Biol Invasions (2022) 24:2613–2634
https://doi.org/10.1007/s10530-022-02798-3

ORIGINAL PAPER

The phytogeography and genetic diversity of the weedy hydrophyte, *Pistia stratiotes* L.

Paul T. Madeira · F. Allen Dray Jr. · Philip W. Tipping

Received: 4 June 2021 / Accepted: 16 December 2021 / Published online: 12 May 2022

Abstract Understanding the origins and genetic relationships of invasive, non-native species is critical to informing conservation and management practices. *Pistia stratiotes* is one such species—a pantropical floating plant that is problematic in many regions of the world, including Florida, USA. Questions surrounding the origins of *P. stratiotes* populations in Florida and elsewhere prompted a molecular investigation using five chloroplast and one mitochondrial DNA sequences. A total of 154 samples were collected from 14 countries. The sequence data was analyzed using haplotype network analysis, maximum likelihood phylogenetics and species delimitation tools. These data show that *P. stratiotes* comprises a minimum of seven distinct haplotypic clades worldwide, three of which differ enough to likely represent different species. Florida, which was more heavily sampled than other regions of the world, contains four of the clades—one of which shows evidence of being pan-Caribbean with sufficient variation to suggest regional (including Florida) nativity. A second clade, present in the U.S. Gulf States and California, may be native within this range, however more sampling is needed to fully describe its distribution and nativity. Another clade, predominant in southern Florida and the St. Johns River, likely originated in South America. Results are discussed in the broader context of the effects of cryptic species on weed management, including biological control efforts.

Keywords *Pistia stratiotes* · Weed management · Biological control · Cryptic species · Molecular phylogenetics · TCS analysis · Species delimitation analysis

Introduction

Managing invasive plants is a challenging endeavor (Liebman et al. 2016; Rolfe and Windle 2014). Funding is generally insufficient, and other non-native species often replace any invader brought under management control (Hanley and Roberts 2019). Management tools that can offer hope also warrant caution. For example, cultural controls effective on a local level (e.g., hand pulling of seedlings) may not translate to a landscape scale (Myers and Bazely 2003). Mechanical controls suffer from a well-known lack of specificity (Center et al. 2002). Chemical controls also lack specificity, can be both mobile and persistent, and can be thwarted by emergence of resistant plant genotypes (Boutin et al. 2014; Gould et al. 2018). Biological controls are generally viewed as environmentally benign, but ecologists debate the

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10530-022-02798-3.

P. T. Madeira · F. A. Dray Jr. (✉) · P. W. Tipping
ARS Invasive Plant Research Laboratory, USDA, 3225 College Avenue, Fort Lauderdale, FL 33314, USA
e-mail: Allen.Dray@usda.gov

potential for unanticipated collateral effects (Hinz et al. 2020; Messing and Wright 2006). These overarching challenges are exacerbated when the invader is a so-called cryptic species.

Canavan et al. (2020) identify two classes of cryptic invaders: intraspecific—foreign genotypes of a locally native species, and interspecific—a non-native misidentified as a native or as another non-native. Cryptic invaders can form expansive populations, at times outcompeting native congeners. Admixture with native genotypes (heterosis) may produce phenotypic variants more aggressive than the natives. Each of these outcomes creates further challenges for managers. Chemical and mechanical controls are likely to impact both invasive and native genotypes, and perhaps ecological congeners as well (Boutin et al. 2014). Employing classical biological controls against invaders with native genotypes is also highly problematic. Most weed biocontrol practitioners and regulatory agencies insist that any agents should be specialists at the genotypic level (targeting only the non-native genotype) as opposed to the more traditional species level (attacking an invasive species, but not its native congeners) (Casagrande et al. 2018; Pemberton 2000). Finally, distinguishing between the native, non-native, and hybrid genotypes in real time can be extremely difficult (Morais and Reichard 2018; Saltonstall 2002). Land managers thus need tools that parse cryptic versus native species to inform their decisions.

Waterlettuce (*Pistia stratiotes* L.; Araceae) is potentially one such cryptic species. Populations of this floating hydrophyte can become sufficiently large to interfere with human endeavors (e.g., transportation, irrigation, rice farming, fishing activities, water treatment and biodiversity) resulting in management efforts on at least four continents and its listing as one of the world's worst weeds (Holm et al. 1977). The genus *Pistia* is dated from 40 (Nauheimer et al. 2012) to 90 (Renner and Zhang 2004) million years old. Although fossil records exist from many regions that are currently temperate, its distribution today is largely pan-tropical (Dray and Center 2002, and references therein). Historically waterlettuce is known at least as early as 300 BCE from Egypt and northern Africa where Theophrastus reported its use medicinally (Theophrastus, translated 1916); Pickering (1879) reports the plant from even earlier Egyptian heiroglyphics. Thus, extant waterlettuce populations have frequently been considered to derive from Old World origins. *Pistia* was first recorded in the United States by eighteenth century naturalists John and William Bartram who described and illustrated it from plants observed in abundance on the St. John's River in northern Florida (Harper 1998). This early sighting led some botanists to consider waterlettuce a North American native (Stuckey and Les 1984). However, its biology, and apparent lack of apparent seed production in North America, caused other twentieth century botanists to consider the plant a non-native invader (Godfrey and Wooten 1979; Weldon et al.1969) that was introduced by transcontinental bird migrations (Stoddard 1989) or in ships' ballast during early European colonization of North America (Stuckey and Les 1984; Schmitz et al. 1993; Dray and Center 2002). Despite this difference of opinion, it's status as a management nuisance led federal and state authorities in 1977 to approve the addition of waterlettuce to the U.S. Corps of Engineers Aquatic Plant Control Program for the state of Florida (U.S. Army Corps of Engineers 1977).

Several lines of evidence point to the presence of multiple genotypes of *P. stratiotes* in Florida. Dray and Center (1989) reported that, counter to prior claims, seed production was present in Florida—at Loxahatchee National Wildlife Refuge and sites around Lake Okeechobee. Further, biomass, leaf production, and flowering phenology of waterlettuce at some sites in St. Lucie County, FL, differed substantially from populations in and around Lake Okeechobee (DeWald and Lounibos 1990; Dray and Center 1992). Additionally, the Argentine weevil *Neohydronomus affinis* Hustache (Coleoptera: Curculionidae), released as a biological control of waterlettuce (Dray et al. 1990), differed in efficacy between these same plant populations (Dray and Center 1992). Another biological control agent, *Spodoptera pectinicornis* Hampson (Lepidoptera: Pyralidae) from Thailand, did not establish (Dray et al. 2001)—possibly as a result of an incorrect genotypic matching between this Asian moth and the plants in Florida. More recently, the discovery of fossil seeds in 12,600 year old sediments from Lake Annie, Florida (Quillen et al. 2013), caused Evans (2013) to suggest that *P. stratiotes* was native to Florida and had persisted in refugia provided by thermal springs throughout the climatic fluctuations associated with the Pleistocene and severe drought conditions that dominated the

state between 11,000–7000 years ago (Grimm et al. 1993; Quillen 2009; Quillen et al. 2013; Watts and Hansen 1994).

The complex nature of the evidence, coupled with the potential to develop additional biological controls against this weedy hydrophyte, led us to investigate the diversity and biogeographical origins of *P. stratiotes*. *Our null hypotheses were (A) P. stratiotes is non-native to Florida, and (B) there is only one extant species of P. stratiotes worldwide. To test these hypotheses*, we:

1. collected samples world-wide, especially in Florida, and determined the haplotypes of these specimens;
2. conducted a haplotype network analysis (TCS) to visualize the relationships between samples and determine the extent of any genetic diversity;
3. mapped the distribution of the haplotypes to parse questions on the nativity of populations in Florida;
4. constructed phylogenies to (a) compare the genetic variation within the genus *Pistia* to variation 'within' and 'between' other genera in the family Araceae, and (b) to generate greater resolution of the *Pistia* clade than has heretofore been available; and
5. examined the degree to which different *Pistia* clades contain evidence of species delimitation.

Methods

Sample acquisition, sequencing and haplotype determination

Molecular analysis of the Araceae phylogeny has been addressed previously by numerous authors utilizing varied sequences: Renner and Zhang (2004) [*trnL intron*, *trnL-F spacer*, *rpl20-rsp12 spacer*, *nad1 b/c intron*], Cabrera et al (2008) [*trnL intron*, *trnL-F spacer*, *trnK/matK*, *rbcL*], Cusimano et al. (2011) [*trnK/matK*], Nauheimer et al. (2012) [*trnL intron*, *trnL-F spacer*, *trnK/matK*, *rbcL*], and others. The phylogenetic analysis presented here utilized NCBI sequences as outgroup (non-*Pistia*) sequences drawn from what Nauheimer et al. (2012) referred to as the "*Pistia* clade", those most closely related to *Pistia*. *Pistia* sequences for the phylogeny and population analyses were generated in-house. Both sets of NCBI sequences are presented in Supplementary Table 1 For the haplotype identification and population analysis of *Pistia* we used the chloroplast *trnL intron*, *trnLF spacer*, *matK* and *rpL32-trnL*. The second, *Pistia* only, phylogenetic analysis was expanded to 6 sequences with chloroplast *rpl20-rsp12* and mitochondrial *nad1 exons b, c and intron* added. Each study utilized separate alignments. The *trnL intron*, *trnL-F spacer*, *mat K* and *rpL32-trnL* were sequenced for all *Pistia* samples. The *rpl20-rsp12* spacer and *nad1 b/c intron* were only sequenced for geographically representative samples of each haplotype obtained in the four sequence study. These samples are identified by bold lettering in Table 1.

