and Brachypodieae (30.9 Ma) lineages, and divergence ages ranging from the more ancestral Mid- to Late Eocene splits of the basal pooids (Brachyelytreae, 44.2 Ma; Phaenospermatae, 38.4 Ma; Meliceae, 36.7 Ma; Stipeae, 35.3 Ma) to the recent Late Oligocene–Early Miocene splits of the core pooids (crown, 27.8 Ma; Poodae, 23.9 Ma; Triticodae, 17.6 Ma) lineages. A Mid- to Late Miocene age (10.1 Ma) was estimated for the *B. stacei*/*B. distachyon* split and a recent Mid-Pleistocene age (0.9 Ma) for the split of the most recent common ancestor (MRCA) of *B. distachyon* (Figs 5, S6a). According to our nested dating analysis, intraspecific divergences within *B. distachyon* occurred very recently, during the last half million years (e.g. EDF+ and S+T+ splits, 0.55 Ma; Figs 3a, Plastome tree, S6b).

Discussion

The plastid genomes of *Brachypodium*

Our study allowed us to construct the first large-scale intraspecific plastome analysis of a grass for the model species *B. distachyon* and a comparative genomics analysis with its close congeners

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

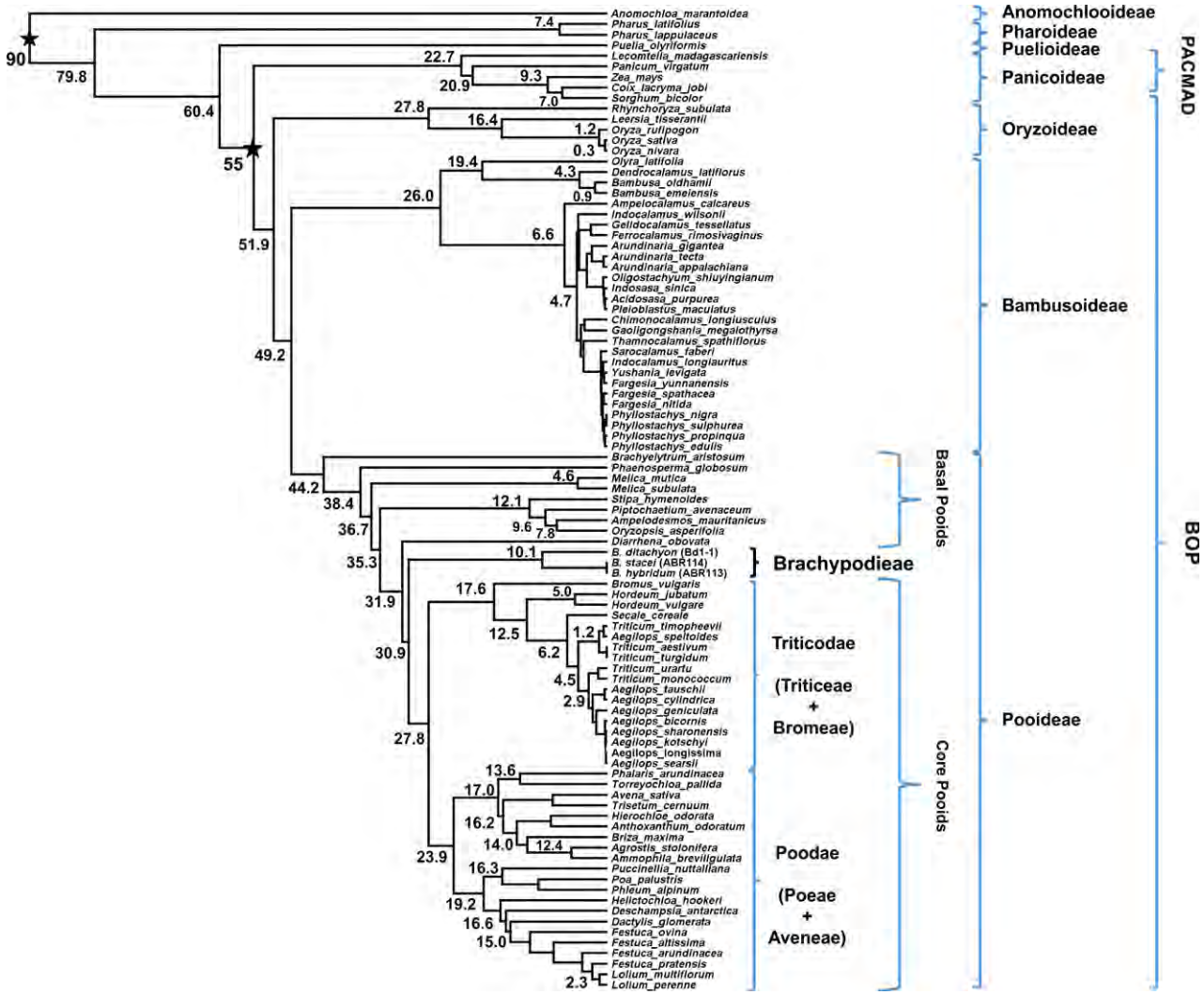

Fig. 5 Beast nested dated chronogram of 93 grass plastomes showing estimated divergence times and posterior probability support values for above-species-level lineages. Stars indicate nodal calibration priors (ages) for the Poaceae and BOP + PACMAD clades. Line thickness indicates posterior probability support, which was > 0.97 in all branches.

