The LINNEAN SOCIETY of London BOTANICAL Journal of the Linnean Society

Botanical Journal of the Linnean Society, 2015, **179**, 421–435. With figures

Molecular phylogenetics of *Kosteletzkya* (Malvaceae, Hibisceae) reveals multiple independent and successive polyploid speciation events

KURT M. NEUBIG[1,2]*, ORLAND J. BLANCHARD Jr[2], W. MARK WHITTEN[2] and STUART F. MCDANIEL[2,3]

[1]*Department of Plant Biology, Southern Illinois University Carbondale, Carbondale, IL 62901-6509, USA*
[2]*Florida Museum of Natural History, University of Florida, Gainesville, FL 32611-7800, USA*
[3]*Department of Biology, University of Florida, Gainesville, FL 32611-8526, USA*

Received 27 January 2015; revised 5 June 2015; accepted for publication 27 July 2015

Kosteletzkya s.s. is a genus of 17 species (excluding the endemic species of Madagascar), found in the New World, continental Africa, Madagascar, and Southeast Asia. Recent chromosome counts revealed diploid, tetraploid, and hexaploid species. To estimate the history of the genus, we sequenced nuclear and plastid loci for nearly all *Kosteletzkya* spp., in the majority of cases, with multiple accessions per species. The African species form a paraphyletic grade relative to a New World clade. Polyploidy has occurred only in some African species, resulting in the relatively ancient formation of one putative autotetraploid species (*K. semota*), one recent allotetraploid species (*K. borkouana*), two relatively ancient allotetraploid species (*K. begoniifolia* and *K. rotundalata*) and one recent allohexaploid species (*K. racemosa*). Our inferences regarding the hypothesized parentage of the polyploids mostly corroborate previous work based on chromosome-pairing patterns in artificial hybrids, highlighting the utility of these complementary data sources. © 2015 The Linnean Society of London, *Botanical Journal of the Linnean Society*, 2015, **179**, 421–435.

ADDITIONAL KEYWORDS: allopolyploidy – autopolyploidy – phylogeny.

INTRODUCTION

Polyploidy is widely believed to play an important role in the generation of plant diversity because it allows for the rapid evolution of reproductive isolation without the need for allopatry (Grant, 1981; Coyne & Orr, 2004). The frequency of genome duplication events in the long history of angiosperms attests to the important role that polyploidy plays in the evolution of plants (Soltis, 2005; Jiao *et al*., 2011). This frequency is reflected at deep and shallow levels of phylogenetic breadth. Detailed studies of recent allopolyploid events have shown that they can be dynamic and repeated (Soltis & Soltis, 1993; Leitch & Bennett, 1997; Soltis *et al*., 2004; Tate *et al*., 2006), suggestive of a polyphyletic origin of at least some polyploids. Hybridization, a frequent initiator of polyploidization, can create entities with a greater ability to invade new habitats and spread geographically (Stebbins, 1985).

One such example of well documented polyploids is the small genus *Kosteletzkya* C. Presl (Malvaceae, Hibisceae; Table 1; Fig. 1). *Kosteletzkya* is a familiar genus among North American botanists because *K. pentacarpos* (L.) Ledeb., until recently known as *K. virginica* (L.) C. Presl ex A. Gray, is a common species in coastal communities that has attracted recent interest as a potential biofuel (Ruan, Xing & Teixeira da Silva, 2012). DNA data have confirmed the placement of *Kosteletzkya* in the complex malvoid tribe Hibisceae, among 25–30 other genera including the species-rich and phylogenetically problematic genera

*Corresponding author. E-mail: kneubig@siu.edu

Table 1. Summary of chromosome numbers (Blanchard, 1974, 2012), genome designations (Blanchard, 2013a),, and general geographical distribution of *Kosteletzkya* spp**.**

Species	Chromosome number (number of collections counted)	Genome	Geographical distribution
K. blanchardii	*n* = 19 (1)	BB	Mexico
K. depressa	*n* = 19 (12)	BB	northern Neotropics
K. hispidula	*n* = 19 (3)	BB	Mexico
K. pentacarpos	*n* = 19 (9)	BB	North America and western Eurasia
K. ramosa	*n* = 19 (1)	BB	Mexico
K. reclinata	*n* = 19 (1)	BB	Mexico
K. tubiflora	*n* = 19 (2)	BB	Mexico
K. adoensis	*n* = 19 (7)	AA	Africa
K. buettneri	*n* = 19 (7)	BB	Africa
K. grantii	*n* = 19 (3)	GG	Africa
K. borkouana	*n* = 38 (4)	AABB	Africa
K. begoniifolia	*n* = 38 (4)	AAGG	Africa
K. rotundalata	*n* = 38 (1)	AAGG	Africa
K. semota	*n* = 37–38 (1)	XXYY	Africa
K. racemosa	*n* = 57 (1)	AABBGG	Africa

Hibiscus L. and *Pavonia* Cav. (La Duke & Doebley, 1995; Pfeil & Crisp, 2005). Other DNA results (Koopman & Baum, 2008; Neubig & Blanchard, unpublished data) and differences in fruit dehiscence (Blanchard, 2012, 2013a) and other morphological characters have prompted us to exclude as unrelated the eight Malagasy endemics that have been described in *Kosteletzkya*. The remaining 17 species {which include the type of the genus, *K. sagittata* C. Presl = [*K. depressa* (L.) O.J. Blanch., Fryxell & D.M. Bates]} form a morphologically well defined clade. Except for one annual [*K. batacensis* (Blanco) Fern.-Vill.], all species of *Kosteletzkya s.s.* are herbaceous perennials. Most have *Hibiscus*-like flowers with petals ranging in colour from white to pink, sometimes with a darker base (Fig. 1A–R, T). The flowers of most species are rotate or broadly campanulate, last for a single day and are presumably bee-pollinated, but two [*K. thurberi* A. Gray and *K. tubiflora* (Moc. & Sessé ex DC.) O.J. Blanch. & McVaugh] have yellow or deep pink flowers with convoluted (tubular) corollas (Fig. 1S) that persist for 2 or 3 days and exhibit protogyny (Blanchard, 2013a, unpublished data). These two species are putatively bird-pollinated (van Devender *et al*., 2004). The capsular fruits are depressed and pentagonal and they disintegrate at maturity to release the five seeds within.

Kosteletzkya is a cytologically well defined genus ($x = n = 19$; Blanchard, 2012) that is distributed primarily in the New World, continental Africa, and Southeast Asia. Polyploidy in the genus is confined to Africa, where, in addition to three known diploid species, there are four tetraploids [*K. begoniifolia* (Ulbr.) Ulbr., *K. borkouana* Quézel, *K. rotundalata* O.J. Blanch., and *K. semota* O.J. Blanch.] and a single hexaploid (*K. racemosa* Hauman) (Blanchard, 1974, 2012, 2013b). In two of the tetraploids and in all three African diploids, chromosome counts have been documented in multiple accessions, indicating that chromosome numbers are constant in *Kosteletzkya* spp. (Table 1). Each of the African diploid species is widely distributed and broadly sympatric with the other two, though they may sometimes be ecologically separated. In contrast, the New World (diploid) species are largely allopatric (Blanchard, 2013a). The African polyploids are allopatric with respect to one another, but they usually overlap with one or both of their putative parents.

An extensive programme of experimental hybridization among 15 species of *Kosteletzkya* spp. and the study of chromosome pairing in hybrids has identified three distinct 'genome types' in the genus and the potential existence of two others (Blanchard, 2013a). The pairing evidence also suggests reticulate evolution built on the African diploids and a colonization of the New World by a carrier of one of the African genomes (Blanchard, 2013a). However, definitive molecular phylogenetic data to support relationships among parental haplotypes have been lacking until now.

The purpose of this study is to use plastid and nuclear DNA sequences to ascertain the phylogenetic

Figure 1. Floral diversity of *Kosteletzkya* spp. African species: (A) *K. adoensis*, (B) *K. adoensis*, (C) *K. adoensis*, *3420*, (D) *K. grantii*, *3427*, (E) *K. begoniifolia*, *3294*, (F) *K. begoniifolia*, *3390*, (G) *K. borkouana*, *3392*, (H) *K. borkouana*, *3460*, (I) *K. rotundalata*, *3417*, (J) *K. semota*, *3437*, (K) *K. racemosa*, *3391*, (L) *K. buettneri*, *3384*; and New World species: (M) *K. pentacarpos*, *3334*, (N) *K. depressa*, *3350*, (O) *K. depressa*, *3161*, (P) *K. hispidula*, *3365*, (Q) *K. reclinata*, *3257*, (R) *K. ramosa*, *3354*, (S) *K. tubiflora*, *3385*, and (T) *K. blanchardii*, *3387*. Scale bars: 1 cm. Numbers indicate voucher information, where available (see Table 2).

relationships in *Kosteletzkya s.s.* and to understand the origins of the polyploid taxa. Secondarily, we aim to relate these findings to the hybridization studies of Blanchard (2013a).