Study samples were solicited from scientists and land managers around Florida, the USA, and the world. Sample packets with 30 g of desiccant, sample ID card, instructions and a return envelope were sent to collectors. Collectors were instructed to sample a single young leaf without the stem and with as little insect damage as possible, blot it with the supplied paper towel and insert it into the sample packet. Upon arrival at our lab, samples were placed on fresh silica gel and stored frozen. Some samples were directly collected by USDA personnel, including the "BA" samples. Dried samples were ground using a 'frozen' mortar and pestle followed by DNA extraction using the E.Z.N.A HP Plant DNA Mini Kit D2485-01 (Omega Bio-Tek, Norcross, GA). Sample information is presented in Table 1 and in more detail in Supplementary Table 2.

Reaction mixes for *matK*, *trnL intron*, *trnLF spacer* and *rpl20-rps12 intergenic spacer* each contained 1X reaction buffer, 1.5 mM $MgCl_2$, 0.35 mM of each primer, 0.2 mM dNTPs, 1X CES (a PCR enhancer, Ralser et al. 2006), 2.5 units/rxn BioReady rTaq DNA polymerase (Bulldog Bio, Inc., Portsmouth, NH); and 40 ng genomic DNA in a 50-ml reaction. *Nad1* and the *rpl32-trnL intergenic spacer* both contained 'premix' 2X Terra PCR Direct Buffer (final concentration at 2 mM Mg^{+2}, 0.4 mM dNTP), an additional 0.5 mM MgCl2 (LAMBDA Biotech Inc. St. Louis, MO), 0.35 µM each primer, and 1.25 units Terra PCR Direct Taq Polymerase Mix. The number of cycles was normally 35X. The primer sets used included: for the *trnL intron*- primers trnC and trnD, and for the *trnLF spacer*- primers trnE and trnF, both

Table 1 Sample haplotype and origin IDs

Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)	Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)
487	A1	GY-ES	M Grodowitz	6.716890	− 58.207980	**CH10**	C1	CN-GX	G Wheeler	24.948700	110.466220
486	A1	GY-ES	R Chan-dranauth	6.845730	− 58.232583	CH2	C1	CN-HB	Y Wang	30.533453	114.418218
US2	A1	US-FL	J Leidi, M Rayamahji	29.569000	− 82.065417	CH11	C1	CN-YN	F A Dray	22.776900	100.970400
FL38	A1	US-FL	P Tipping	26.175690	− 80.447260	CH6	C1	CN-YN	G Wheeler	24.350510	102.762290
FL61	A1	US-FL	J Thompson	26.907777	− 82.288333	1b	C1	FR-OCC	M Julien	43.653417	3.879536
FL18	A1	US-FL	M Rumbach, S Clem	26.373320	− 81.608970	**2**	C1	FR-OCC	MJulien	43.580556	3.943333
FL5	A1	US-FL	M Owen, K Relish,	26.028819	− 81.397331	**9**	C1	ZA-EC	J Coetzee	− 33.772223	25.385345
US3	A1	US-FL	J Leidi, M Rayamahji	29.328250	− 83.140083	5	C1	ZA-GP	A Boiunes	− 25.727500	28.236944
FL42	A1	US-FL	S Tonjes	29.299130	− 81.372580	10	C1	ZA-KZN	J Coetzee	− 29.708626	30.872055
US1b	A1	US-FL	P Madeira	26.084103	− 80.240419	7	C1	ZA-KZN	J Coetzee	− 30.066667	30.866667
FL49	A1	US-FL	E L de la Vega	26.720680	− 81.693370	6	C1	ZA-LP	A Boiunes	− 23.828056	30.218611
FL53	A1	US-FL	P Madeira	26.932728	− 81.342350	**ThC**	C1	TH-10	M Rayamahji	13.852150	100.574983
FL10	A1	US-FL	D Sowell	26.529190	− 81.299540	**FL56**	C1	US-FL	D Fussell	30.251681	− 84.150468
FL17	A1	US-FL	J Huckabee	26.602440	− 81.368980	**SC1**	C1	US-FL	D Hood	32.952466	− 80.021330
FL9	A1	US-FL	J Huckabee	26.490080	− 81.140180	**Mx1**	D1	MX-JAL	M M Jiménez	20.282276	− 103.321006
BA15b	A1	US-FL	FA Dray, S Goldstein	27.388016	− 81.434278	**Br7**	D2	BR-BA	G Wheeler	− 12.548780	− 38.682410
BA16b	A1	US-FL	FA Dray, S Goldstein	27.447020	− 81.264768	**Cu4**	E1	CU-15	D Giardina	22.793361	− 83.008254
FL16	A1	US-FL	E Boughton, E Menges	27.391328	− 81.426228	Mx4	E1	MX-TAM	A F Cabo	22.463663	− 97.910062
FL24	A1	US-FL	D Stone	28.903045	− 81.825453	**Mx2**	E1	MX-YUC	J Callahan	20.089075	− 89.551101
FL25	A1	US-FL	S Tonjes	28.923170	− 81.549770	N1	E1	NI-BO	E Sequeira	12.442500	− 85.517100
US4	A1	US-FL	J Leidi, M Rayamahji	30.431060	− 83.597500	**N2**	E1	NI-BO	E Sequeira	12.382731	− 85.879180
FL3	A1	US-FL	M Perez	27.476389	− 82.301811	FL62	E1	US-FL	J Thompson	27.034250	− 82.203022
FL14	A1	US-FL	R Rossmanith,	27.028830	− 80.172960	BA19a	E1	US-FL	FA Dray, S Goldstein	27.620493	− 81.804351
FL11	A1	US-FL	L Proenza	28.500930	− 80.915270	FL1	E1	US-FL	J Mosley	27.620500	− 81.804190

Table 1 (continued)

Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)	Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)
FL32	A1	US-FL	L Proenza	28.447990	− 80.937790	FL8	E1	US-FL	J Huckabee	26.596278	− 81.409624
FL33	A1	US-FL	T Boyette	29.664970	− 81.695250	**BA13B**	E1	US-FL	FA Dray, S Goldstein	26.571784	− 81.823944
BA17b	A1	US-FL	FA Dray, S Goldstein	27.912949	− 81.311392	FL45	E1	US-FL	L Clark	26.571670	− 81.823890
FL37	A1	US-FL	R Young	26.382620	− 80.880970	FL48	E1	US-FL	EL de la Vega	26.471530	− 81.852320
FL6	A1	US-FL	T Wolf	26.713797	− 80.321675	FL43	E1	US-FL	S Hayes, A Lloyd	28.254371	− 82.241356
US5	A1	US-FL	T Center, FA Dray	26.706583	− 80.800017	**FL36**	E1	US-FL	D Donaghy	27.265610	− 82.288330
US6	A1	US-FL	T Center, FA Dray	26.752183	− 80.729233	FL60	E1	US-FL	B Nelson	27.064399	− 80.286990
FL63	A1	US-FL	J Howard	27.663717	− 81.377633	FL64	E1	US-FL	B Nelson	27.135223	− 82.255799
FL46	A1	US-FL	P Tipping	29.691060	− 81.572110	PR1	E1	US-PR	J Crossland	18.331678	− 67.249575
FL50	A1	US-FL	G Eby	28.726280	− 81.263950	PR2	E1	US-PR	J Crossland	18.325525	− 65.981464
BA05b	A1	US-FL	FA Dray, P Madeira	28.714520	− 81.035847	FL57	E2	US-FL	C Cummins	28.403351	− 80.787633
FL51	A1	US-FL	G Eby	28.731620	− 81.034570	FL12	E2	US-FL	C Werner	28.821886	− 82.181792
FL52	A1	US-FL	H Morgenstern	29.386110	− 81.362500	FL59	E2	US-FL	D Morse	30.339610	− 83.991210
BA06b	A2	US-FL	FA Dray, P Madeira	27.381246	− 80.305199	FL7	E2	US-FL	S Schmidt	27.800930	− 82.144870
BA07b	A2	US-FL	FA Dray, P Madeira	27.272589	− 80.354861	**BA12B**	E2	US-FL	FA Dray, P Madeira	30.339712	− 83.991000
FL66	A2	US-FL	D Rodgers	27.381152	− 80.306610	FL67	E2	US-FL	D Fussell	30.474528	− 84.358389
US7	A2	US-FL	T Center, FA Dray	27.271483	− 81.364250	FL44	E2	US-FL	K Morin	28.904410	− 82.593250
US8	A2	US-FL	T Center, FA Dray	27.272483	− 80.354718	FL28	E2	US-FL	R Lacy	30.280900	− 84.150520
492	A3	UG-E	M Grodowitz	0.414144	33.207603	FL15	E2	US-FL	TLGraves	29.488817	− 82.975378
493	A3	UG-E	M Grodowitz	0.414144	33.207603	FL40	E2	US-FL	M Sowinski	27.437270	− 82.483820
494	A3	UG-E	M Grodowitz	0.414144	33.207603	FL2	E2	US-FL	J Sowards	29.098590	− 82.436420
Arg12	B0	AR-B	W Cabrera-Walsh	− 34.220756	− 58.895720	FL65	E2	US-FL	B Nelson	27.439667	− 82.484667

Table 1 (continued)

Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)	Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)
Arg17	B1	AR-W	W Cabrera-Walsh	−28.130230	−58.997780	FL26	E2	US-FL	J Evans	28.953637	−82.233031
Arg15	B1	AR-E	W Cabrera-Walsh	−30.746530	−58.017780	FL27	E2	US-FL	J Evans	28.958814	−82.231531
Arg11	B1	AR-B	W Cabrera-Walsh	−34.220756	−58.895720	FL55	E3	US-FL	D Fussell	30.225333	−85.882588
Arg14	B1	AR-G	W Cabrera-Walsh	−27.479720	−64.828610	BA10b	E3	US-FL	FA Dray, P Madeira	30.281507	−84.150778
Arg18	B1	AR-G	W Cabrera-Walsh	−27.479720	−64.828610	BA02b	E3	US-FL	FA Dray, P Madeira	29.216027	−82.048760
Arg19	B1	AR-G	W Cabrera-Walsh	−27.479720	−64.828610	FL68	E3	US-FL	E Egensteiner	27.940083	−81.345017
CA1	B2	US-CA	P Barbara, SQ Raymond	32.765250	−117.168780	FL69	E3	US-FL	E Egensteiner	27.943267	−81.356583
CA2	B2	US-CA	P Barbara, SQ Raymond	32.770275	−117.154362	**BA11b**	E3	US-FL	FA Dray, P Madeira	30.248799	−84.148555
LA5	B3	US-LA	R Diaz	30.408812	−91.163080	PR3	E3	US-PR	J Crossland	18.390281	−66.055436
BA01b	B4	US-FL	FA Dray, P Madeira	29.605174	−82.302775	FL19	E4	US-FL	K Aliengena	29.212780	−82.053250
FL13	B4	US-FL	M McLaughlin, M Perez	29.604269	−82.301969	**FL20**	E4	US-FL	K Aliengena	29.216331	−82.049241
FL29	B4	US-FL	R Lacy	30.432440	−84.542830	Cu3	E5	CU-15	D Giardina	22.793361	−83.008254
BA18a	B5	US-FL	FA Dray, S Goldstein	27.503656	−81.805925	BA03b	E5	US-FL	FA Dray, P Madeira	29.084876	−81.577654
FL41	B5	US-FL	M Sowinski	27.503650	−81.804760	**FL31**	E5	US-FL	S Simmons	29.080590	−81.578470
BA08b	B5	US-FL	FA Dray, P Madeira	30.466066	−84.488319	**BA04b**	E5	US-FL	FA Dray, P Madeira	29.214082	−81.654982
FL23	B5	US-FL	D Fussell	30.465780	−84.488420	FL35	E5	US-FL	J Evans	29.213872	−81.655236
LA1	B5	US-LA	S Johnson	29.920050	−90.452650	Cu2	E6	CU-15	D Giardina	22.793361	−83.008254
LA3	B5	US-LA	M Ferro	30.422200	−91.169100	**CH1**	F1	CN-HB	Y Wang	31.108583	115.714194
LA4	B5	US-LA	M Ferro	29.920050	−90.452650	**ThA**	F1	TH-10	M Rayamahji	13.852150	100.574983
LA6	B5	US-LA	R Diaz	29.562086	−90.792820	ThD	F2	TH-83	T Center, A, Wright	9.149917	99.672650

Table 1 (continued)

Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)	Sample ID	TCS haplo-type	Country/ region code	Collector(s)	Latitude (°N)	Longitude (°W)
LA7	B6	US-LA	R Diaz	29.375211	− 90.726607	**ThB**	F2	TH-80	M Rayamahji	8.524133	99.950000
3	B7	AU-QLD	A Wright	− 26.576517	152.923151	ThE	F2	TH-81	T Center, A, Wright	8.518800	99.958083
4	B7	AU-QLD	B Brown	− 27.511667	152.996833	**Arg6**	G1	AR-B	M Rayamahji	− 26.435260	− 58.276990
Arg16	C1	AR-N	W Cabrera-Walsh	− 27.420820	− 55.900070	Br1A	G1	BR-SP	G Wheeler	− 22.680660	− 48.253120
Br3	C1	BR-PR	P Tipping	− 25.557870	− 49.226100	Br1B	G1	BR-SP	G Wheeler	− 22.680660	− 48.253120
Br4	C1	BR-PR	P Tipping	− 25.532650	− 49.225550	491	G1	GY-EB	M Grodowitz	6.212853	− 57.390367
Br8	C1	BR-SC	G Wheeler	− 27.449578	− 48.501644	490	G1	GY-EB	M Grodowitz	6.090540	− 57.177990
Br2	C1	BR-SC	P Tipping	− 27.028848	− 49.093395	**488**	G1	GY-EB	M Grodowitz	5.823100	− 57.163010
CH7	C1	CN-GX	G Wheeler	24.996790	110.451480	Equation 1	G2	EC-U	T Center	− 0.473662	− 76.460188
CH8	C1	CN-GX	G Wheeler	24.987860	110.442820	**Equation 2**	G2	EC-U	T Center	− 0.473662	− 76.460188
CH9	C1	CN-GX	G Wheeler	24.975940	110.447260	**489**	G3	GY-EB	M Grodowitz	5.894760	− 57.143400

Country codes (ISO 3166–2): Argentina—AR, Australia—AU, Brazil—BR, Cuba—CU, China—CN, Ecuador—EC, France—FR, Guyana—GY, Mexico—MX, Nicaragua—NI; South Africa—ZA, Thailand—TH, Uganda—UG, United States of America—US

Country and region (ISO 3166–2): code locations available using search engines: enter "code" and the actual ISO 3166 code

Bold letters indicate the sample was used in the Pistia6seq analysis

of Taberlet et al. (1991); for *rpl20-rps12*- primers rpl20 and rps12 of Hamilton (1999); for *nad1*- primers 1b and 1c of Demesure et al. (1995); for *rpl32-trnL*- primers trnL (UAG) and rpL32-F of Shaw et al. (2007) with additional primers (ours)- *Pistia* R, external (AGC TCA TTC ATC TTG AAT CGT CGA G) and *Pistia* F, internal (CGA GAT AAT AAT TTT GAC TTA CGA CA CTC); *matK*- our primers- *Pistia* MatK F'' (CAT TGC GAT TAT CTT CCC TCG AAG) and ''*Pistia* MatK R'' (CTT ACT AAT GGG ATT CCC CGA TAC GT). Annealing temperatures were for: *matK*-60 °C, *trnL intron*-63 °C, *trnLF spacer*-63 °C, *rpl20-rps12 intergenic spacer*-53 °C, *nad1*-57.5 °C, *rpL32-trnL*-62 °C.

Amplification products were electrophoresed on 1.6% agarose gels and visualized with ethidium bromide. PCR products were usually purified with the Zymogen DNA Clean and Concentrator™-5 Kit (Zymo Research, Irvine, CA). Occasionally the PCR product was excised from the gel and cleaned with the Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Recovered PCR product was quantitated (Qubit fluorimeter, Life Technologies, Carlsbad, CA), adjusted to 20 µg/µl, and used as template for cycle sequencing performed by Eurofins MWG Operon (Huntsville, AL, USA) using BigDye™ terminator technology (Life Technologies Corp., Carlsbad, CA, USA). Sequencing was performed in both directions using the PCR primers and, in the *rpL32-trnL* case, when needed, with the secondary *Pistia* primers.

Sequences were edited using Sequencher v5.0 (Gene Codes USA). The *matK*, *rpl20-rsp12*, and *nad1* sequences used for phylogenetic analysis aligned unambiguously using MAFFT (Katoh et al. 2002). The *trnL intron*, *trnLF spacer*, and *rpL32-trnL* sequences for phylogeny were aligned with MCoffee (Wallace et al. 2006) and evaluated using the Transitive Consistency Score (TCS), which identifies the most correct positions in an MSA then strips the most ambiguous parts (Chang et al. 2014). Finally, some remaining gap-containing regions were removed manually to further reduce missing data. For within-*Pistia* species comparisons, both the four sequences used for the population analysis and the six sequences for the internal phylogeny aligned unambiguosly using the MAFFT aligner (Katoh et al. 2002).

Haplotype network analysis and mapping

The TCS 1.21 program (Templeton et al. 1992) (Clement et al. 2000), not to be confused with the Transitive Consistency Score (TCS) used in the alignment, was used to generate a haplotype network. *Pistia* haplotypes were analyzed using three concatenated chloroplast sequences: *matK* (800 bases), *trnLF* (1048 bases containing the intron, *trnL*, and spacer, with gaps removed), and *trnL-rpL32* (1032 bases). Gaps were also coded (as bases) to extract additional information. Different but overlapping gaps were coded differently but as the same character. The *trnLF* sequence produced 3 gap characters while *rpL32-trnL* produced 12. The analysis in TCS was run using "Gaps = missing" so that the gap changes would be considered as a single change without order. During an early analysis, after comparing actual distances with TCS distances, a link through the B7 haplotype (Australia) appeared to distort distance (actual vs TCS) relationships. After removing B7 and rerunning TCS "minus Australia" the distortion disappeared. The B7 loop was then manually re-added. Haplotype locations were then geolocated onto maps generated using ArcGIS.

Phylogenetic analysis

Phylogenetic analysis was carried out using the edge-linked partition model of maximum likelihood in IQ-TREE (Chernomor et al. 2016; Nguyen et al. 2015; Trifinopoulos et al. 2016) which allows the modeling of each partition. For concatenation-based tree inference, Zhou et al. (2018) showed IQ-TREE consistently achieved the best-observed likelihoods among IQ-TREE, RAxML/ExaML, PhyML, and FastTree. The model for each sequence within IQ-TREE was chosen based on BIC (allowing "free rate heterogeneity"). A partition file was created for both a 4 sequence (4seq) and a 6 sequence '*Pistia* only' analysis (Pistia6seq). The '4seq' analysis (*matK*, *trnL intron*, *trnLF spacer*, *rpl20-rsp12*) compares interspecific variation from NCBI outgroup sequences versus the haplotype variation within *Pistia*. The 'Pistia6seq' analysis adds *nad1* and *rpL32-trnL* to the alignment while also aligning unambiguously, adding characters and including gap coding.