B. stacei and *B. hybridum* (Fig. 2; Table S5). We detected two main indels between *B. distachyon* and *B. stacei*/*B. hybridum* plastomes (Fig. S2), and no structural changes but a total of 415 polymorphisms (298 without indels) among the 53 *B. distachyon* ecotypes (Table S7a,b). A 1161 bp insert and the deletion of one copy of the *rps*19 gene, discovered in both the *B. stacei* and the *B. hybridum* ecotypes, indicates that the former is probably the maternal diploid plastome donor of the *B. hybridum* accession used in this study, which is consistent with previous findings reporting *B. stacei* as the maternal progenitor of most, but not all, wild *B. hybridum* populations (López-Alvarez *et al.*, 2012). The scarce number of polymorphisms (six) found in the *B. hybridum* as compared to the *B. stacei* plastome (Table S7e) indicates either that the *B. hybridum* plastome has remained almost intact since the formation of *B. hybridum* or that there has been continuous gene flow from *B. stacei* to *B. hybridum* (e.g. in Pleistocene–Holocene times, after the dated split of *B. distachyon* parent; Figs 3a, S6b).

The 1161 bp insert found in the *B. stacei*/*B. hybridum* plastomes contains an *rpl*23 pseudogene of 225 bp located around position 56 335 bp (Table S7c; Figs 2, S2a,b). The presence of an *rpl*23 pseudogene in this region has been reported in several monocots and in a large number of grasses, with insert sizes ranging from 40 to 243 bp (Morris & Duvall, 2010), whereas other authors have detected a functional *rpl*23 copy in *Agrostis stolonifera* (NC_008591) and *Sorghum bicolor* (NC_008602) (Saski *et al.*, 2007). In this study, all the assessed *B. distachyon* plastomes lack the insert and show two annotated *rpl*23 functional copies and no pseudogene, whereas the *B. stacei*/*B. hybridum* plastomes have also two functional *rpl*23 copies plus the *rbc*L-*psa*I insert *rpl*23 pseudogene (Table S7c; Fig. 2a,b).

In monocots, the *trn*H-*rps*19 cluster is located near the junctions of LSC and the two inverted repeats (Borsch & Quandt, 2009, and references therein). Wang *et al.* (2008) described three types of IR–LSC junctions based on the organization of their flanking genes in several monocots and dicots. While the studied

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

B. distachyon plastomes fit the type III class typical of monocots (*trn*H-*rps*19 clusters contain the *rps*19 gene in both IRs), the *B. stacei*/*B. hybridum* plastomes show a single *rps*19 copy near the *rpl*22 functional LSC flanking gene, and the lack of the second *rps*19 copy (Fig. S2c), fitting best the type I junction model. The type I class is mostly found in basal angiosperms, Magnoliids and Eudicots (Wang *et al.*, 2008). Thus, the *rbc*L-*psa*I insert *rpl*23 pseudogene and the *trn*H-*rps*19 type I cluster constitute landmarks of the more ancestral *B. stacei* chloroplast genome.

Flowering time divergence, chloroplast capture and introgression in *B. distachyon* plastomes

Our genealogical and haplotypic network analyses have detected a main split of two intraspecific *B. distachyon* lineages (EDF+ vs S+T+) that are not primarily connected with geography but with flowering time phenotypic traits, although the second clade is further separated into two geographically disjunct western (S+) and eastern (T+) circum-Mediterranean groups (Figs 3a, plastome tree, S3a,b; Table S3). Although our geographic sampling is biased towards Spain, Turkey and Iraq, these regions span the entire native distribution area of *B. distachyon* (López-Alvarez *et al.*, 2012, 2015), and our results are comparable with those obtained by Tyler *et al.* (2016) using nuclear SNPs from genotyping-by-sequencing (GBS) data. Haplotypic divergence data confirm the isolation of the EDF+ clade from the S+ and T+ genomic groups and similar haplotypic diversity values of EDF+ and S+ (Table 1a,b). Intraspecific evolutionary studies of organisms tend to recover the spatiotemporal divergence of populations, which are usually associated with a geographical distribution, detecting a typical isolation-by-distance (IBD) pattern (Wright, 1943; Jenkins *et al.*, 2010). However, long-distance dispersal events and biological and ecological traits have influenced the population structure in *B. distachyon* (Vogel *et al.*, 2009; Mur *et al.*, 2011; López-Alvarez *et al.*, 2012; Tyler *et al.*, 2016). Here, we have detected a strong influence of flowering time in the ancestral divergence of the *B. distachyon* EDF+ and S+T+ lineages, as several EDF+ lines (BdTR7A, BdTR8I, Tek2, Tek4) flower considerably later than the S+T+ lines (Fig. 3a, Plastome tree; Table S3). Our parallel nuclear pan-genome study of *B. distachyon* has also recovered a main EDF+ clade, including all the extremely delayed flowering (EDF) lines of our plastome clade (Fig. 3a, nuclear tree), and recent population genetic studies of *B. distachyon* based on GBS data (Tyler *et al.*, 2016) have also found it. Thus, flowering time is a main biological factor controlling the divergence of the major annual *B. distachyon* clades since the late Pleistocene (0.9–0.55 Ma) (Figs 3a, Plastome and Nuclear trees, S6b). Flowering time has been extensively studied in temperate cereals (barley, wheat), which have winter and spring races governed by vernalization and photoperiod requirements analogous to the delayed and rapid flowering phenotypes observed in *B. distachyon* (Vogel & Bragg, 2009; Schwartz *et al.*, 2010; Colton-Gagnon *et al.*, 2014; Ream *et al.*, 2014; Woods *et al.*, 2014). Although inflorescence heading-date phenotypic data in this work come from growth chamber experiments (Gordon *et al.*, 2017), they parallel the outcomes observed in field experiments (e.g. variation in flowering time was detected between winter-annual and spring-annual wild accessions of *B. distachyon*; Manzaneda *et al.*, 2015; A. J. Manzaneda, pers. comm.). Our study highlights the evolutionary importance of flowering time in driving intraspecies divergence.