MATERIALS AND METHODS

Taxon sampling

Seeds were obtained from wild-collected plants or greenhouse-grown descendants of these plants and then grown for tissue used for DNA extraction and vouchers (Table 2). Sampling of *Kosteletzkya* included 15 of the 17 species (*K. thurberi* and *K. batacensis* were not available) and a spontaneously tetraploidized hybrid between *K. grantii* (Mast.) Garcke and *K. adoensis* (Hochst. ex A. Rich.) Mast. (Blanchard, 2013a). We excluded the Malagasy endemic species from this analysis because molecular data place them in another clade of Hibisceae; the remaining species form a clade that includes taxa from continental Africa and the New World (Neubig & Blanchard, unpublished data). The outgroup (*Hibiscus vitifolius* L.) was chosen because of its close relationship to *Kosteletzkya* in tribe Hibisceae based on previous phylogenetic work (Pfeil *et al*., 2002; Koopman & Baum, 2008). Much of the material used in this phylogenetic study is from the same collections as were used in the chromosome-pairing study of Blanchard (2013a).

Extractions, amplification, and sequencing

All freshly collected materials were preserved in silica gel (Chase & Hills, 1991). Genomic DNA was extracted using a modified cetyl trimethylammonium bromide (CTAB) technique (Doyle & Doyle, 1987), scaled to a 1.5 mL volume reaction. Approximately 10 mg dried tissue was ground in 1.5 mL CTAB 2× buffer and 20 μg proteinase K. Amplifications were performed using a Biometra T gradient or an Eppendorf Mastercycler EP Gradient S thermocycler and reagents in ~25 μL volumes. Loci were chosen based on previous evidence demonstrating relatively high nucleotide variation (Taberlet *et al*., 1991; Cronn *et al*., 2002b; Small, 2004; Shaw *et al*., 2007; Neubig *et al*., 2009; Fazekas *et al*., 2012).

Amplification and sequencing primers are listed in Table 3. The plastid regions *matK*, *trnL–F*, *rpl32–trnL*, and *trnQ–rps16* and the nuclear regions *A1341* [see (Cronn *et al*., 2002b) for details on this locus] and *GBSSI* were amplified using Jumpstart (Sigma) reagents: 0.5–1.0 μL template DNA (~10–100 ng), 17.5 μL water, 2.5 μL 10× buffer, 2.0–2.5 μL $MgCl_2$ (25 mM), 0.5 μL of 10 μM dNTPs, 0.5 μL of each 10 μM primers, and 0.5 units *Taq*. For the plastid region *ycf1* (hypothetical plastid reading frame 1) and the nuclear locus *CesA1* (cellulose synthase), high-fidelity Phusion reagents were used: 0.5–1.0 μL template DNA (~10–100 ng), 15.5 μL water, 5 μL 5× buffer, 1 μL $MgCl_2$ (25 mM), 0.5 μL 10 μM dNTPs, 0.5 μL of each 10 μM primers and 0.5 units polymerase.

A *c.* 800-bp portion of *matK* was amplified with the parameters 94 °C, 3 min; 33 × (94 °C, 30 s; 55 °C, 30 s; 72 °C, 2 min); 72 °C, 3 min. The *rpl32–trnL* intergenic spacer was amplified with the parameters 94 °C, 3 min; 38 × (94 °C, 30 s; 52 °C, 30 s; 72 °C, 2 min); 72 °C, 3 min. The *trnL–trnF* region (including the *trnL* intron and the *trnL–trnF* intergenic spacer) was amplified with the same parameters as *matK*. The *trnQ-rps16* intergenic spacer was amplified with the parameters 94 °C, 3 min; 30 × (94 °C, 45 s; 58 °C, 45 s; 72 °C, 1 min); 72 °C, 3 min. The nuclear regions *A1341* and *GBSSI* were amplified with the parameters 94 °C, 3 min; 8 × (94 °C, 30 s; 60 °C, 1 min, reducing to 1 °C per cycle; 72 °C, 2 min 30 s); 30 × (94 °C, 30 s; 50 °C, 45 s; 72 °C, 2 min 30 s); 72 °C, 3 min. For *ycf1*, we sequenced a *c.* 2700-base pair (bp) portion from the 3′ end of the open reading frame (ORF), amplifying with primers 2860F and 5700R and sequencing with these and 3670F, 4400F and 4565R. For *CesA1*, we sequenced an c. 950-bp portion of the gene, including exons and introns using primers CelAF and CelAR (Cronn *et al*., 2002b). Both loci were amplified with the parameters 98 °C, 2 min; 33 × (98 °C, 10 s; 55 °C, 15 s; 72 °C, 1 min, 45 s); 72 °C, 3 min.

Polymerase chain reaction (PCR) products of nuclear loci in polyploid taxa were cloned using a TopoTA Top10 cloning kit (Invitrogen) according to manufacturer's protocols. Ligation reactions were made by combining 0.2 μL of 10× ligation buffer, 0.4 μL vector, 0.8 μL water, 0.2 μL ligase, and 0.4 μL PCR product, then incubating for 30 min at room temperature. Tubes of competent cells were divided into ¼ reactions after thawing on ice, with 1 μL of the ligation reaction, then incubating on ice for 30 min. The competent cell/ligation mix was then heat-shocked at 42 °C for 30 sec, then placed on ice for 2 min. Transformed cells were then incubated at 37 °C with 150 μL SOC medium for 1 h. The solution was then spread on plates of Lysogeny Broth (LB) media [stock mix contains 5 g tryptone, 2.5 g yeast extract, 2.5 g NaCl, 7.5 g agar, and 500 mL of H_2O], then autoclaved; 1 mL of ampicillin (stock: 0.5 g in 5 mL water) was then added, then plated; 40 μL X-Gal (stock: 0.15 g in 5 mL dimethyl formamide) was then spread onto the surface of individual plates. Transformed *Escherichia coli* cells were streaked on fresh plates, incorporation of the target amplicon was checked by PCR following the above protocols. Multiple clones for each locus in each polyploid individual were amplified and sequenced. It was not necessary to clone any of the PCRs of diploid taxa, except for two accessions: *K. depressa* (*Blanchard*

Table 2. Species names, voucher information (all specimens deposited at FLAS), and GenBank accessions for material used in this study [Species, voucher (additional voucher specimens); country of origin; *A1341*; *CesA1*; *GBSSI*; *matK*; *rpL32–trnL*; *trnL–trnF*; *trnQ–rps16*; *ycf1*]