Sequence criteria for the 4seq data (alignment length: # informative characters) returned for *matK*

(800:71), *trnL intron* (470:32), *trnLF spacer* (410:30), *rpl20-rsp12* (738:44). The Pistia6seq analysis yielded *matK* (800:41), *trnL intron* (470:10), *trnLF spacer* (410:15), *rpl20-rsp12* (738:20), *NAD* (1293:39), *rpL32-trnL* (638:87). Models were chosen using IQ-TREE "model selection" and entered into the partition file. Models for 4seq: *matK*-K3Pu + F + G4, *trnL intron*-F81 + F + R2, *trnLF spacer*-TVM + F + R2, *rpl20-rsp12*-HKY + F + G4, *NAD*-JC + I, *rpL32-trnL*-HKY + F + G4; for Pistia6seq: *matK*-F81 + F + I, *trnL intron*-TN + F + I, *trnLF spacer*-HKY + F, *rpl20-rsp12*-K3Pu + F + I, *NAD*-JC + I, *rpL32-trnL*-F81 + F. Gaps were coded as in the population analysis and included in the partition file as model JC.

Clade support was analyzed using 5000 replicates for aBAYES (Anisimova et al. 2011), ultrafast bootstrap [UFBoot] (Hoang et al. 2018), and the Shimodaira–Hasegawa approximate likelihood ratio test [SH-aLRT) (Guindon et al. 2010). The majority rule (MR) consensus percentage was calculated from 500 trees. Branch support was considered strong for: aBAYES ≥ 0.95; SH-aLRT ≥ 0.85 (stronger ≥ 0.90); UFBoot BS ≥ 95%; MR ≥ 80%.

Finally, as it became apparent that there were cryptic differences in Florida's *Pistia* populations, the need for a simple molecular tool to differentiate among these cryptic genotypes became clear. Natural area managers may want to identify the specific haplotype they are managing (c.f. Saltonstall 2003). Thus, each of the six sequences used in the phylogenetic analysis were processed individually through a Neighbor Joining analysis in MEGAX using complete deletion and the "number of differences" model to see if a single sequence could serve the purpose.

Species delimitation

The definition of 'species' has varied widely. Mayden (1997) names 24 different species concepts and presents even more definitions. De Queiroz (2007), in summarizing the major categories of definitions (e.g., biological, ecological, evolutionary, cohesion, phylogenetic and phenetic), indicated general agreement that a species represents a "separately evolving metapopulational lineage". However, there is discord on which secondary properties best define species separation. A "Unified Species Concept" (de Queiroz 2007) separates issues of 'species conceptualization' from those of 'species delimitation'. Species delimitation, as a secondary criteria, is relevant in proportion to the number and strength of its evidence for lineage separation.

Ross et al. (2008) reviewed molecular tools used for species delimitation including BLAST, distance, and tree-based methods. The predictor with the best identification success was the 'average within-species genetic distances' divided by the 'average between-species genetic distances' (Intra/Inter ratio) presented under both strict and relaxed cladistic criteria [P ID (strict) or P ID (liberal)]. The conservative strict criterion which requires the query to fall 'within' a monospecific clade was not useful for our analysis where small clades occur. The liberal method requires only that the query forms a clade 'with' a monospecific group, either sister to or within the reference clade. If all species are represented within the reference data the liberal method will make more correct identifications, however also more false positive ones.

Masters et al. (2011) have designed useful tools (Kiewnick et al. 2014; Gutiérrez et al. 2017; Xu et al. 2017) to explore species boundaries in a 'Species Delimitation' plugin for Geneious software. The user assigns taxa to clades and the plugin computes probability statistics. Output metrics here include the Intra and Inter distances, the Intra/Inter ratio, the P ID (liberal)with its 95% confidence. It additionally calculates Rosenberg's PAB statistic (Rosenberg 2007), the probability of reciprocal monophyly. These analyses were conducted using a Nexus tree generated from the 'Pistia6seq' dataset (without haplotype B7, Australia samples) as run and modeled in IQ-TREE.

Results

Sample acquisition

Table 1 presents the collection data for 155 samples collected worldwide from fourteen countries. Ninety-four samples were from the USA (excluding Puerto Rico), eleven from Mexico, Central America and the Caribbean, twenty-four from South America, nine from Africa, two from Europe, thirteen from Asia and two from Australia. Figures 1 and 2 show their distribution.

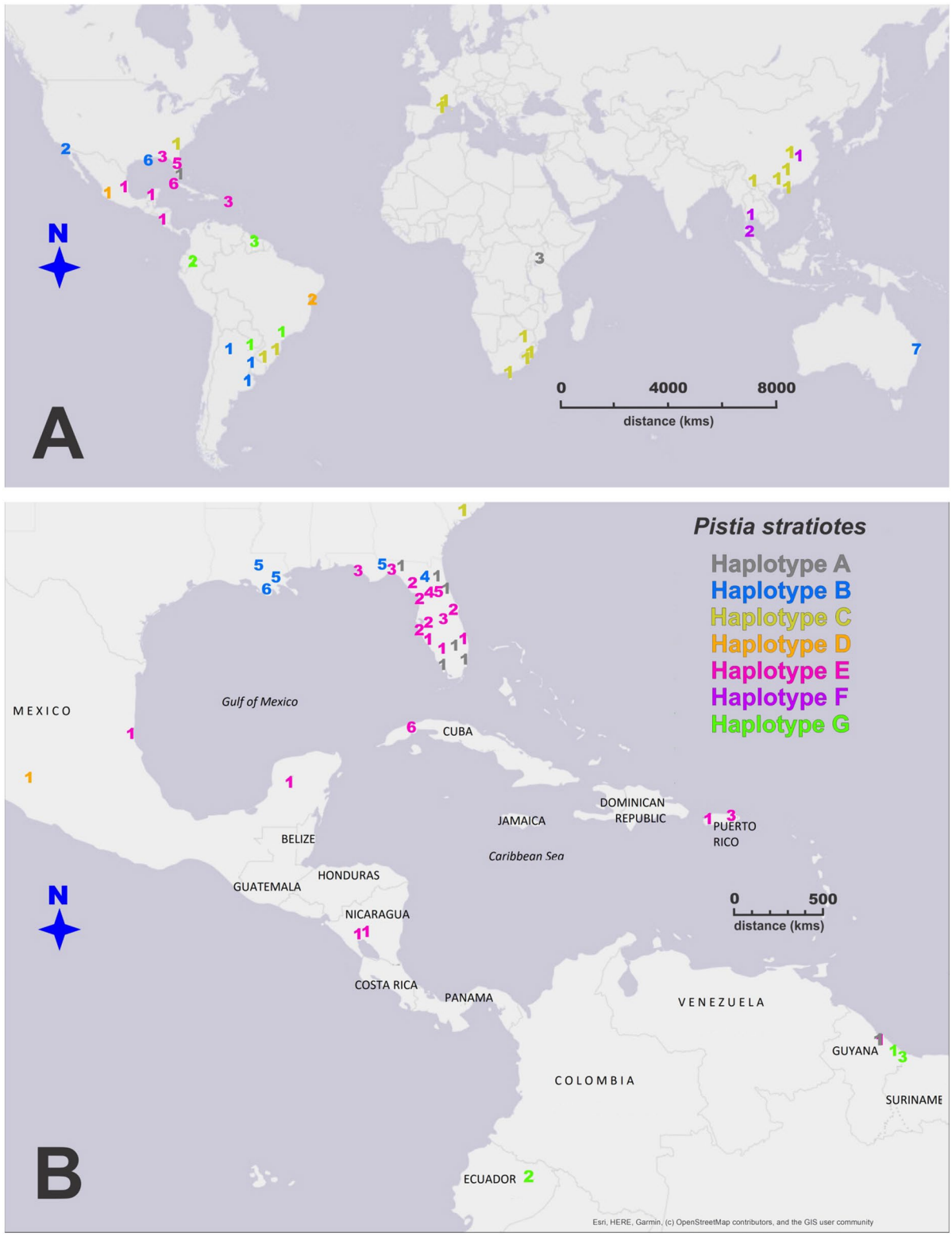

Fig. 1 Map of *Pistia stratiotes* haplotypes worldwide (**a**) and in the Caribbean (**b**). Clades are represented by different colors, with haplotypes within a clade separated numerically following Table 1