It could be expected that flowering time isolation would create a barrier to gene flow, which might ultimately lead to (micro)speciation (Silvertown *et al.*, 2005; Lowry *et al.*, 2008; Noirot *et al.*, 2016). However, our study has demonstrated that it is not the case in *B. distachyon*, where frequent introgressions have apparently occurred between the EDF+ and S+T+ clades during the last half million years (Figs 3a, S6b). Topological comparison between the plastome and nuclear trees (Fig. 3a) indicated that seven Turkish accessions (BdTR11A, BdTR11I, BdTR11G, BdTR13A, BdTR13C, BdTR3C, Bis1) that are deeply and strongly nested within the eastern group of the S+T+ clade in the nuclear tree are, however, deeply and strongly nested within the eastern group of the EDF+ clade in the plastome tree and network. Similarly, two Spanish accessions (ABR3, Uni2) deeply nested within the western group of the S+T+ clade in the nuclear tree are instead nested within the eastern group of the S+T+ clade in the plastome tree, but with low support (Figs 3a,b, S3a,b). Moreover, two Spanish accessions (Arn1, Mon3) which are part of the EDF+ clade in the nuclear tree are nested within the S+T+ clade in the plastome tree, and form a loop with an EDF+ subgroup in the plastome haplotypic network (Figs 3a,b, S3a,b). Interestingly, genomic structure analyses indicated considerable introgression signals in the Arn1 and Mon3 nuclear and plastid genomes, whereas the seven Turkish accessions and the two Spanish accessions do not show evidence of introgression to the other genetic group in their chloroplast or nuclear genomes (Figs 3a, plastome genomic structure, S4). These results support the occurrence of two different introgression events. An early introgression of an S+T+ Spanish lineage with a member of the EDF+ clade could have originated the admixed ancestor of the Arn1/Mon3 lineage that kept most of its maternal S+T+ plastome but two-thirds of its paternal nuclear EDF+ genome over generations (Gordon *et al.*, 2017). According to our dating analysis, this introgression probably occurred in Ionian–Upper Pleistocene times (0.55–0.02 Ma) (Figs 3a, S6b). By contrast, more recent late Pleistocene–Holocene (0.025–0.007 Ma) introgressions between geographically close Turkish EDF+ and S+T+ lines probably resulted in the seven lines that show chloroplast capture for their intact EDF+ plastomes in combination with their intact paternal nuclear S+T+ genomes, the latter probably originating through repeated back-crossing to paternal S+T+ individuals (Figs 3a, S4, S6b). A similar late Pleistocene–Holocene scenario of introgressions and repeated backcrossing, but between geographically distant S+ and T+ lines, probably resulted in the two Spanish lines that show chloroplast capture for their intact T+ maternal plastomes and their paternal nuclear S+ genomes (Figs 3a, S4). These observations support previous evidence of long-distance dispersal of eastern *B. distachyon* seeds to the West across the Mediterranean basin (cf. López-Alvarez *et al.*, 2012, 2015). Additionally, Uni2 shows a significantly smaller inbreeding coefficient ($F_{is} = 0.48$) than the remaining

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

highly selfed *B. distachyon* accessions (median $F_{is} = 0.88$) (Gordon *et al.*, 2017), suggesting than the reduced F_{is} might reflect recent potential interpopulation crosses.

Our analyses also point towards the potential existence of heteroplasmic recombination in the Arn1 and Mon3 plastomes (Fig. 3a, plastome structure; Table S9). Also, visual inspection of the polymorphic data matrix identified a large proportion of their plastomes as S+T+ type and a smaller proportion of them (e.g. micro-recombinations) as EDF+ type (Fig. S4). Natural chloroplast heteroplasmy originated from biparentally inherited chloroplasts is infrequent in angiosperms (but see Mogensen, 1996). While plastid inheritance is considered to be mostly maternal (Jansen & Ruhlman, 2012), evidence of ptDNA biparental inheritance and of introgression has been documented in flowering plants (Mason-Gamer *et al.*, 1995; Mogensen, 1996), including potential low levels of sexual organelle recombination (Greiner *et al.*, 2015). For instance, heteroplasmy and potential inter- or intraspecific recombination have been detected in the plastomes of the highly hybridogenous genus *Citrus* (Carbonell-Caballero *et al.*, 2015). Also, interspecific chloroplast recombination was observed after somatic cell fusion in *Nicotiana* (Medgyesy *et al.*, 1985). Our study reports the first case of potential intraspecific recombination between different plastome types in these two introgressed *B. distachyon* accessions.