Hibiscus vitifolius L., *O.J. Blanchard, Jr. 3352 (3331)*; Australia; NA; KM463316; KM463369; KM463417; KM463458; KM463499; KM463538; KM463578. *Kosteletzkya adoensis* (Hochst. ex A. Rich.) Mast., *O.J. Blanchard, Jr. 3405*; Angola; KM463255; KM463297; KM463347; KM463401; KM463442; KM463483; KM463522; KM463563. *K. adoensis*, *O.J. Blanchard, Jr. 3402*; Malawi; KM463256; KM463298; KM463348; KM463402; KM463443; KM463484; KM463523; KM463564. *K. adoensis*, *O.J. Blanchard, Jr. 3410*; Tanzania; KM463257; KM463299; KM463349; KM463403; KM463444; KM463485; KM463524; KM463565. *K. adoensis*, *O.J. Blanchard, Jr. 3401*; Sierra Leone; KM463261; KM463303; NA; KM463407; KM463448; KM463489; KM463528; NA. *K. adoensis*, *O.J. Blanchard, Jr. 3404*; Ethiopia; KM463262; NA; KM463354; KM463408; KM463449; KM463490; KM463529; KM463569. *K. adoensis*, *O.J. Blanchard, Jr. 3420 (3403)*; Congo-Kinshasa; KM463274; KM463322; KM463375; KM463420; KM463461; KM463502; KM463541; KM463581. *K. adoensis*, *O.J. Blanchard, Jr. 3413 (3396)*; Kenya; KM463286; KM463336; KM463390; KM463432; KM463473; KM463512; KM463553; KM463593. *K. begoniifolia* Ulbr., *O.J. Blanchard, Jr. 3389*; Kenya; KM463258; KM463300; KM463350, KM463351; KM463404; KM463445; KM463486; KM463525; KM463566. *K. begoniifolia*, *O.J. Blanchard, Jr. 3399*; Ethiopia; KM463263; KM463304; KM463355, KM463356; KM463409; KM463450; KM463491; KM463530; KM463570. *K. begoniifolia*, *O.J. Blanchard, Jr. 3388*; Ethiopia; KM463266; KM463308; KM463360, KM463361; KM463412; KM463453; KM463494; KM463533; KM463573. *K. begoniifolia*, *O.J. Blanchard, Jr. 3294*; Kenya; KM463275; KM463323, KM463324; KM463376, KM463377; KM463421; KM463462; KM463503; KM463542; KM463582. *K. begoniifolia*, *O.J. Blanchard, Jr. 3390 (3411)*; Tanzania; NA; KM463330; KM463383, KM463384; KM463427; KM463468; KM463507; KM463548; KM463588. *K. blanchardii* Fryxell, *O.J. Blanchard, Jr. 3387 (3378)*; Mexico; KM463265; KM463307; KM463359; KM463411; KM463452; KM463493; KM463532; KM463572. *K. borkouana* Quézel, *O.J. Blanchard, Jr. 3460*; Uganda; KM463264; KM463305, KM463306; KM463357, KM463358; KM463410; KM463451; KM463492; KM463531; KM463571. *K. borkouana*, *O.J. Blanchard, Jr. 3361*; Chad; KM463268; KM463310, KM463311; KM463363, KM463364; KM463414; KM463455; KM463496; KM463535; KM463575. *K. borkouana*, *O.J. Blanchard, Jr. 3376*; Congo-Kinshasa; KM463269; KM463312, KM463313; KM463365, KM463366; KM463415; KM463456; KM463497; KM463536; KM463576. *K. borkouana*, *O.J. Blanchard, Jr. 3377*; Congo-Kinshasa; KM463270; KM463314, KM463315; KM463367, KM463368; KM463416; KM463457; KM463498; KM463537; KM463577. *K. borkouana*, *O.J. Blanchard, Jr. 3392*; Tanzania; KM463271; KM463317, KM463318; KM463370, KM463371; KM463418; KM463459; KM463500; KM463539; KM463579. *K. buettneri* Gürke, *O.J. Blanchard, Jr. 3379 (3337)*; Malawi; KM463267; KM463309; KM463362; KM463413; KM463454; KM463495; KM463534; KM463574. *K. buettneri*, *O.J. Blanchard, Jr. 3384 (3339)*; Tanzania; KM463279; KM463329; NA; KM463425; KM463466; KM463505; KM463546; KM463586. *K. buettneri*, *O.J. Blanchard, Jr. 3426*; Guinea; KM463280; NA; KM463382; KM463426; KM463467; KM463506; KM463547; KM463587. *K. buettneri*, *O.J. Blanchard, Jr. 3383 (3355)*; Central African Republic; KM463282; KM463332; KM463386; KM463429; KM463470; KM463509; KM463550; KM463590. *K. buettneri*, *O.J. Blanchard, Jr. 3380 (3363)*; Congo-Kinshasa; KM463283; KM463333; KM463387; KM463430; KM463471; KM463510; KM463551; KM463591. *K. buettneri*, *O.J. Blanchard, Jr. 3357*; Congo-Kinshasa; KM463288; KM463338; KM463392; KM463434; KM463475; KM463514; KM463555; KM463595. *K. buettneri*, *O.J. Blanchard, Jr. 3381*; Zambia; KM463289; KM463339; KM463393; KM463435; KM463476; KM463515; KM463556; KM463596. *K. depressa* (L.) O.J. Blanch., Fryxell & D.M. Bates, *O.J. Blanchard, Jr. 3438*; Peru; KM463260; KM463302; KM463353; KM463406; KM463447; KM463488; KM463527; KM463568. *K. grantii* (Mast.) Garcke, *O.J. Blanchard, Jr. 3424*; Congo-Kinshasa; KM463277; KM463327; KM463380; KM463423; KM463464; NA; KM463544; KM463584. *K. grantii*, *O.J. Blanchard, Jr. 3422*; Nigeria; KM463278; KM463328; KM463381; KM463424; KM463465; NA; KM463545; KM463585. *K. grantii*, *O.J. Blanchard, Jr. 3423*; Congo-Kinshasa; KM463281; KM463331; KM463385; KM463428; KM463469; KM463508; KM463549; KM463589. *K. grantii*, *O.J. Blanchard, Jr. 3425*; Kenya; KM463293; KM463343; KM463397; KM463438; KM463479; KM463518; KM463559; KM463599. *K. grantii*, *O.J. Blanchard, Jr. 3427*; Nigeria; KM463294; KM463344; KM463398; KM463439; KM463480; KM463519; KM463560; KM463600. *K. hispidula* (Spreng.) Garcke, *O.J. Blanchard, Jr. 3429 (3364, 3416)*; Mexico; KM463259; KM463301; KM463352; KM463405; KM463446; KM463487; KM463526; KM463567. *K. pentacarpos* (L.) Ledeb., *O.J. Blanchard, Jr. 3336 (3335)*; USA; KM463287; KM463337; KM463391; KM463433; KM463474; KM463513; KM463554; KM463594. *K. racemosa* Hauman, *O.J. Blanchard, Jr. 3391*; Congo-Kinshasa; KM463272, KM463273; KM463319, KM463320, KM463321; KM463372, KM463373, KM463374; KM463419; KM463460; KM463501; KM463540; KM463580. *K. ramosa* Fryxell, *O.J. Blanchard, Jr. 3354*; Mexico; KM463253; KM463295; KM463345; KM463399; KM463440; KM463481; KM463520; KM463561. *K. reclinata* Fryxell, *O.J. Blanchard, Jr. 3257 (3351)*; Mexico; KM463254; KM463296; KM463346; KM463400; KM463441; KM463482; KM463521; KM463562. *K. rotundalata* O.J. Blanch., *O.J. Blanchard, Jr. 3417 (3398)*; Congo-Kinshasa; KM463276; KM463325, KM463326; KM463378, KM463379; KM463422; KM463463; KM463504; KM463543; KM463583. *K. semota* O.J. Blanch., *O.J. Blanchard, Jr. 3437*; Nigeria; KM463290, KM463291; KM463340, KM463341; KM463394, KM463395; KM463436; KM463477; KM463516; KM463557; KM463597. *K. adoensis* × *K. grantii*, *spontaneously tetraploidized*, *O.J. Blanchard, Jr. 3400 (3418)*; KM463284, KM463285; KM463334, KM463335; KM463388, KM463389; KM463431; KM463472; KM463511; KM463552; KM463592. *K. tubiflora* (DC.) O.J. Blanch. & McVaugh, *O.J. Blanchard, Jr. 3386 (3360)*; Mexico; KM463292; KM463342; KM463396; KM463437; KM463478; KM463517; KM463558; KM463598.

Table 3. Primers used in this study and the associated papers from which they were obtained

Locus/primer	Primer sequence	Reference
Nuclear primers:		
A1341		
A1341F	GCATGCTGAATTGACAGAACCAGCY	Cronn *et al*. (2002b)
A1341R	CACTCACAAAGTTATGCCGGATGY	Cronn *et al*. (2002b)
CesA1		
CelAF	GATGGAATCTGGGGTTCCTGTTTGC	Cronn *et al*. (2002b)
CelAR	GGGAACTGATCCAACACCCAGGA	Cronn *et al*. (2002b)
GBSSI		
GBSSI1F	CTGGTGGACTCGGTGATGTTCTTG	Evans *et al*. (2000)
GBSSI9R	CTCTTCTAGCCTGCCAATGAACC	Evans *et al*. (2000)
set1GBSSI1560R	CTAGTAACAAAAATAACACCAGC	This paper
set1GBSSI1625R	GCTTTTAATTCCTCRACTGAAG	This paper
set1GBSSI10Fadoe	GGCTTGATTTTGGTTAATTTATCATT	This paper
set1GBSSI10Fbuett	GGCTTGATTTTGGTTAGTTATCG	This paper
set1GBSSI10Fgrant	GGCTTGATTTTGGTTAATTTATCATG	This paper
set1GBSSI1490R	AACAAAAATAACACCAGCGGCA	This paper
set1GBSSI1490Radoe	AACAAAAATAACACCAGCA	This paper
set1GBSSI1490Rbuett	AACAAAAATAACACCAGCGG	This paper
set1GBSSI1490Rgrant	AACAAAAATAACACCAGCGGT	This paper
Plastid primers:		
matK		
3F	CGTACAGTACTTTTGTGTTTACGAG	Fazekas *et al*. (2012)
1R	ACCCAGTCCATCTGGAAATCTTGGTTC	Fazekas *et al*. (2012)
trnL–F		
c	CGAAATCGGTAGACGCTACG	Taberlet *et al*. (1991)
f	ATTTGAACTGGTGACACGAG	Taberlet *et al*. (1991)
rpl32–trnL		
trnL(UAG)	CTGCTTCCTAAGAGCAGCGT	Shaw *et al*. (2007)
rpL32–F	CAGTTCCAAAAAAACGTACTTC	Shaw *et al*. (2007)
trnQ–rps16		
trnQ UUG	GCGTGGCCAAGYGGTAAGGC	Shaw *et al*. (2007)
rps16x1	GTTGCTTTYTACCACATCGTTT	Shaw *et al*. (2007)
ycf1		
ycf1Malv2860F	TTCGTTTTTTGAGCCTATTTTTAAAGAAC	This paper
ycf1Malv2970F	TCAAAAGAAACAAAAAAATGGATC	This paper
ycf1Malv3670F	TCCTCCCTCTCACAAGCATATG	This paper
ycf1Malv4400F	GTCGATATTGAGTCCTGGGTCGATACC	This paper
ycf1Malv4565R	GATTGGATGGGACTGAATGAAGAAA	This paper
ycf1Malv5600R	AAAGTTCTTTCTTTGGCCCAAT	This paper
ycf1Malv5700R	GGTTTAATACTAATAAYGGCAGTCGTT	This paper