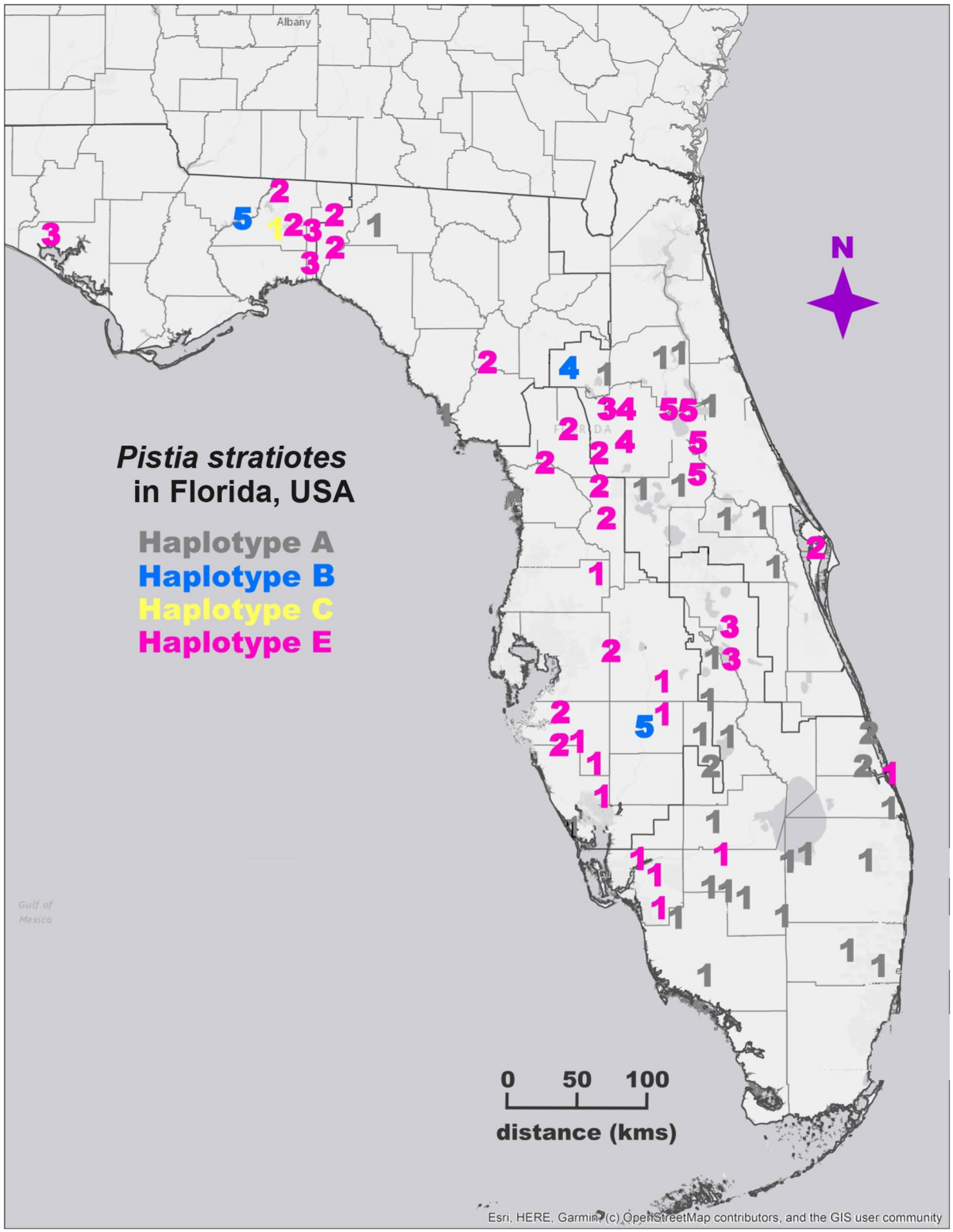

Fig. 2 Map of *Pistia stratiotes* haplotypes in Florida. Clades are represented by different colors, with haplotypes within a clade separated numerically following Table 1

Haplotype network analysis and mapping

Figure 3 displays the results of the TCS network construction. Oval colors represent (assigned) groupings of similar haplotypes separated from other groups by larger numbers of differences. An exception to these delineations is groups **A** and **B** (hereafter **bold** letters represent clades, whereas individual haplotypes remain plain text) which, as haplotypes were added, clustered together (**A/B**).

Information on individual samples (Table 1, Supplemental Table 2) may be visualized in the maps of Figs. 1 and 2.

Haplotype A1 appears in Florida (37 times: 37X) and Guyana (2X). Haplotype A2 is in Florida (5X) and haplotype A3 in Lake Victoria, Uganda (3X). Four **B** haplotypes (B3-6) were primarily found along the US Gulf Coast States [Louisiana (1X); Florida (7X) & Louisiana (6X)]. The **B** haplotypes also appeared in California (haplotype B2). The Haplotype B1, basal in the **B** clade, indicating it diverged early, was found in six samples from Argentina. Haplotype B0, also from Argentina, is basal to both the **B** and **A** clades. It is unique in displaying the **G** type (rather than an **A/B** type) matK sequence. Haplotype B7 from Australia is quite different and intermediate between the **B** and **E** haplotypes in the TCS analysis (Fig. 3).

The **C** group consists of a single haplotype, C1, found in 22 samples and 6 countries, predominantly in Eurasia. There are two **F** group haplotypes, F1 (China, 1X; Thailand, 1X) and F2 (Thailand, 3X). There are three **G** group haplotypes, all located within South America.The **D** group has two haplotypes, found widely separated geographically, in Mexico and Brazil.

Six haplotypes were found within the **E** clade. Haplotype E1 is widespread around the Caribbean [Florida (12X sites), Mexico (2X, 1X in the Yucatan), Nicaragua (2X), Puerto Rico (2X), Cuba (1X)]. Haplotypes E2 (14X), E3 (6X), E4 (2X), and E5 (3X) appeared in Florida. E3 also appeared in Puerto Rico (1X). E5 (1X) and E6 (1X) were found in Cuba.

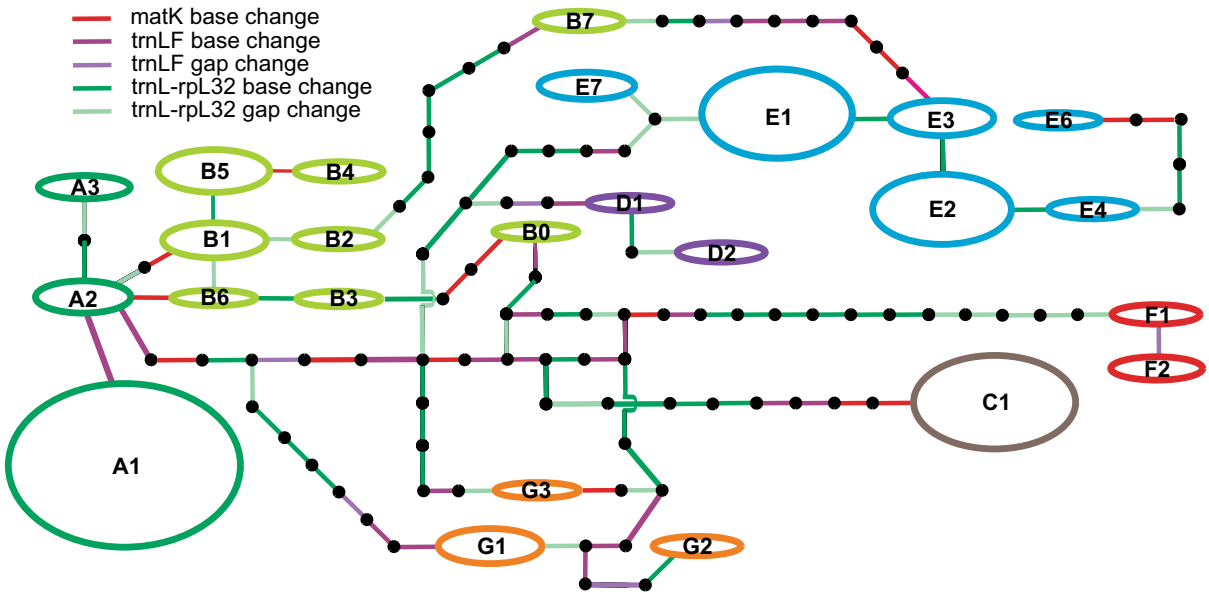

Fig. 3 TCS network construction for Florida, Caribbean and world-wide samples. Oval symbols represent each haplotype (25 total) for the concatenated matK, trnLF and rpL32-trn sequences, including gaps. The haplotypes may also be cross-referenced with the corresponding NCBI accession number in Supplemental Table 1. Each oval is sized in proportion to the number of samples obtained of that type. The connections between each pair of ovals, between dots, or between ovals and dots represent either a single base change or a gap change. The changes are represented by color coded lines. The ovals representing haplotypes are color coded in seven colors representing (assigned) groupings of similar haplotypes, usually separated from other groups by larger numbers of differences. Information on individual sample IDs, haplotype and sample location may be accessed in Table 1 and visualized in the maps of Figs. 1 and 2

Phylogenetic analysis and speciation

The phylogenetic analysis is presented using maximum likelihood trees generated by IQ-TREE (Fig. 4). Figure 4a represents the 4 sequence (4seq) phylogenetic analysis. Despite different combinations of sequences and final alignments, the output tracks closely that of both Nauheimer et al. (2012; using RAxML) and Renner and Zhang (2004; using MrBayes). Figure 4a allows the comparison of patristic distances across related Araceae species with the variation within our *Pistia* clade. Note that from the crown of the *Pistia* clade the distance axis has been expanded to allow better visualization (see distance bars). Within the Alocasias, the smallest patristic distance between is *Arisaema cucullata* and *A. gageana* at 0.00084, followed by *A. triphyllum* and *A. amurense* at 0.00643. Looking