Evolutionary placement of a model genus for both temperate and tropical grasses

The phylogenomic analysis of 145 grass plastomes allowed us to infer the phylogenetic placement of *Brachypodium* and to calculate its genetic and patristic distances to other grass lineages (Table S10; Figs 4, 5, S5a–d, S6a). The intermediate nesting of *Brachypodium* within the Pooideae clade and the relationships of the other Poaceae lineages agree with previous studies based on nuclear or plastid genes (Bouchenak-Khelladi *et al.*, 2008; Schneider *et al.*, 2011; Hochbach *et al.*, 2015; Soreng *et al.*, 2015) or whole plastome sequences (Saarela *et al.*, 2015). The sister but non-inclusive relationship of *Brachypodium* to the core pooid clade (Triticodae (Triticeae + Bromeae)/Poodae (Poeae + Aveneae)), originally proposed by Davis & Soreng (1993), was abandoned in favor of the inclusion of *Brachypodium* within the 'core pooids', a nontaxonomic but independently evolved natural group, in some recent analyses (Davis & Soreng, 2007; Saarela *et al.*, 2015; Soreng *et al.*, 2015). Our ML and BI analyses support the sister relationship proposed by Davis & Soreng (Fig. S5a–d) as well as divergence times intermediate between those of the basal ancestral pooids and the recently evolved core pooids (Figs 5, S6a). Additionally, our pairwise ptDNA genetic and patristic distances have further confirmed that *Brachypodium* is closer to some basal pooid lineages than to the core pooid lineages (Table S10; Fig. 4), corroborating similar results based on nuclear single copy genes (Minaya *et al.*, 2015). Also, our genetic and phylogenetically based patristic data indicate that *Brachypodium* is similarly close to some core pooid groups than to more distant Oryzoideae and Puelioideae lineages. The evolutionary placement of *Brachypodium* in the Poaceae supports its utility as a model system for the monocots, as has been recently manifested in functional genomic studies of regulation of vernalization and flowering time. *B. distachyon* shows either seasonal response to flowering mechanisms close to those of core pooid grasses adapted to cold and temperate climates (Fjellheim *et al.*, 2014), or new flowering repressor vernalization genes shared with basal pooids, other tropical and subtropical grasses and less related Musaceae and Arecaceae (Woods *et al.*, 2016). Under the sampling in this study, the isolated and 'bridging' intermediate position of *Brachypodium* within the Pooideae supports its value as a model genus for many types of grasses, particularly for bioenergy crops (Brkljacic *et al.*, 2011) from different grass subfamilies (e.g. *Miscanthus*, *Paspalum* (Panicoideae), *Thinopyrum* (Pooideae)).

Our estimated divergence times for the main Poaceae lineages (Oryzoideae, 52 Ma; Bambusoideae 49 Ma; Pooideae, 44 Ma) (Figs 5, S6a) are in agreement with those calculated by Bouchenak-Khelladi *et al.* (2010) and Christin *et al.* (2014) but slightly older than those estimated by Wu & Ge (2012). Our results support early Oligocene (32 Ma) and late Miocene (10 Ma) splits for the respective stem and crown nodes of *Brachypodium*, which are also slightly older than those calculated by Catalán *et al.* (2012), although the highest posterior density range intervals overlap in both studies. The relatively old divergence inferred for the annual *B. stacei* and *B. distachyon* lineages in the late Miocene contrasts with the very recent burst of the intraspecific *B. distachyon* lineages. The estimated time of the late radiation (0.9 Ma) is in agreement with the estimated age of *B. hybridum* (*c.* 1 Ma; cf. Catalán *et al.*, 2012), the allotetraploid derivative of crosses between *B. stacei* and *B. distachyon*. Thus, the two complementary dating analyses fit a Mid-Pleistocene scenario for the almost contemporary origins of both parent and hybrid species.

Conclusion

Our comparative genomic study of whole plastome sequences of *B. distachyon* and its close relatives allowed us to detect intraspecific introgressions and other associated evolutionary events (e.g. biparental plastome inheritance, heteroplasmy) that could not be detected with single genes. The observed plastome admixture that goes along with the nuclear genome admixture in the *B. distachyon* Arn1 and Mon3 lines, and the essential swapping of plastomes among the three different *B. distachyon* plastome groups (EDF+, S+, T+), probably resulted from random backcrossing followed by stabilization through selection pressure. The chloroplast genome of *B. distachyon* is much more constrained as compared to its nuclear genome as we do not observe variation in the plastome genes.

Acknowledgements

We thank Drs Daniel Woods and Weilon Hao for fruitful discussions about flowering time features of the studied *B. distachyon* accessions and potential plastome recombination events, respectively, and Dr Tim Langdon and three anonymous referees for valuable comments on an early version of the manuscript. P.C., B.C-M., R.S. and D.L-A. received funding from the Spanish

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

Ministry of Economy and Competitivity (Mineco) grant projects (CGL2012-39953-C02-01, CSIC13-4E-2490 and CGL2016-79790-P). B.C-M. was funded by Fundación ARAID. R.S. and D.L-A. were funded by their respective Spanish Mineco PhD fellowships. P.C., R.S. and D.L-A. were partially funded by a Bioflora grant cofunded by the Spanish Aragon Government and the European Social Fund. The work conducted by the US DOE Joint Genome Institute is supported by the Office of Science of the US Department of Energy under Contract no. DE-AC02-05CH11231.