3438) and *K. hispidula* (Spreng.) Garcke (*Blanchard 3429*) for *A1341*.

Amplification of *GBSSI* (Evans *et al*., 2000; Small, 2004) yielded three distinct and divergent paralogous copies, but only one copy was used in this study (that copy having distinct homoeologues in polyploids). This copy was amplified using GBSSI1F and GBSSI9R, or alternatively using reverse primers set1GBSSI1560R or set1GBSSI1625R. For a select set of polyploids, GBSSI was cloned and sequenced. After editing those sequences, followed by phylogenetic analysis, the different homoeologues were determined based on phylogenetic positions and primers were designed to differentially amplify those copies. Specifically, primers were designed for homoeologues most closely related to *K. adoensis* (set1GBSSI10-Fadoe and set1GBSSI1490Radoe), *K. grantii* (set1GBSSI10Fgrant and set1GBSSI1490Rgrant) and *K. buettneri* Gürke (set1GBSSI10Fbuett and set1GBSSI1490Rbuett). These primers were then used to amplify all potential homoeologues and to directly sequence them using the same primers

(see Supporting Information Table S1 for further details).

Purified cycle sequencing products were directly sequenced on an ABI 3130 automated sequencer according to the manufacturer's protocols (Applied Biosystems, Foster City, CA, USA). Electropherograms were edited and assembled using Sequencher 4.9 (GeneCodes, Ann Arbor, MI, USA). All sequences were deposited in GenBank (Table 2) and data sets were deposited in Dryad (http://dx.doi.org/10.5061/dryad.2h5j7).

Data analysis

Sequence data were manually aligned using Se-Al v2.0a11 (Rambaut, 1996). No sequence data were excluded from analyses. Indels (insertions/deletions) were not coded as characters. Analyses were performed using PAUP*4.0b10 (Swofford, 1999). Fitch parsimony (unordered characters with equal weights) analyses (Fitch, 1971) used a heuristic search strategy consisted of branch swapping by tree bisection reconnection (TBR), Deltran character optimization, stepwise addition with 1000 random-addition replicates holding five trees at each step and saving multiple trees (MulTrees). Levels of support were assessed using the bootstrap (Felsenstein, 1985). Bootstrap percentages under parsimony were estimated with 1000 bootstrap replicates, using TBR swapping for 50 random-addition replicates per bootstrap replicate. For maximum likelihood (ML), jModelTest2 (Darriba *et al*., 2012) was used to determine the appropriate model for analysis using all combined data under the Akaike information criterion. ML analyses were performed using a HKY + Γ model for the combined nuclear data set and a TVM + I + Γ model for the combined plastid data set. Bootstrap percentages under ML were estimated with 100 bootstrap replicates, using TBR swapping for one random-addition replicate per bootstrap replicate. All analyses were performed for data sets including combined nuclear only, combined plastid only and all individual data sets.

All high-quality sequences from clones for the polyploid species (i.e. putative homoeologues) were initially retained in the analysis. Because the cloning process results in PCR errors and chimaeras (Liesack, Weyland & Stackebrandt, 1991; Cronn *et al*., 2002a), we took a conservative approach to editing homoeologues of polyploids. We first conducted a phylogenetic analysis incorporating all clones in order to place them in the phylogenetic tree. Their placement indicated homoeologue designation, depending on their closest relative in the gene tree. All clones for that homoeologue designation were then combined through a majority rule consensus for the sake of simplicity in presentation and to reduce the phylogenetic noise introduced by sequence error propagated through cloning at the potential cost of omitting allelic variation (McDaniel & Shaw, 2005). To validate this process, we took the GBSSI data (the most variable and complicated data set in terms of homoeologue presence) produced by cloning and directly compared the homoeologues observed to the direct-sequenced homoeologue (amplified by specific primers outlined above) sequences (see Supporting Information Table S1 for further details).

RESULTS

Results of phylogenetic analyses are highly congruent by visual comparisons of phylogenetic structure (Figs 2, 3, S1–S8). All loci, plastid and nuclear, converged upon highly similar topologies as evidenced by a visual inspection, with the obvious caveat that the plastid phylogenetic tree includes only the retained maternally inherited plastid haplotype. Maximum parsimony and ML analyses supported similar basic interspecific relationships. However, because of the nature of biparental inheritance in nuclear markers vs. uniparental inheritance in plastid markers, the plastid signal is, of course, incomplete (Fig. 3).

Our method of consolidating sequences of clones by an iterative process of phylogenetic analysis, then taking a majority rule consensus of those clones that are most closely related for each individual accession, was found to be an effective method for removing PCR-related error in individual sequences. For accessions where there were clone sequence data and direct-sequence data (Supporting Information Table S1), putative spurious nucleotide substitutions specific to single clones were confirmed as spurious. Therefore, we interpret this method as a conservative approach to reduce noise in the sequence data of clones.

Plastid data exclusively represent a maternally inherited genome of a plastid lineage that tells only part of reticulate evolution in *Kosteletzkya*, if such species interactions have occurred. Direct sequencing of nuclear loci revealed that species known to be polyploids produced 'noisy' electropherograms due to simultaneous sequencing of multiple haplotypes (i.e. putative homoeologues). These samples were cloned and sequenced to reveal the parentage of the polyploid species. These three loci (*CesA1*, *A1341*, and *GBSSI*) showed congruent topologies and indicated parental haplotypes that were consistent among them. However, of the three loci, some haplotypes were never recovered in the *A1341* locus and therefore the data are incomplete, either representing true biological absence or artefacts of data acquisition. Phylogenetic relationships among the diploid species did not differ significantly among the different loci.