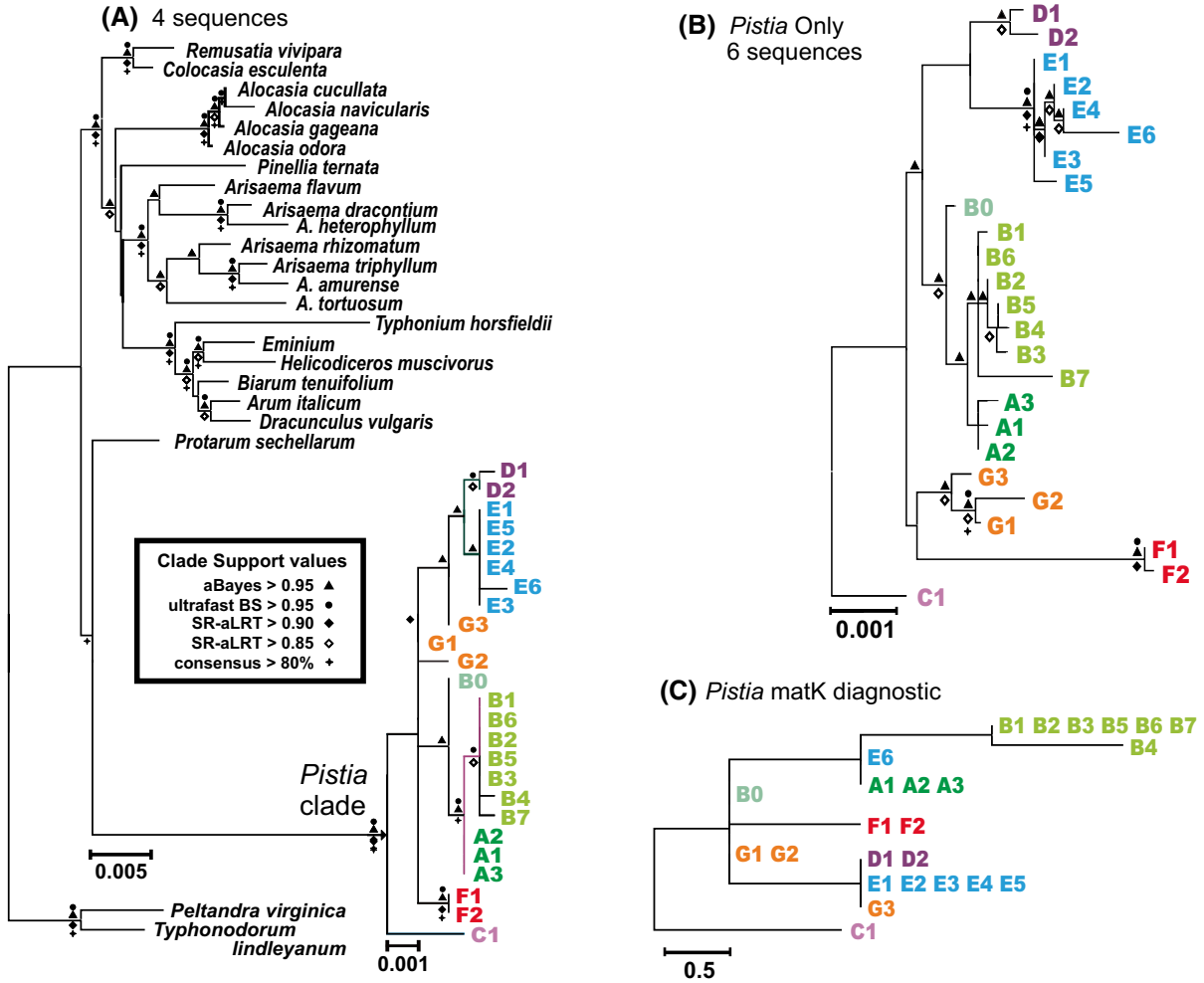

Fig. 4 Phylogenetic analyses of Pistia samples. **a** IQ-TREE Maximum Likelihood analysis of 4 concatenated sequences (4seq) including matK, trnL intron, trnLF spacer, and rpl20-rsp12, with gaps coded. It is presented to allow comparison of the variation across related Araceae species with variation in the Pistia clade. Note that from the crown of the Pistia clade the distance axis has been expanded to allow better visualization (see distance bars). Patristic distances are the sum of the horizontal lengths between OTUs. The **b** IQ-TREE Maximum Likelihood analysis of 6 concatenated sequences (Pistia6seq) adding nad1 and rpL32-trnL to the alignment. **c** Pistia "clade" diagnostic using a Neighbor Joining Tree of matK sequences offers the best single sequence diagnostic for haplotype clades in Florida. Use with great caution outside of Florida where unique haplotypes not found here may occur, noting the exceptions to classification from other regions. Please see supplementary Table 1 for the corresponding NCBI accessions

within the *Pistia* groups, the distance between G1 and C1 is 0.00299, between G1 and F is 0.00086, and between D2 and most E's (E1-5) is 0.00086. All of these are greater than the interspecies distance between *A. cucullata* and *A. gageana*, but less than all other interspecies distances in the *Pistia* "outgroup" clade. The smallest inter-clade distance is between the **A** and **B** groups at 0.00044, suggesting they separated evolutionarily relatively recently. However, **A** and **B** groups do form distinct sub-clades.

Even though C1 appears as an outgroup when examining the 4seq phylogeny, there is limited branch support—especially near the root. The *Pistia* 6seq (Fig. 4b) phylogeny strengthens support values for **C** (as the outgroup) as well as for other clades and serves as the template for our species delimitation analysis. Low support values for the **B** clade reflect the instability of B7 placement in this analysis, as in the TCS analysis, with repeated IQ-TREE runs (not presented) showing B7 switching positions within the clade.

Pistia "clade" diagnostic

Of the six 'single sequence' Neighbor Joining Trees examined, *the matK sequence stood out as the best diagnostic* (Fig. 4c) *for distinguishing the haplotype clades* ***in Florida***. The diagnostic should be used with caution outside of Florida where unique haplotypes likely occur. Two criteria make matK a useful diagnostic. First, as a coding gene, there were no gaps/indels to consider in alignment. Second, the sequence amplifies robustly. The matK sequence successfully created a clade distinguishing all the **B** haplotypes. It also distinguished between the **A** and **B** clades despite their genetic similarity. The **A** clade presented as a single haplotype for matK, though it also included haplotype E6. However, E6 was only found in a single sample from Cuba, so it is unlikely to confuse an analysis of Florida samples. The third major clade in Florida, containing the **E** haplotypes, appeared with an identical sequence with the **D** clade and with the G3 haplotype. However, there is no evidence those haplotypes are in Florida with the analyzed sample G3 from Guyana and the **D** haplotypes from Mexico and Brazil.

Species delimitation

Results of the species delimitation tests are presented in Table 2. Rosenberg's PAB determined that all tested groups were statistically monophyletic, expected since haplotypes were defined in monophyletic groups. We examined the Intra/Inter ratios (Ross et al. 2008; Masters et al. 2011) using relaxed (Prob ID, Liberal) criteria. This test was designed to compare the "closest" sister clades, but comparisons between clades that are not "closest" are also informative and thus here presented. Table 2a presents these statistics. Additionally, the comparison probabilities were sorted (Table 2b) from lowest significance to greatest. Because liberal probabilities are unidirectional they are influenced by the number of members in the alternative clade. Those clades which contain low numbers and singletons generally display lower significance as the alternate clade so the bi-directional calculations should be considered. A Prob ID, Liberal ≥ 0.93 is used here as the cutoff for significance (Hamilton et al. 2014). Comparisons are provided with and without the singleton B0 (Table 2), which lies close to the root (Fig. 4a). When clades **A** and **B** are combined, B0 is included in the third analysis (**AB***) and omitted in the second analysis (**AB**).

*The Pistia clade **C** is phylogenetically the most distinct clade.* Despite both being singletons, C1 and B0 were bi-directionally different at ≥ 0.95. The **G** clade was not distinguishable from **C** (**G** distinguishable from **C**: **G**$\succ$**C**), a singleton, with a probability at 0.85, however **C** was distinguishable from **G**: (**C**$\succ$**G**) at 0.96 (G has 3 haplotypes). Clade **C** was reciprocally distinguishable from the next nearest clade as **AB** (excluding B0) or **AB***(including B0). These results suggest **C** may be a different species.

The next most isolated clade, **F**—found in SE Asia, was also reciprocally distinguishable from B0, from **AB** and from **AB***. **G**$\succ$**F** was not significant, however **F**$\succ$**G** was significant. Again, and for similar reasons to **C**, *this evidence suggests that **F** is a separate species.*

Given that the **E**, **A** and **B** clades are in Florida, it is important for weed control practitioners to understand whether these represent different species. In many comparisons (**E**$\succ$**A**, **A**$\succ$**E**, **E**$\succ$**B**, **B**$\succ$**E**, **E**$\succ$**AB**, **AB**$\succ$**E**, **E**$\succ$**AB***, **AB***$\succ$**E**, B0$\succ$**E**, **E**$\succ$B0) the clades were distinguishable. The clear