Author contributions

B.C-M. and P.C. designed the experiment, S.P.G. and J.P.V. collected the data, R.S., C.P.C., D.L-A., B.C-M. and P.C. performed the analyses, R.S., C.P.C., B.C-M. and P.C. wrote the paper, S.P.G. and J.P.V. contributed to the writing. All authors read, commented and approved the paper. The authors declare no conflicts of interest.

References

Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**: 578–579.

Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biology* **13**: R56.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Borsch T, Quandt D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution* **282**: 169–199.

Bortiri E, Coleman-Derr D, Lazo GR, Anderson OD, Gu YQ. 2008. The complete chloroplast genome sequence of *Brachypodium distachyon*: sequence comparison and phylogenetic analysis of eight grass plastomes. *BMC Research Notes* **1**: 61.

Bouchenak-Khelladi Y, Salamin N, Savolainen V, Forest F, van der Bank M, Chase MW, Hodkinson TR. 2008. Large multi-gene phylogenetic trees of the grasses (Poaceae): progress towards complete tribal and generic level sampling. *Molecular Phylogenetics and Evolution* **47**: 488–505.

Bouchenak-Khelladi Y, Verboom GA, Savolainen V, Hodkinson TR. 2010. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. *Botanical Journal of the Linnean Society* **162**: 543–557.

Brkljacic J, Grotewold E, Scholl R, Mockler T, Garvin DF, Vain P, Brutnell T, Sibout R, Bevan M, Budak H *et al.* 2011. Brachypodium as a model for the grasses: today and the future. *Plant Physiology* **157**: 3–13.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Carbonell-Caballero J, Alonso R, Ibañez V, Terol J, Talon M, Dopazo J. 2015. A phylogenetic analysis of 34 chloroplast genomes elucidates the relationships between wild and domestic species within the genus *Citrus*. *Molecular Biology and Evolution* **32**: 2015–2035.

Catalán P, Chalhoub B, Chochois V, Garvin DF, Hasterok R, Manzaneda AJ, Mur LAJ, Pecchioni N, Rasmussen SK, Vogel JP *et al.* 2014. Update on the genomics and basic biology of *Brachypodium*. International *Brachypodium* Initiative (IBI). *Trends in Plant Science* **19**: 414–418.

Catalán P, Kellogg EA, Olmstead RG. 1997. Phylogeny of Poaceae subfamily Pooideae based on chloroplast *ndhF* gene sequences. *Molecular Phylogenetics and Evolution* **8**: 150–166.

Catalán P, López-Alvarez D, Bellosta C, Villar L. 2016a. Updated taxonomic descriptions, iconography, and habitat preferences of *Brachypodium distachyon*, *B. stacei*, and *B. hybridum* (Poaceae). *Anales del Jardín Botánico de Madrid* **73**: 1–14.

Catalán P, López-Alvarez D, Díaz-Pérez A, Sancho R, López-Herranz ML. 2016b. Phylogeny and evolution of the genus *Brachypodium*. In: Vogel JP, ed. *Genetics and genomics of* Brachypodium. Cham, Switzerland: Springer, 9–38.

Catalán P, Müller J, Hasterok R, Jenkins G, Mur LAJ, Langdon T, Betekhtin A, Siwinska D, Pimentel M, López-Alvarez D. 2012. Evolution and taxonomic split of the model grass *Brachypodium distachyon*. *Annals of Botany* **109**: 385–405.

Catalán P, Olmstead RG. 2000. Phylogenetic reconstruction of the genus *Brachypodium* P. Beauv. (Poaceae) from combined sequences of chloroplast *ndhF* gene and nuclear ITS. *Plant Systematics and Evolution* **220**: 1–19.

Chaw SM, Chang CC, Chen HL, Li WH. 2004. Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *Journal of Molecular Evolution* **58**: 424–441.

Christin PA, Spriggs E, Osborne CP, Strömberg CAE, Salamin N, Edwards EJ. 2014. Molecular dating, evolutionary rates, and the age of the grasses. *Systematic Biology* **63**: 153–165.

Clement M, Posada D, Crandall KA. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9**: 1657–1660.

Colton-Gagnon K, Ali-Benali MA, Mayer BF, Dionne R, Bertrand A, Do Carmo S, Charron JB. 2014. Comparative analysis of the cold acclimation and freezing tolerance capacities of seven diploid *Brachypodium distachyon* accessions. *Annals of Botany* **113**: 681–693.

Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* **9**: 772.

Davis JI, Soreng RJ. 1993. Phylogenetic structure in the grass family (Poaceae) as inferred from chloroplast DNA restriction site variation. *American Journal of Botany* **80**: 1444–1454.

Davis JI, Soreng RJ. 2007. A preliminary phylogenetic analysis of the grass subfamily Pooideae (Poaceae), with attention to structural features of the plastid and nuclear genomes, including an intron loss in GBSSI. *Aliso: A Journal of Systematics and Evolutionary Botany* **23**: 335–348.

Dinh Thi VH, Coriton O, Le Clainche I, Arnaud D, Gordon SP, Linc G, Catalan P, Hasterok R, Vogel JP, Jahier J *et al.* 2016. Recreating stable *Brachypodium hybridum* allotetraploids by uniting the divergent genomes of *B. distachyon* and *B. stacei*. *PLoS ONE* **11**: e0167171.