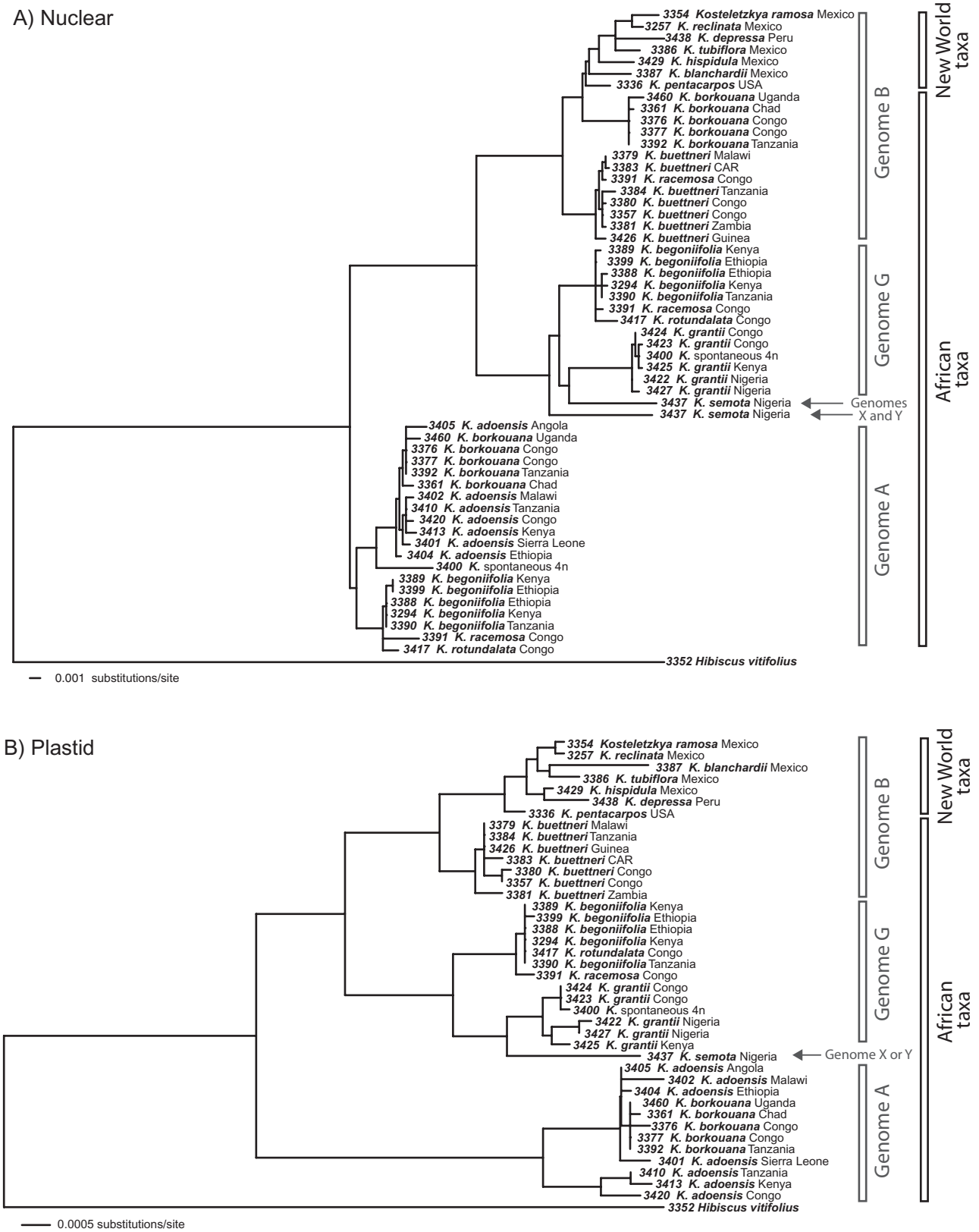

Figure 2. See caption on next page.

Figure 2. Phylogenetic trees from ML analysis of (A) combined nuclear data (*A1341*, *CesA1*, and *GBSSI*) and (B) combined plastid data (*matK*, *trnL–F*, *rpl32–trnL*, *trnQ–rps16*, and *ycf1*), respectively. Major biogeographic groups are delimited in black to the right; genome designations previously described by Blanchard and outlined in Table 1 are delimited in gray to right. Abbreviations: CAR, Central African Republic; Congo, Democratic Republic of the Congo. Numbers indicate voucher information (see Table 2).

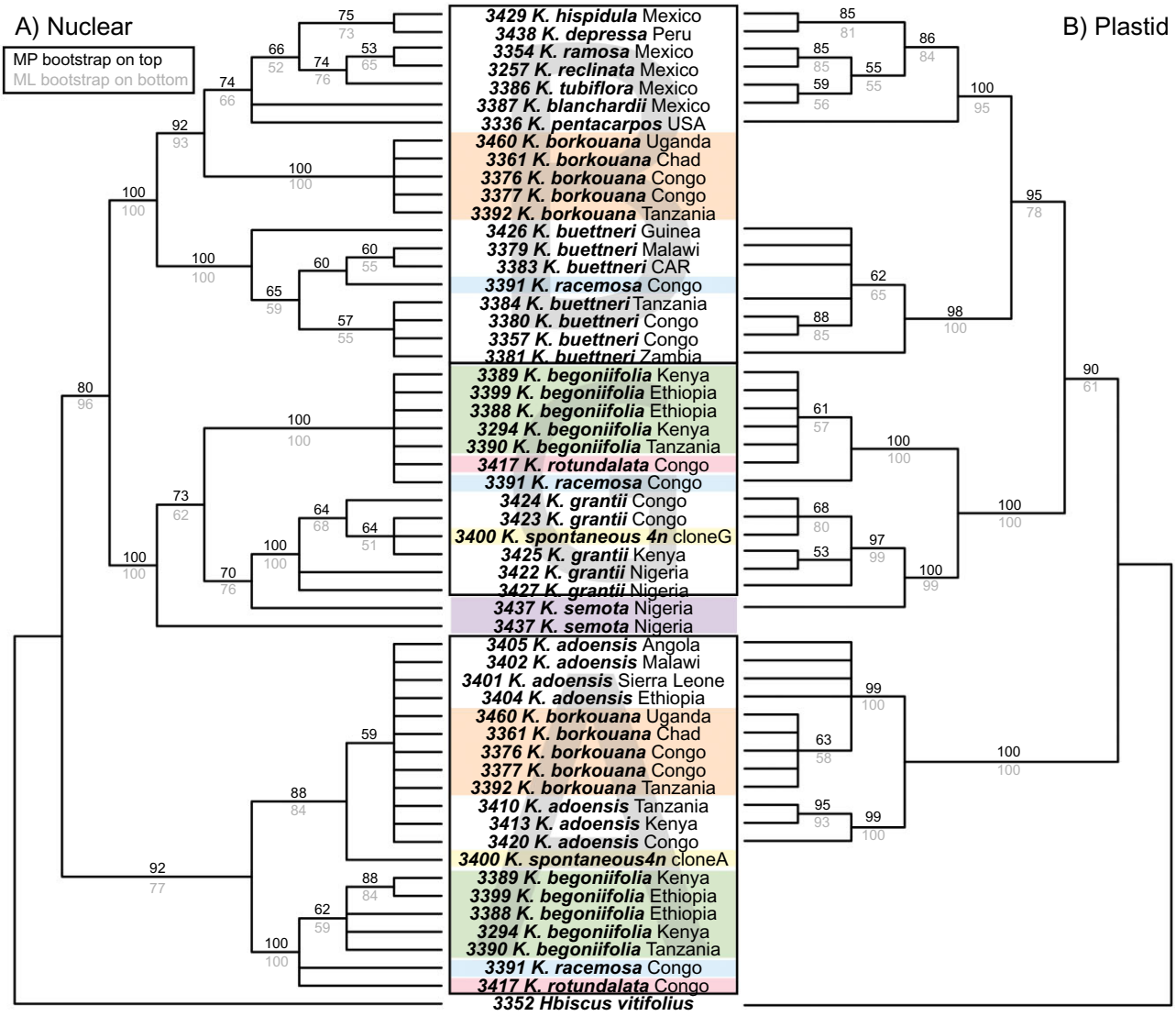

Figure 3. Bootstrap consensus trees, reconciling relationships using plastid (*matK*, *trnL–F*, *rpl32–trnL*, *trnQ–rps16*, and *ycf1*) vs. nuclear (*A1341*, *CesA1*, and *GBSSI*) data. Diploid taxa are uncoloured and each polyploid taxon is denoted by a different colour. Abbreviations: CAR, Central African Republic; Congo, Democratic Republic of the Congo. Black squares with large gray background letters indicate the genome designations of Blanchard in Table 1 (omitted are the X and Y genomes of *K. semota*). Numbers indicate voucher information (see Table 2).

Our results indicate that the African diploid *K. adoensis* is sister to all other diploids in the genus (Fig. 4). *Kosteletzkya grantii*, also from Africa, is the next diploid species to diverge phylogenetically in the genus. The third African diploid, *K. buettneri*, is sister to all of the New World species.

The nuclear data support a hypothesis that *K. semota* has two haplotypes that were closely related