Table 2 Species delimitation analysis

A. Results of Geneious delimitation analysis

Clade	Alternative Clade	Intra Clade distances[1]	Inter Clade distances[2]	Ratio of Intra/ Inter distances[3]	Prob ID (Liberal)[4,5] (R^2 95% limits)	Rosenberg's Probability[6]
initial analysis - B0 as distinct clade from other Bs						
C	B0	0.00E+00	3.00E–03	0.00	0.96 (0.83, 1.00)	3.33E–03
B0	C	0.00E+00	3.00E–03	0.00	0.96 (0.83, 1.00)	2.00E–02
C	G	0.00E+00	3.00E–03	0.00	0.96 (0.83, 1.00)	3.33E–03
G	C	9.34E–04	4.00E–03	0.28	0.85 (0.70, 0.99)	5.00E–02
F	B0	1.23E–04	4.00E–04	0.03	0.96 (0.81, 1.00)	5.00E–02
B0	F	0.00E+00	4.00E–03	0.00	0.96 (0.83, 1.00)	2.00E–02
F	G	1.23E–04	4.00E–03	0.03	0.97 (0.81, 1.00)	5.00E–02
G	F	9.34E–04	4.00E–03	0.28	0.85 (0.70, 0.99)	5.00E–02
B0	A	0.00E+00	6.63E–04	0.00	0.96 (0.83, 1.00)	2.00E–02
A	B0	2.66E–04	6.63E–04	0.40	0.78 (0.63, 0.93)	2.98E–03
B0	B	0.00E+00	7.54E–04	0.00	0.96 (0.83, 1.00)	2.00E–02
B	B0	2.77E–04	7.54E–04	0.37	0.91 (0.81, 1.00)	2.98E–03
A	B	2.66E–04	6.22E–04	0.43	0.76 (0.62, 0.91)	2.98E–03
B	A	2.77E–04	6.22E–04	0.45	0.89 (0.78, 0.99)	2.98E–03
B0	G	0.00E+00	2.00E–03	0.00	0.96 (0.83, 1.00)	2.00E–02
G	B0	9.34E–04	2.00E–03	0.50	0.73 (0.58, 0.88	5.00E–02
B0	D	0.00E+00	2.00E–03	0.00	0.96 (0.83, 1.00)	2.00E–02
D	B0	5.75E–04	2.00E–03	0.30	0.79 (0.64, 0.95)	1.00E–02
E	B0	5.35E–04	2.00E–03	0.23	0.94 (0.87, 1.00)	1.00E–02
B0	E	0.00E+00	2.00E–03	0.00	0.96 (0.83, 1.00)	2.00E–02
E	B	5.35E–04	3.00E–03	0.19	0.95 (0.88, 1.00)	1.00E–02
B	E	2.77E–04	3.00E–03	0.10	0.97 (0.87, 1.00)	2.98E–03
E	A	5.35E–04	3.00E–03	0.19	0.95 (0.88, 1.00)	1.00E–02
A	E	2.66E–04	3.00E–03	0.10	0.95 (0.81, 1.00)	2.98E–03
E	D	5.34E–04	2.00E–03	0.26	0.93 (0.87, 1.00)	1.00E–02
D	E	5.35E–04	2.00E–03	0.28	0.81 (0.66, 0.96)	1.00E–02
2nd analysis - AB (A & B as combined clade but without B0)						
C	AB	0.00E+00	3.00E–03	0.00	0.96 (0.83, 1.00	3.33E–03
AB	C	4.48E–04	3.00E–03	0.14	0.97 (0.92, 1.00)	2.00E–02
F	AB	1.42E–04	5.00E–03	0.03	0.97 (0.82, 1.00	5.00E–02
AB	F	4.48E–04	5.00E–03	0.10	0.98 (0.93, 1.00)	2.00E–02
E	AB	5.34E–04	3.00E–03	0.19	0.95 (0.88, 1.00)	1.00E–02
AB	E	4.48E–04	3.00E–03	0.16	0.97 (0.91, 1.00)	2.00E–02
G	AB	9.34E–04	2.00E–03	0.39	0.79 (0.64, 0.93)	5.00E–02
AB	G	4.48E–04	2.00E–03	0.19	0.96 (0.91, 1.00	2.00E–02
D	AB	5.75E–04	2.00E–03	0.24	0.83 (0.68, 0.99)	1.00E–02
AB	D	4.48E–04	2.00E–03	0.19	0.96 (0.91, 1.00	2.00E–02
3rd analysis - AB* (A & B as combined clade, including B0)						
C	AB*	0.00E+00	3.00E–03	0.00	0.96 (0.83, 1.00	3.33E–03
AB*	C	5.03E–04	3.00E–03	0.16	0.97 (0.91, 1.00)	1.20E–06
F	AB*	1.23E–04	5.00E–03	0.03	0.97 (0.82, 1.00	5.00E–02
AB*	F	5.03E–04	5.00E–03	0.11	0.98 (0.92, 1.00)	1.20E–06
G	AB*	9.34E–04	2.00E–03	0.40	0.78 (0.64, 0.93)	5.00E–02
AB*	G	5.03E–04	2.00E–03	0.22	0.95 (0.90, 1.00)	1.20E–06
D	AB*	5.75E–04	2.00E–03	0.24	0.83 (0.68, 0.98)	1.00E–02
AB*	D	5.03E–04	2.00E–03	0.21	0.95 (0.90, 1.00	1.20E–06

B. Clade relationships sorted by Prob ID

Clade	Alternative Clade	Prob ID (Liberal) mean
< 0.93 [5]		
G	B0	0.73
A	B	0.76
A	B0	0.78
G	AB*	0.78
D	B0	0.79
G	AB	0.79
D	E	0.81
D	AB	0.83
D	AB*	0.83
G	C	0.85
G	F	0.85
B	A	0.89
B	B0	0.91
>= 0.93 [5]		
E	D	0.93
E	B0	0.94
A	E	0.95
E	B	0.95
E	A	0.95
E	AB	0.95
AB*	G	0.95
AB*	D	0.95
B0	E	0.96
F	B0	0.96
C	B0	0.96
B0	C	0.96
C	G	0.96
B0	F	0.96
B0	A	0.96
B0	B	0.96
B0	G	0.96
B0	D	0.96
C	AB	0.96
C	AB*	0.96
AB	G	0.96
AB	D	0.96
F	G	0.97
F	AB	0.97
F	AB*	0.97
B	E	0.97
AB	E	0.97
AB*	C	0.97
AB	C	0.97
AB*	F	0.98
AB	F	0.98

[1] Average within-clade patristic distances

[2] Average between-Sclade patristic distances

[3] (Average within-clade patristic distances)/(Average between–clade patristic distances)

[4] unidirectional probability calculated using Species Delimitation plugin for Geneious, see Masters et al. (2011)

[5] probability above 0.95 (or 0.93, Hamilton et al. 2014) indicate clades are likely separate species

[6] probability of reciprocal monophyly under the null model of random coalescence, see Rosenberg (2007)

differentiation of **AB*** from **E** suggests that ***E** and **AB*** clades may approach speciation*.

The species delimitation analysis is inconclusive for the closest clade comparisons. For example, **E**≻**D** (0.93) but **D**≻**E** (0.81). Similarly, clades **A** and **B**, both in Florida, are distinct phylogenetically but fail the metric used here to separate species (**B**≻**A**, 0.89; **A**≻**B**, 0.76). B0≻**A** and B0≻**B** are both significant whereas **A**≻B0 (singleton) is not, and **B**≻B0 is only marginally significant.

Discussion

Does haplotype divergence among sample sites inform nativity?

Biological invasions present a plethora of management challenges that can be amplified if the invader is a cryptogenic species (Canavan et al. 2020). *Pistia stratiotes* (waterlettuce) *provides an example of a heretofore unrecognized cryptic species, with the potential to complicate local management decisions.* After determining the haplotypes of all samples, we present (using TCS) 25 distinct haplotypes clustered into seven separate haplotype clades. Given the sparsity or absence of samples from many regions of the world (e.g., Egypt, where the plant was historically recorded) it seems reasonable to assume that more haplotypes (and perhaps clades) remain to be discovered.

Figure 2 and Table 1 make it clear that *two groups, **A** and **E**, dominate the Florida population, with the **B** group also present.* The A1 haplotype probably has its origins in South America, with two samples found in Guyana. The A3 haplotype showed up in Uganda, causing us to consider a possible African origin. However, the closely related **B** group is clearly situated in the western hemisphere. B0 and B1 are from Argentina with several samples of B1 taken inland close to the Andes where human mediated introduction is less likely. Additionally, the related B2–B6 were only found in North America. Therefore, A3's Ugandan location is likely anthropogenic. Other introductions of American species to Africa, which often become invasive, are well documented. For example, Henderson and Wilson (2017) list the thirty exotic plants that have shown the greatest range expansion in South Africa, of which twenty originated in the Americas. Two of these, *Sagittaria platyphylla* (Engelm.) J.G. Sm. and *Egeria densa* Planch., are aquatic. Focusing on Uganda, Witt et al. (2018) present seventeen plants of "greatest impact", of which fifteen have American origins and one, *Eichhornia crassipes* (Mart.) Solms is aquatic.

The distribution of A1 and A2 in Florida is primarily either in southern Florida or in the St. John's watershed, corresponding with where early Spanish trade ships could have unloaded it with their ballast. Based on the available samples, the **A** haplotypes are probably not native in Florida. Haplotype B0, found in Argentina, might better have been designated as haplotype AB1, as it is "ancestral" (close to the *Pistia* crown) to both clades. Haplotype B1, also found in Argentina, is basal to the **B** clade. Other **B** haplotypes were found in Florida, Louisiana and California, indicating a possibility that some **B** haplotypes are native to North America. Broader sampling of the **B** haplotype might provide clarity. Nevertheless, the number of **B** haplotypes in North America, relative to Argentina, is greater than would be expected (i.e., an introduced type generally shows less phylogenetic variation than the native range).

*The **E** clade, present in Florida, the Yucatan, Central America, Cuba and Puerto Rico, appears to be a Caribbean clade* (Fig. 1b). Examining Fig. 2, the **E** clade dominates in the spring fed rivers of north central Florida emptying towards the Gulf as well as upstream in the Ocklawaha River system (a tributary of the St Johns River) where it has persisted despite the dominance of the introduced A1 haplotype downstream. The genetic diversity of the **E** haplotypes in Florida, along with its occurrence primarily upstream in spring-fed systems, suggests a 'native' status.

This analysis provides an answer to our null hypothesis (A): *P. stratiotes is non-native to Florida. There is at least one haplotypic clade, **E**, which is native. Clade **A** is likely not native. There is a third haplotypic clade, **B**, for which nativity is debatable.*

Do phylogenetics and the 'species delimitation' plugin inform speciation?

The shortest patristic distance between G1 and C1 (Fig. 4a) is 0.00299, clearly greater than the distance between *Arisaema cucullata* and *A. gageana*, and by this comparison sufficient to call C1 a different species. We have seen previously, apart from the **G**

clade, species delimitation analysis indicates C1 is a separate species, despite the lack of power from the C1 singleton. Given these comparisons, and the geographic separation between 'Old World' clade C and the 'New World' clades, C1 probably represents a separate species.

Similarly, the distance between G1 and the **F** clade (red path, Fig. 4a) is 0.00066, in the range of the inter-species distance between *A. cucullate* and *A. gageana*. Species delimitation also indicated potential speciation for the **F** clade, although again the separation from the basal **G** clade is unclear. Given potentially different evolutionary rates within *Pistia* as compared to outgroup species, and that this comparison is only to the 'smallest' Alocasia inter-species distance, it is debatable whether the **F** clade represents a separate species. However, the geographic separation of this clade from American clades still suggests probable speciation. This conclusion is further supported by the patristic distance displayed in the 6 sequence phylogeny (Fig. 4a), which shows **F** at a much greater distance from the other *Pistia* clades than was displayed in the 4 sequence phylogeny (Fig. 4b).