Döring E, Schneider J, Hilu KW, Röser M, Hilu W, Rserl M, Doringl E. 2007. Phylogenetic relationships in the Aveneae/Poeae complex (Pooideae, Poaceae). *Kew Bulletin* **62**: 407–424.

Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**: 1969–1973.

Earl DA, vonHoldt BM. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359–361.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611–2620.

Filiz E, Ozdemir BS, Budak F, Vogel JP, Tuna M, Budak H. 2009. Molecular, morphological, and cytological analysis of diverse *Brachypodium distachyon* inbred lines. *Genome* **52**: 876–890.

Fjellheim S, Boden S, Trevaskis B. 2014. The role of seasonal flowering responses in adaptation of grasses to temperate climates. *Frontiers in Plant Science* **5**: 431.

Garvin DF. 2007. *Brachypodium distachyon*: a new model system for structural and functional analysis of grass genomes. In: Varshney RK, Koebner RMD, eds. *Model plants and crop improvement*. Boca Raton, FL, USA: CRC Press, 109–123.

Garvin DF, Gu YQ, Hasterok R, Hazen SP, Jenkins G, Mockler TC, Mur LAJ, Vogel JP. 2008. Development of genetic and genomic research resources for *Brachypodium distachyon*, a new model system for grass crop research. *Crop Science* **48**: S-69–S-84.

Gordon SP, Contreras-Moreira B, Woods DP, Des Marais DL, Burgess D, Shu S, Stritt C, Roulin A, Schackwitz W, Tyler L *et al.* 2017. Extensive gene content variation in the Brachypodium distachyon pan-genome correlates with population structure. *Nature Communications*. doi: 10.1038/s41467-017-02292-8.

Greiner S, Sobanski J, Bock R. 2015. Why are most organelle genomes transmitted maternally? *Bioessays* **37**: 80–94.

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology* **52**: 696–704.

Hao W. 2010. OrgConv: detection of gene conversion using consensus sequences and its application in plant mitochondrial and chloroplast homologs. *BMC Bioinformatics* **11**: 114.

Hochbach A, Schneider J, Röser M. 2015. A multi-locus analysis of phylogenetic relationships within grass subfamily Pooideae (Poaceae) inferred from sequences of nuclear single copy gene regions compared with plastid DNA. *Molecular Phylogenetics and Evolution* **87**: 14–27.

Jansen RK, Ruhlman TA. 2012. Plastid genomes of seed plants. In: Bock R, Knoop V, eds. *Genomics of chloroplast and mitochondria. Advances in photosynthesis and respiration (including bioenergy and related processes), vol.* 35. Dordrecht, the Netherlands: Springer, 103–126.

Jenkins DG, Carey M, Czerniewska J, Fletcher J, Hether T, Jones A, Knight S, Knox J, Long T, Mannino M *et al.* 2010. A meta-analysis of isolation by distance: relic or reference standard for landscape genetics? *Ecography* **33**: 315–320.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**: 772–780.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C *et al.* 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**: 1647–1649.

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. *Genome Research* **19**: 1639–1645.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**: 1870–1874.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Liu Y, Schröder J, Schmidt B. 2013. Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. *Bioinformatics* **29**: 308–315.

Lohse M, Drechsel O, Kahlau S, Bock R. 2013. OrganellarGenomeDRAW – a suite of tools for generating physical maps of plastid and mitochondrial genomes and visualizing expression data sets. *Nucleic Acids Research* **41**: 575–581.

López-Alvarez D, López-Herranz ML, Betekhtin A, Catalán P. 2012. A DNA barcoding method to discriminate between the model plant *Brachypodium distachyon* and its close relatives *B. stacei* and *B. hybridum* (Poaceae). *PLoS ONE* **7**: e51058.

López-Alvarez D, Manzaneda AJ, Rey PJ, Giraldo P, Benavente E, Allainguillaume J, Mur L, Caicedo AL, Hazen SP, Breiman A *et al.* 2015. Environmental niche variation and evolutionary diversification of the *Brachypodium distachyon* grass complex species in their native circum-Mediterranean range. *American Journal of Botany* **102**: 1073–1088.

Lowry DB, Modliszewski JL, Wright KM, Wu CA, Willis JH. 2008. The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **363**: 3009–3021.

Ma PF, Zhang YX, Zeng CX, Guo ZH, Li DZ. 2014. Chloroplast phylogenomic analyses resolve deep-level relationships of an intractable bamboo tribe Arundinarieae (Poaceae). *Systematic Biology* **63**: 933–950.

Mairal M, Pokorny L, Aldasoro JJ, Alarcón ML, Sanmartín I. 2015. Ancient vicariance and climate-driven extinction explain continental-wide disjunctions in Africa: the case of the Rand Flora genus *Canarina* (Campanulaceae). *Molecular Ecology* **24**: 1335–1354.

Manzaneda AJ, Rey PJ, Anderson JT, Raskin E, Weiss-Lehman C, Mitchell-Olds T. 2015. Natural variation, differentiation, and genetic trade-offs of ecophysiological traits in response to water limitation in Brachypodium distachyon and its descendent allotetraploid *B. hybridum* (Poaceae). *Evolution* **69**: 2689–2704.

Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. 2015. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**: 1–5.

Mason-Gamer RJ, Holsinger KE, Jansen RK. 1995. Chloroplast DNA haplotype variation within and among populations of *Coreopsis grandiflora* (Asteraceae). *Molecular Biology and Evolution* **12**: 371–381.