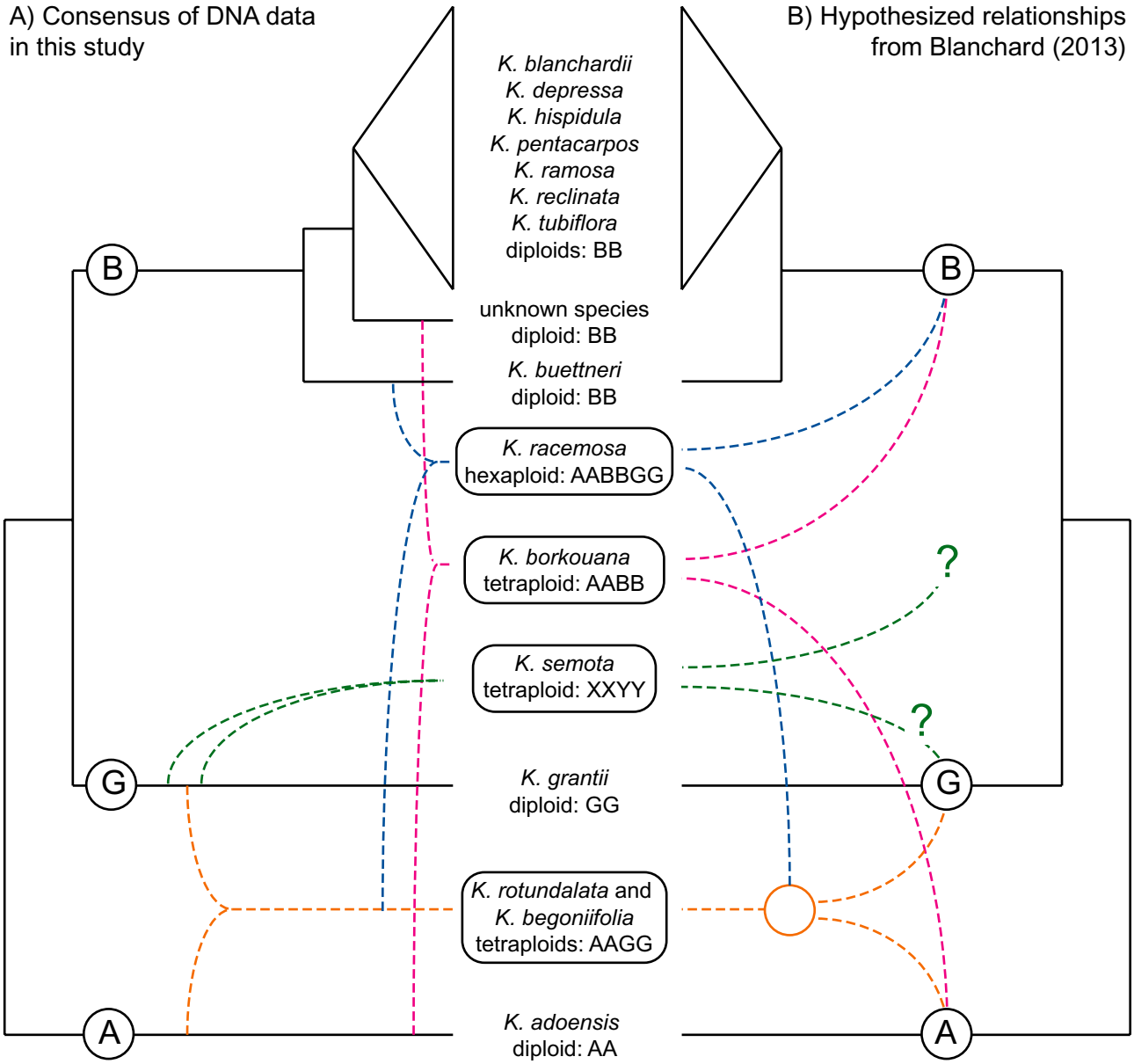

Figure 4. Comparison of (A) a consensus of the phylogenetic analysis of DNA data gathered in this study, and (B) the hypothesized relationships based on chromosome-pairing studies in Blanchard (2013a). Note that the genome notation from Blanchard (2013a) is followed here for consistency (i.e. A, B, and G). Boxes indicate polyploid species.

to *K. grantii*, a diploid species. In *CesA1* and *GBSSI*, these two haplotypes are paraphyletic, whereas in *A1341* the two haplotypes are monophyletic. Plastid data show a sister relationship of *K. semota* to *K. grantii*.

Kosteletzkya borkouana shows two haplotypes, according to *CesA1* and *GBSSI*, one of which corresponds well to extant *K. adoensis* haplotypes and another haplotype that does not match any known *Kosteletzkya* spp., but which is most closely related to the New World species. Clones of *A1341* only recovered the *K. adoensis* haplotype. Plastid data show a clade for this species that is phylogenetically embedded in *K. adoensis*, as in the *A1341* data.

The tetraploids *K. begoniifolia* and *K. rotundalata* are virtually indistinguishable according to the DNA data and display two haplotypes for both *CesA1* and *GBSSI* that are sister to *K. adoensis* and *K. grantii*, respectively. For *A1341*, there is only one haplotype, forming a polytomy that is most closely related to

K. grantii and *K. semota*. Plastid data group these two species together, with *K. racemosa*, in a clade that is sister to *K. grantii* and *K. semota*.

Finally, *K. racemosa* is a hexaploid that has the three haplotypes, as indicated by *CesA1* and *GBSSI*: the two shared by *K. begoniifolia* and *K. rotundalata* and a haplotype identical to *K. buettneri*. According to *A1341*, this species shares the *K. buettneri* haplotype and the haplotype in *K. begoniifolia*/*rotundalata* that is most closely related to *K. grantii*, but is missing the other expected haplotype that would be sister to *K. adoensis*.

The spontaneous tetraploid derived from an artificial cross between *K. adoensis* and *K. grantii* shows two haplotypes, one of which is closely similar to the parental *K. grantii* sample (*Blanchard 3423*) from the Democratic Republic of the Congo according to nuclear and plastid data. However the second haplotype, rather than being close to the other parental sample, *K. adoensis* (*Blanchard 3401*) from Sierra Leone, is sister to a clade containing all seven *K. adoensis* samples and all five *K. borkouana* samples.

DISCUSSION

Chromosomal and genomic makeup

Blanchard (1974, 2012, 2013a) studied chromosome numbers and chromosome pairing in interspecific hybrids for many *Kosteletzkya* spp. The genus (as circumscribed here) has a single known base chromosome number of $x = n = 19$ (Table 1). Experimental crosses among ten diploid species identified three genomes: A (for *adoensis*), B (for *buettneri*), and G (for *grantii*). All seven available New World species were shown to share the B genome with *K. buettneri*. Additional experimental crosses incorporating all of the polyploid species yielded a pattern of chromosome pairing that strongly suggested that each of the polyploids is an allopolyploid (Table 1). The degree of chromosome pairing in interspecific hybrids supports the contention that the African *K. adoensis* was the earliest species to diverge, followed by the two other African diploids *K. grantii* and *K. buettneri*. We postulate a relatively recent trans-Atlantic dispersal to the New World followed by a radiation that gave rise to the seven known diploids in that hemisphere, first suggested by Blanchard (2013a) and confirmed in this study. Blanchard's three genome designations are found to correspond to the three major clades in both plastid and nuclear trees presented here; however, the X and Y genomes of *K. semota* were most closely related to the G genome (Figs 2, 3).

Phylogenetic relationships of diploids

Phylogenetic analysis of plastid data and nuclear data converge on a similar history. When the *Kosteletzkya* spp. from Madagascar are excluded, the genus in the strict sense is monophyletic (Koopman & Baum, 2008) revealing a paraphyletic grade of African species relative to a monophyletic assemblage of New World taxa. The African diploids and their ancestors are the foundation of the polyploid series and of the reticulate diversification of the genus.

The New World *Kosteletzkya* spp. represent a relatively recently derived clade in the genus. Most examples of these species are restricted to Mexico, except *K. pentacarpos* and *K. depressa*. *Kosteletzkya pentacarpos* has long been known as *K. virginica*, a widespread North American coastal species, but it had in fact been merged by Cavanilles in 1787 with *K. pentacarpos*, a species known from western Eurasia (see Blanchard, 2008 for a more in-depth discussion of taxonomy). The other widespread species, *K. depressa*, is found throughout the northern Neotropics and south along the Pacific coast as far as northern Peru. Three other New World diploids, all of restricted distribution and either tabulated or illustrated here but not specifically mentioned in the text, are *K. blanchardii* Fryxell, *K. ramosa* Fryxell and *K. reclinata* Fryxell. The relationships among these New World species will be the focus of a separate paper.

Of the 17 known extant species, two were not sampled: *K. batacensis* and *K. thurberi*. *Kosteletzkya thurberi* is a poorly known Mexican endemic. We hypothesize that it is closely related to *K. tubiflora* because of similarities in geography and morphology. The other unsampled species, *K. batacensis*, is even more poorly known. It is native to the island of Luzon in the Philippines and is known from few collections. Its distribution is unique in the genus. Several workers (Merrill, 1909; van Borssum-Waalkes, 1966) hypothesized that *K. batacensis* was simply a naturalized introduction from the New World. However, Blanchard (2008) noted that the morphology of *K. batacensis* did not correspond to any other known *Kosteletzkya* spp.

Phylogenetic relationships of polyploids

The combinatory, biparental signal of hybrids can complicate the understanding of both their parentage and evolutionary placement (McDade, 1992). However, the data in this study were produced by extensively sampling clones and careful selection of different parental haplotypes, so the various phylogenetic signals produced are fairly clear.

Polyploidy is a common characteristic of many malvaceous groups including *Gossypium* L. (Small & Wendel, 2000; Cronn *et al*., 2002b; Alvarez, Cronn & Wendel, 2005), *Tarasa* Phil. (Tate & Simpson, 2003), *Malva* L. (Escobar García *et al*., 2009), *Palaua* Cav. (Schneider *et al*., 2011), and in *Hibiscus* section

Furcaria DC. (Wilson, 1994, 2006). The study of allopolyploidy has received much more attention than autopolyploidy and although it is intuitive to assume that autopolyploidy could lead to greater genomic reduction (i.e. of homologous gene copies) because of its relatively higher level of redundancy, there is little evidence to support this idea (Parisod, Holdregger & Brochmann, 2010). Some gene loss following polyploidization appears common in various organisms over different spans of time (Soltis & Soltis, 1995, 1999; Adams & Wendel, 2005; Sehrish *et al*., 2014). The absence (i.e. lack of recovery via PCR) of some homoeologues in the nuclear loci sampled in this study may represent the loss of some homoeologues (Figs S1–S3).

Polyploidy in *Kosteletzkya* has been well documented by Blanchard (2012; Table 1). Diploid (n = 19), tetraploid (n = 38), and hexaploid (n = 57) species exist in this genus. In general, the number of haplotypes (as determined by phylogenetic analysis) according to the nuclear DNA data in the polyploids was directly related to the ploidy (i.e. diploids have one copy, tetraploids have two copies and hexaploids have three copies). Therefore, these data represent an independent confirmation of the number of chromosome sets in the polyploids, and of hypotheses of relationships. Based on divergent unique haplotypes, most examples of *Kosteletzkya* polyploids are of hybrid origin, each independent and occurring at different times in the past.

Kosteletzkya borkouana is a tetraploid species. Analyses of these DNA data support an allotetraploid origin that was relatively recent. The five accessions of this species share haplotypes that are little differentiated from extant individuals of *K. adoensis*. We hypothesize that the other haplotype from *K. borkouana* is from an unknown B-genome species, either an extinct species or an extant species that remains unknown. Morphological intermediacy of this species between the two parental species is also impossible to evaluate because of the unknown parent. Nonetheless, the hypothesis that *K. borkouana* is an allotetraploid with a genomic signature of AABB is supported by these data. It is noteworthy that all five samples of *K. borkouana* share a unique plastid haplotype consistent with a single hybridization event that led to the creation of this species.

Kosteletzkya semota is a tetraploid. Blanchard (2013a) was uncertain of the parental origin of this species, since neither set of chromosomes clearly corresponded to other diploid species. For this reason, he designated two additional genome types: X and Y. Both haplotypes recovered in phylogenetic analyses are closely related to *K. grantii*. Because nuclear haplotypes of this species are closely related according to the phylogenetic data, we hypothesize that *K. semota* might be of autotetraploid origin. Because of the relatively large phylogenetic divergence of both plastid and nuclear data, we infer that the polyploidy event that led to the isolation of this species is relatively ancient.

Kosteletzkya begoniifolia and *K. rotundalata* clearly share a common allotetraploid hybrid origin. Their most recent common ancestor is shared with *K. adoensis* and *K. grantii*, but they do not share identical haplotypes with these two diploid species. In fact, each haplotype shows a similar signature of phylogenetic divergence from haplotypes sampled of the diploid species. Therefore, we hypothesize that this hybrid lineage consists of the two species *K. begoniifolia* and *K. rotundalata* resulting from a relatively ancient hybridization. If these are distinct species, there is little divergence between them and their monophyly is inconclusive. Plastid data fail to distinguish them, but they have a distinct haplotype that is mutually exclusive of any other species, except for *K. racemosa*.

Kosteletzkya racemosa is the only known hexaploid in the genus. The data strongly support the parentage of this species as being between the tetraploid lineage of either *K. begoniifolia* or *K. rotundalata* and the diploid *K. buettneri*. The species haplotypes are not strongly diverged from any of the parental haplotypes, supporting our hypothesis that the hybridization event that led to this species is recent relative to the lineage that led to *K. begoniifolia* and *K. rotundalata*.

One of the highly sterile interspecific hybrids grown by Blanchard (2013a), *K. adoensis* × *K. grantii*, spontaneously produced fertile offspring from seeds of a single flower and they proved on cytological examination to be tetraploids ('*K.* spontaneous 4*n*' in Figs 2, 3 here). Test backcrosses of these offspring to their two parents confirmed that the tetraploids were true allopolyploids. Moreover, hybrids produced from crosses between these allotetraploids and the two tetraploid species *K. begoniifolia* and *K. rotundalata* showed nearly complete chromosome pairing. This strongly suggested that *K. grantii* and *K. adoensis*, or at least plants of G and A genomic ancestry, were the progenitors of the two tetraploids. Results from the present study corroborate Blanchard's (2013a) results in that the two haplotypes recovered reside in the A and G clades. However as would be expected, the spontaneous tetraploid plant clusters more closely with its two parent species than with either *K. begoniifolia* or *K. rotundalata*, testimony to the recent nature of the spontaneous plant vs. the antiquity of the two wild species.

A question arises as to whether chromosome-pairing evidence is consistent with molecular phylogenetic data in other genera. The well studied malvaceous genus *Gossypium*, like *Kosteletzkya*, has

an African–New World distribution, multiple African genomes, both diploid and polyploid species, and evidence of a trans-Atlantic dispersal of one of its genomes. (Unlike *Kosteletzkya*, *Gossypium* includes in addition several Australian species.) An examination of known chromosome-pairing relationships in African diploid *Gossypium* indicates a pattern among genomes A, B, and E (Konan *et al*., 2009) that is similar to the pattern among the genomes B, G, and A in *Kosteletzkya*. This same pattern can be found in molecular phylogenetic studies of *Gossypium* (Cronn & Wendel, 2004) and is moreover evidenced by morphological analysis (Fryxell, 1971).

CONCLUSIONS

Our data provide further evidence of the importance of polyploidy as a significant mode of speciation in the evolution of Malvoideae (Malvaceae), a subfamily in which at least 31% of the non-monotypic genera include species at different ploidies (Blanchard, 2012). By estimating relationships among plant species, especially in the context of polyploid species, we can better understand the rich and complex evolutionary history that can develop from hybridization and we can better understand the mechanisms driving evolution in groups such as *Kosteletzkya*.

ACKNOWLEDGEMENTS

We thank Barbara Sue Carlsward for technical support. Computation time was provided by the Florida Museum of Natural History Phyloinformatics Cluster for High Performance Computing in the Life Sciences funded by grants from the United States National Science Foundation awarded to Pamela and Douglas Soltis with technical assistance provided by Matt Gitzendanner. Savita Shanker and Patrick Thimote at the Interdisciplinary Center for Biotechnology Research at University of Florida provided sequencing services and the University of Florida Department of Biology and Jeff Hubbard provided greenhouse resources. Specimen curation was provided by Kent Perkins and Norris Williams at the University of Florida Herbarium (FLAS) at the Florida Museum of Natural History. Stuart McDaniel was supported by start-up funds from the University of Florida.

REFERENCES

Adams KL, Wendel JF. 2005. Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology* **8:** 135–141.

Alvarez I, Cronn R, Wendel JF. 2005. Phylogeny of the New World diploid cottons (*Gossypium* L., Malvaceae) based on sequences of three low-copy nuclear genes. *Plant Systematics and Evolution* **252:** 199–214.

Blanchard OJ Jr. 1974. Chromosome numbers in *Kosteletzkya* Presl (Malvaceae). *Rhodora* **76:** 64–66.

Blanchard OJ Jr. 2008. Innovations in *Hibiscus* and *Kosteletzkya* (Malvaceae, Hibisceae). *Novon* **18:** 4–8.

Blanchard OJ Jr. 2012. Chromosome numbers, phytogeography, and evolution in *Kosteletzkya* (Malvaceae). *Rhodora* **114:** 37–49.

Blanchard OJ Jr. 2013a. Experimental hybridization, chromosome pairing, phylogeny and phytogeography in *Kosteletzkya* (Malvaceae: Malvoideae). *Comparative Cytogenetics* **7:** 73–101.

Blanchard OJ Jr. 2013b. A new species of *Kosteletzkya* (Hibisceae, Malvoideae, Malvaceae) and notes on a second species, both from eastern Democratic Republic of the Congo. *Phytotaxa* **88:** 31–37.

van Borssum-Waalkes J. 1966. Malesian Malvaceae revised. *Blumea* **14:** 1–213.

Chase MW, Hills HG. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* **40:** 215–220.

Coyne JA, Orr HA. 2004. *Speciation*. Sunderland: Sinauer Associates.

Cronn R, Cedroni M, Haselkorn T, Grover C, Wendel JF. 2002a. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theoretical and Applied Genetics* **104:** 482–489.

Cronn R, Small RL, Haselkorn T, Wendel JF. 2002b. Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. *American Journal of Botany* **89:** 707–725.