Medgyesy P, Fejes E, Maliga P. 1985. Interspecific chloroplast recombination in a *Nicotiana* somatic hybrid. *Proceedings of the National Academy of Sciences, USA* **82**: 6960–6964.

Middleton CP, Senerchia N, Stein N, Akhunov ED, Keller B, Wicker T, Kilian B. 2014. Sequencing of chloroplast genomes from wheat, barley, rye and their relatives provides a detailed insight into the evolution of the triticeae tribe. *PLoS ONE* **9**: e85761.

Minaya M, Díaz-Pérez A, Mason-Gamer R, Pimentel M, Catalán P. 2015. Evolution of the beta-amylase gene in the temperate grasses: non-purifying selection, recombination, semiparalogy, homeology and phylogenetic signal. *Molecular Phylogenetics and Evolution* **91**: 68–85.

Mogensen H. 1996. The hows and whys of cytoplasmic inheritance in seed plants. *American Journal of Botany* **83**: 383–404.

Morris LM, Duvall MR. 2010. The chloroplast genome of *Anomochloa marantoidea* (Anomochlooideae; Poaceae) comprises a mixture of grass-like and unique features. *American Journal of Botany* **97**: 620–627.

Mur LAJ, Allainguillaume J, Catalán P, Hasterok R, Jenkins G, Lesniewska K, Thomas I, Vogel J. 2011. Exploiting the *Brachypodium* tool box in cereal and grass research. *New Phytologist* **191**: 334–347.

Nadalin F, Vezzi F, Policriti A. 2012. GapFiller: a *de novo* assembly approach to fill the gap within paired reads. *BMC Bioinformatics* **13**: S8.

Nock CJ, Waters DLE, Edwards MA, Bowen SG, Rice N, Cordeiro GM, Henry RJ. 2011. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnology Journal* **9**: 328–333.

Noirot M, Charrier A, Stoffelen P, Anthony F. 2016. Reproductive isolation, gene flow and speciation in the former *Coffea* subgenus: a review. *Trees* **30**: 597–608.

Peterson DG, Tomkins JP, Frisch DA, Wing RA, Paterson AH. 2000. Construction of plant bacterial artificial chromosome (BAC) libraries: an illustrated guide. *Journal of Agricultural Genomics* **5**: 1–100.

Pokorny L, Oliván G, Shaw AJ. 2011. Phylogeographic patterns in two Southern Hemisphere species of *Calyptrochaeta* (Daltoniaceae, Bryophyta). *Systematic Botany* **36**: 542–553.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Ream TS, Woods DP, Schwartz CJ, Sanabria CP, Mahoy JA, Walters EM, Kaeppler HF, Amasino RM. 2014. Interaction of photoperiod and vernalization determines flowering time of *Brachypodium distachyon. Plant Physiology* **164**: 694–709.

Ronen R, Boucher C, Chitsaz H, Pevzner P. 2012. SEQuel: improving the accuracy of genome assemblies. *Bioinformatics* **28**: 188–196.

Ronquist F, Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**: 1572–1574.

Ronquist F, Huelsenbeck J, Teslenko M. 2011. *MrBayes Version 3.2 manual: tutorials and model summaries.* [WWW document] URL http://mrbayes.sourceforge.net/mb3.2_manual.pdf [accessed 1 March 2015].

Saarela JM, Wysocki WP, Barrett CF, Soreng RJ, Davis JI, Clark LG, Kelchner SA, Pires JC, Edger PP, Mayfield DR *et al.* 2015. Plastid phylogenomics of the cool-season grass subfamily: clarification of relationships among early-diverging tribes. *AoB PLANTS* **7**: 1–27.

Saski C, Siri SL, Chittibabu F, Jansen RK, Luo H, Je V, Odd T, Rognli A, Daniell H, Liu J. 2007. Complete chloroplast genome sequences of *Hordeum vulgare, Sorghum bicolor* and *Agrostis stolonifera*, and comparative analyses with other grass genomes. *Theoretical and Applied Genetics* **115**: 571–590.

Schippmann U. 1991. Revision der europäischen Arten der Gattung *Brachypodium* Palisot de Beauvois (Poaceae). *Boissiera* **45**: 1–250.

Schneider J, Winterfeld G, Hoffmann MH, Röser M. 2011. Duthieeae, a new tribe of grasses (Poaceae) identified among the early diverging lineages of subfamily Pooideae: molecular phylogenetics, morphological delineation, cytogenetics, and biogeography. *Systematics and Biodiversity* **9**: 27–44.

Schwartz CJ, Doyle MR, Manzaneda AJ, Rey PJ, Mitchell-Olds T, Amasino RM. 2010. Natural variation of flowering time and vernalization responsiveness in *Brachypodium distachyon. Bioenergy Research* **3**: 38–46.

Silvertown J, Servaes C, Biss P, Macleod D. 2005. Reinforcement of reproductive isolation between adjacent populations in the Park Grass Experiment. *Heredity* **95**: 198–205.

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust

Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, Filgueiras TS, Davis JI, Morrone O. 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae); a worldwide phylogenetic classification of the Poaceae (Gramineae). *Journal of Systematics and Evolution* **53**: 117–137.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Strömberg CAE. 2011. Evolution of grasses and grassland ecosystems. *Annual Review of Earth and Planetary Sciences* **39**: 517–544.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**: 178–192.

Tyler AL, Lee SJ, Young ND, Deiulio GA, Benavente E, Reagon M, Sysopha J, Baldini RM, Troìa A, Hazen SP *et al.* 2016. Population structure in the model grass *Brachypodium distachyon* is highly correlated with flowering differences across broad geographic areas. *Plant Genome* **9**: 1–55.

Vogel J, Bragg J. 2009. *Brachypodium distachyon*, a new model for the Triticeae. In: Muehlbauer GJ, Feuillet C, eds. *Genetics and genomics of the Triticeae*. New York, NY, USA: Springer, 427–449.

Vogel J, Hill T. 2008. High-efficiency *Agrobacterium*-mediated transformation of *Brachypodium distachyon* inbred line Bd21-3. *Plant Cell Reports* **27**: 471–478.

Vogel JP. 2016. The rise of *Brachypodium* as a model system. In: Vogel JP, ed. *Genetics and genomics of* Brachypodium. Cham, Switzerland: Springer, 1–8.

Vogel JP, Garvin DF, Leong OM, Hayden DM. 2006. *Agrobacterium*-mediated transformation and inbred line development in the model grass *Brachypodium distachyon*. *Plant Cell, Tissue and Organ Culture* **84**: 199–211.

Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon-Smith M, Lail K *et al.* 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**: 763–768.

Vogel JP, Tuna M, Budak H, Huo N, Gu YQ, Steinwand MA. 2009. Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon*. *BMC Plant Biology* **9**: 88.

Wang R-J, Cheng C-L, Chang C-C, Wu C-L, Su T-M, Chaw S-M. 2008. Dynamics and evolution of the inverted repeat-large single copy junctions in the chloroplast genomes of monocots. *BMC Evolutionary Biology* **8**: 36.

Waters DLE, Nock CJ, Ishikawa R, Rice N, Henry RJ. 2012. Chloroplast genome sequence confirms distinctness of Australian and Asian wild rice. *Ecology and Evolution* **2**: 211–217.

Woods DP, Mckeown MA, Dong Y, Preston JC, Amasino RM. 2016. Evolution of *VRN2/GhD7*-like genes in vernalization-mediated repression of grass. *Plant Physiology* **170**: 1–12.

Woods DP, Ream TS, Amasino RM. 2014. Memory of the vernalized state in plants including the model grass *Brachypodium distachyon*. *Frontiers in Plant Science* **5**: 99.

Wright S. 1943. Isolation by distance. *Genetics* **28**: 114–138.

Wu ZQ, Ge S. 2012. The phylogeny of the BEP clade in grasses revisited: evidence from the whole-genome sequences of chloroplasts. *Molecular Phylogenetics and Evolution* **62**: 573–578.

Wysocki WP, Clark LG, Attigala L, Ruiz-Sanchez E, Duvall MR. 2015. Evolution of the bamboos (Bambusoideae; Poaceae): a full plastome phylogenomic analysis. *BMC Evolutionary Biology* **15**: 50.

Zerbino DR. 2010. *Using the Columbus extension to Velvet*. [WWW document] URL http://gensoft.pasteur.fr/docs/velvet/1.1.02/Columbus_manual.pdf [accessed 1 April 2014].

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Pipeline used for the assembly of the *Brachypodium* plastomes.

Fig. S2 Evidence of major indels found among the *Brachypodium distachyon*, *B. stacei* and *B. hybridum* plastomes.

Fig. S3 Phylogenomic analysis of *Brachypodium distachyon* plastomes.

Fig. S4 Potential recombination events detected in the plastomes of the introgressed *Brachypodium distachyon* Arn1 and Mon3 ecotypes.

Fig. S5 Plastome phylogenomic analysis of Poaceae.

Fig. S6 BEAST nested dating analysis of Poaceae (above-species) and *Brachypodium distachyon* (below-species) plastome sequences.

Table S1 List of *Brachypodium distachyon*, *B. stacei* and *B. hybridum* accessions studied

Table S2 Grass plastomes employed in evolutionary and genomic analyses

Table S3 Flowering time classes classified according to Ream *et al.* (2014)

Table S4 Bioinformatic tools used in the assembly and annotation of *Brachypodium* plastomes and in their evolutionary and genomic analyses

Table S5 Comparative ptDNA data of the assembled *Brachypodium distachyon*, *B. hybridum* and *B. stacei* plastomes and EMBL/ENA accession numbers

Table S6 Primers used for amplification and sequencing of IRa and IRb junction regions and of the IR *rps*19 copy

Table S7 Polymorphisms found in inter- and intraspecific comparisons of the *Brachypodium distachyon*, *B. stacei* and *B. hybridum* plastomes

Table S8 List of *Brachypodium distachyon* ptDNA haplotypes found across the 53 analyzed ecotypes' plastomes

Table S9 Percentages of membership of 53 *B. distachyon* ecotypes' plastome profiles to optimal $K=2$ and $K=4$ Bayesian genomic groups

Table S10 Pairwise Tamura–Nei raw and phylogenetically based patristic genetic distances between three *Brachypodium* and 91 grass plastomes

Methods S1 Detailed description of the plastome automated assembly pipeline, annotation and validation processes.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

© 2017 The Authors
New Phytologist © 2017 New Phytologist Trust