Cronn R, Wendel JF. 2004. Cryptic trysts, genomic mergers, and plant speciation. *New Phytologist* **161:** 133–142.

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

van Devender TR, Calder WA, Krebbs K, Reina-G AL, Russell SM, Russell RO. 2004. Hummingbird plants and potential nectar corridors for the rufous hummingbird in Sonora, Mexico. In: Nabhan GP, ed. *Conserving migratory pollinators and nectar corridors in western North America*. Tucson: University of Arizona Press and Arizona-Sonora Desert Museum, 96–121.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* **19:** 11–15.

Escobar García P, Schönswetter P, Fuertes Aguilar J, Nieto Feliner G, Schneeweiss GM. 2009. Five molecular markers reveal extensive morphological homoplasy and reticulate evolution in the *Malva* alliance (Malvaceae). *Molecular Phylogenetics and Evolution* **50:** 226–239.

Evans RC, Alice LA, Campbell CS, Kellogg EA, Dickinson TA. 2000. The granule-bound starch synthase (GBSSI) gene in the Rosaceae: multiple loci and phylogenetic utility. *Molecular Phylogenetics and Evolution* **17:** 388–400.

Fazekas AJ, Kuzmina ML, Newmaster SG, Hollingsworth PM. 2012. DNA barcoding methods for land

plants. In: Kress WL, Erickson DL, eds. *DNA barcodes: methods and protocols*. New York: Humana Press.

Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39:** 783–791.

Fitch WM. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology* **20:** 406–416.

Fryxell PA. 1971. Phenetic analysis and the phylogeny of the diploid species of *Gossypium* L. (Malvaceae). *Evolution* **25:** 554–562.

Grant VP. 1981. *Polyploidy*. New York: Columbia University Press.

Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS, Soltis DE, Clifton SW, Schlarbaum SE, Schuster SC, Ma H, Leebens-Mack J, dePamphelis CW. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature* **473:** 97–100.

Konan NO, Baudoin J-P, D'Hont A, Mergeai G. 2009. Bridging classical and molecular cytogenetics of *Gossypium*. In: Paterson AH, ed. *Plant genetics and genomics: crops and models*, Vol. *3*. New York: Springer.

Koopman MM, Baum DA. 2008. Phylogeny and biogeography of tribe Hibisceae (Malvaceae) on Madagascar. *Systematic Botany* **33:** 364–374.

La Duke JC, Doebley J. 1995. A chloroplast DNA based phylogeny of the Malvaceae. *Systematic Botany* **20:** 259–271.

Leitch IJ, Bennett MD. 1997. Polyploidy in angiosperms. *Trends in Plant Science* **2:** 470–476.

Liesack W, Weyland H, Stackebrandt E. 1991. Potential risks of gene amplification by PCR as determined by 16S rDNA analysis of a mixed-culture of strict barophilic bacteria. *Microbial Ecology* **21:** 191–198.

McDade LA. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* **46:** 1329–1346.

McDaniel SF, Shaw AJ. 2005. Selective sweeps and intercontinental migration in the cosmopolitan moss *Ceratodon purpureus* (Hedw.) Brid. *Molecular Ecology* **14:** 1121–1132.

Merrill ED. 1909. New or noteworthy Philippine plants, VII. *Philippine Journal of Science* **4:** 247–330.

Neubig KM, Whitten WM, Carlsward BS, Blanco MA, Endara L, Williams NH, Moore M. 2009. Phylogenetic utility of *ycf1* in orchids: a plastid gene more variable than *matK*. *Plant Systematics and Evolution* **277:** 75–84.

Parisod C, Holdregger R, Brochmann C. 2010. Evolutionary consequences of autopolyploidy. *New Phytologist* **186:** 5–17.

Pfeil BE, Brubaker CL, Craven LA, Crisp MD. 2002. Phylogeny of *Hibiscus* and the tribe Hibisceae (Malvaceae) using chloroplast DNA sequences of *ndhF* and the *rpl16* intron. *Systematic Botany* **27:** 333–350.

Pfeil BE, Crisp MD. 2005. What to do with *Hibiscus*? A proposed nomenclatural resolution for a large and well known genus of Malvaceae and comments on paraphyly. *Australian Systematic Botany* **18:** 49–60.

Rambaut A. 1996. Se-Al: sequence alignment editor, v2.0a11. Available at: http://tree.bio.ed.ac.uk/software/seal/ (last accessed 21 August 2015).

Ruan C-J, Xing W-H, Teixeira da Silva J. 2012. Potential of five plants growing on unproductive agricultural lands as biodiesel resources. *Renewable Energy* **41:** 191–199.

Schneider JV, Schulte K, Fuertes Aguilar J, Huertas ML. 2011. Molecular evidence for hybridization and introgression in the Neotropical coastal desert-endemic *Palaua* (Malveae, Malvaceae). *Molecular Phylogenetics and Evolution* **60:** 373–384.

Sehrish T, Symonds VV, Soltis DE, Soltis PS, Tate JA. 2014. Gene silencing via DNA methylation in naturally occurring *Tragopogon miscellus* (Asteraceae) allopolyploids. *BMC Genomics* **15:** 701.

Shaw J, Lickey EB, Schilling EE, Small RL. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* **94:** 275–288.

Small RL. 2004. Phylogeny of *Hibiscus* sect. *Muenchhusia* (Malvaceae) based on chloroplast *rpL16* and *ndhF*, and nuclear ITS and GBSSI sequences. *Systematic Botany* **29:** 385–392.

Small RL, Wendel JF. 2000. Phylogeny, duplication, and intraspecific variation of *Adh* sequences in New World diploid cottons (*Gossypium* L., Malvaceae). *Molecular Phylogenetics and Evolution* **16:** 73–84.

Soltis DE, Soltis PS. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* **12:** 243–273.

Soltis DE, Soltis PS. 1995. The dynamic nature of polyploid genomes. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 8089–8091.

Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology and Evolution* **14:** 348–352.

Soltis DE, Soltis PS, Pires JC, Kovarik A, Tate JA, Mavrodiev E. 2004. Recent and recurrent polyploidy in *Tragopogon* (Asteraceae): cytogenetic, genomic and genetic comparisons. *Biological Journal of the Linnean Society* **82:** 485–501.

Soltis PS. 2005. Ancient and recent polyploidy in angiosperms. *New Phytologist* **166:** 5–8.

Stebbins GL. 1985. Polyploidy, hybridization, and the invasion of new habitats. *Annals of the Missouri Botanical Garden* **72:** 824–832.

Swofford DL. 1999. *PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0b10*. Sunderland: Sinauer Associates.

Taberlet P, Gielly L, Pautou G, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17:** 1105–1109.

Tate JA, Ni Z, Scheen A-C, Koh J, Gilbert CA, Lefkowitz D, Chen ZJ, Soltis PS, Soltis DE. 2006. Evolution and

expression of homeologous loci in *Tragopogon miscellus* (Asteraceae), a recent and reciprocally formed allopolyploid. *Genetics* **173:** 1599–1611.

Tate JA, Simpson BB. 2003. Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Botany* **28:** 723–737.

Wilson FD. 1994. The genome biogeography of *Hibiscus* L. section *Furcaria* DC. *Genetic Resources and Crop Evolution* **41:** 13–25.

Wilson FD. 2006. A distributional and cytological survey of the presently recognized taxa of *Hibiscus* section *Furcaria* (Malvaceae). *Bonplandia (Corrientes)* **15:** 53–62.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. Maximum likelihood phylogram from the analysis of the nuclear *A1341* data set.
Figure S2. Maximum likelihood phylogram from the analysis of the nuclear *CesA1* data set.
Figure S3. Maximum likelihood phylogram from the analysis of the nuclear *GBSSI* data set.
Figure S4. Maximum likelihood phylogram from the analysis of the plastid *matK* data set.
Figure S5. Maximum likelihood phylogram from the analysis of the plastid *rpl32–trnL* data set.
Figure S6. Maximum likelihood phylogram from the analysis of the plastid *trnL–F* data set.
Figure S7. Maximum likelihood phylogram from the analysis of the plastid *trnQ–rps16* data set.
Figure S8. Maximum likelihood phylogram from the analysis of the *ycf1* data set.
Table S1. Tabulation of clones sequenced for the three nuclear loci of the polyploid taxa. For *A1341* and *CesA1*, all nuclear PCR products were cloned. For *GBSSI*, some accessions for each species were cloned and sequenced; then after homoeologue-specific primers were designed, amplification of all three copies (i.e. from genomes A, B, and G using primer names that end with adoe, buett and grant, respectively; see Table 3) via PCR was attempted with the recovered number of copies amplified and sequenced in parentheses.