The largest distance between adjacent clades in the Americas is between D2 and most **E**'s (green path, Fig. 4a) at 0.00086, again greater than the inter-species distance between *A. cucullata* and *A. gageana*. However, in the 'species delimitation' analysis only one of the closest clade comparisons, **E≻D** (0.93), was significant. The other, **D≻E** (0.81), was not. Similarly, **AB*≻D** (0.95) is significant while **D≻AB***(0.83) falls short of significance. So, while there is evidence of potential speciation between the AB* and E clades, the presence of **D** in Mexico and Brazil and **E** in Central America suggests avenues of gene flow may still exist through these intermediate clades. We would conclude that the American clades might best be conservatively considered as different sub-species. It will take a population diversity analysis with more extensive sampling to further inform the speciation question within the Americas.

Our null hypothesis, B, posited that there is only one extant species of P. stratiotes worldwide. The species delimitation evidence here, using both phylogenetics and the Prob ID (Liberal) tool, in addition to the pre-anthropogenic geographic separation, suggest an alternate result, that there are likely three different species of P. stratiotes worldwide.

An overview of the worldwide haplotype distribution of Pistia

While the number of samples clearly has an impact on perceived genetic variation, *it is evident that there has been a tremendous radiation of American clades and haplotypes*, encompassing the **A**, **B**, **D**, **E** and **G** clades. In contrast, the 15 samples of C1 which appear in Old World samples show no genetic variation. C1 does also appear in limited numbers in Argentina, Brazil, and the USA, creating potential disagreement on its native origin. However, *C1 is widely distributed and predominant in the Old World*, whereas the C1 samples in the Americas are all found in coastal areas indicating probable introduction. Also, C1 is phylogenetically very divergent from the American clades, suggesting a different geographic origin. The CABI Invasive Species Compendium [https://www.cabi.org/isc/datasheet/41496], henceforth CABI ISC, lists *Pistia*'s presence in North Africa, Mediterranian Europe, as well as some non-Mediterranian countries. Additionally, the lack of samples from all tropical Asia west of China and Thailand, where CABI ISC lists broad presence, severely limits any analysis of the distribution of 'Old World' *Pistia* clades, confounding determination of whether the C1 haplotype is native to China.

SE Asia and Oceania present another lack of geographic samples. On the northwestern edge of this region (China and Thailand samples) the **F** clade, genetically very divergent from any other clades, appears native. We wonder why the **F** clade is genetically so distant from the 'Old World' **C** clade while closer to 'New World' clades? According to CABI ISC, *Pistia* is broadly distributed in SE Asia. To the west/southwest of the mainland *Pistia* is found in Japan, Philippines, Singapore, Taiwan, Indonesia and Brunei. It is reasonable that the **F** clade might be represented in parts of this range and present additional haplotype diversity.

Recall that B7 created distortion in the TCS analysis and also low phylogenetic support values for the **B** clade, to which it may have been drawn because of both its uniqueness and 'long branch' attraction. Blake (1954) *suggested that Pistia was introduced into Australia. However, later analysis by* Gillet et al. (1988) *indicated it is probably indigenous in the Northern Territory. The haplotypic uniqueness*

of B7 supports the contention of its native status in Australia.

Between the **F** clade in Thailand and B7 in Australia lies Oceania, where many landmasses also host *Pistia*. While CABI ISC lists *Pistia* as 'introduced' in most of Oceania it is considered native in the isolated Solomon Islands. We suggest caution should be exercised on the 'introduced' designation for much of this range until haplotypic divergence is analyzed or unless historical evidence is clear.

Implications of multiple haplotypes, both alien and native, for management

Canavan et al. (2020, see also references therein) note problems which occur from invasions by a cryptic species. Our data suggest that the occurrence of non-native *Pistia* in Florida (except haplotype **C**) should conservatively be considered as an interspecific cryptic invasion. However, Canavan et al. (2020, following Morais and Reichard 2018) define intraspecific cryptic invasions as those where a non-native lineage is introduced to a region where the species is considered native. Our analysis indicates that some *Pistia* haplotypes are native in Florida. *Thus, this research presents an unusual case where a non-native lineage was introduced into a system containing a cryptic native which was being managed as an invasive weed.*

Pistia has been managed for decades in Florida as a Class I invasive non-native species for its deleterious impacts (USACE 1977; FDEP 2007; FWCC 2017). These can probably be attributed primarily to the non-native **A** haplotypes which predominate in Southeastern Florida and the Saint John's waterway. The native **E** haplotypes, often found in spring-fed waters of Florida, are only occasionally aggressive. This also comports with Evans (2013) contention that *Pistia* survived the cooler temperatures of the most recent glacial maximum (~18,000 BP) in thermally regulated spring systems on the Florida peninsula. It should be noted, however, that Quillen (2013, citing Watts 1969, 1975) states "most lakes in Florida did not begin Holocene sediment accumulation until after ~ 8000 YBP" which argues that the proposed paleo-refugia would have been exceedingly rare, suggesting the possibility of later natural reinvasion from Meso-America or the Caribbean.

Although management of *Pistia* (where problematic) will likely continue, it is possible that the nativity of some populations will prompt natural resource managers to forgo herbicide treatments in some waterways. One suggestion for future research would be for land managers to overlay on a map, areas where herbicides are used to treat *Pistia* with areas containing the invasive **A** haplotypes. The overlap may show that populations being treated are primarily non-native. Further, *the matK sequence can be used as a tool with which land managers can check individual populations to determine nativity before deciding on the type of management.*

This study utilized mostly chloroplast (1 mitochondrial) sequences for the determination of genetic variation. Since these sequences are maternally inherited, they offer no information on the potential hybridization of divergent haplotypes, and especially here among **A**, **B** and **E** haplotypes. One would presume, given their close relationship, this would most likely occur between **A** and **B**. Ellstrand and Schierenbeck (2006) have proposed that hybridization may act as a stimulus for increased success in non-native populations, citing numerous cases. Clearly, a better understanding of *Pistia*'s reproductive biology and potential for hybridization is needed. Hybridization could be tested through laboratory interbreeding of the haplotypes. It also could be examined through a population genetics study using microsatellites or genome-wide tools such as RAD or GBS sequencing.

Implications of multiple haplotypes, both alien and native, for biological control

In their review, Canavan et al. (2020) *state that "to date there have been no biological control agents released on a cryptic species". Pistia stratiotes provides a newly recognized exception to this observation.* Because of its classification and behavior as a non-native invasive weed, two insect biological control agents (as mentioned earlier) have been released in the USA: the South American weevil *N. affinis* and the Asian moth *S. pectinicornis* Hampson (Dray et al. 1990, 2001).

It is generally agreed that finding the origin of a non-native invasive increases the chances of finding co-evolved, efficacious agents (Smith et al. 2017, and literature therein). Perhaps *S. pectinicornis* failed to persist in Florida (Dray et al. 2001) because it was collected from *Pistia* haplotypes in Thailand and was thus poorly adapted to succeed on Florida haplotypes.

In contrast, *N. affinis* (the South American weevil) was collected on haplotype(s) genetically closer to Florida **A** and **B** haplotypes, perhaps explaining why the weevil was moderately successful (Dray et al. 1990). The success of these agents on the **E** haplotype remains unknown.

A third agent, *Lepidelphax pistiae* Remes Lenicov (Hemiptera: Delphacidae), has been extensively tested both in its native range (Cabrera Walsh et al. 2014) and at the USDA ARS Invasive Plant Research Laboratory (Goode et al. 2019). Combined, these host range studies showed that *L. pistiae* was specific enough to petition for release. Additionally, quarantine studies showed it was quite damaging at medium and high densities (Goode et al. 2019). However, these studies overlapped with Quillen's (2013) report of fossil seeds in Lake Annie, Florida, with its implications for potential *Pistia* nativity. This caused reflection by USDA and other scientists about the advisability of conducting biological control activities against *Pistia*.

Many biological control practitioners would argue that biocontrol agents released on intraspecific cryptic species should be specialists at the subspecific level. While some herbivores do show specificity below the species level, Canavan et al. (2020) point out that potential rapid adaption and/or hybridization increase the danger of host shifts over time in such systems. A similar situation to *Pistia* is found in biological control efforts against European *Phragmites australis* (Cav.) Trin. ex Steud (Poaceae), a cryptic invader in the US, where it out-competes the native *P. australis* ssp. *americanus* (Saltonstall 2002). Testing showed native lineages were within the fundamental host range of two European moths (*Archanara* spp.) but were outside the native's ecological host range, with much lower impacts on the native (Blossey et al. 2018). A vigorous debate ensued, reflecting both the potential dangers and benefits of releasing agents against weeds at the subspecies level (Kiviat et al. 2019; Blossey et al., 2020). It is important to note that current management methods have been unable to blunt the invasiveness of *P. australis australis* and its threats to rare and endangered native species.

*In conclusion, the weight of evidence points to the presence of native (**E** and possibly **B** haplotypes) and non-native (**A** and **C** haplotypes) P. stratiotes in Florida.* Thus, the hesitancy of biological control practitioners to release foreign control agents on native plant species prompted a pause in efforts to obtain release permits for *L. pistiae*. Studies are underway to ascertain whether *L. pistiae* demonstrates differential specificity at the cryptic interspecies level.

Finally, we suggest that the monospecific status of P. stratiotes should be reconsidered. Following de Queiroz's (2007) approach, we have presented four secondary properties (TCS network, phylogenetic distance, species delimitation and geographic separation) as evidence that at least the **C** and **F** haplotypes are distinct species within *Pistia.* The remaining clades (all New World), while showing substantial differentiation, collectively represent a third species. The unique Australian B7 haplotype points to the possibility that other clades (and perhaps species) exist in under-sampled areas (e.g., trans-Asia and Oceania).

Funding This research was conducted under USDA CRIS Project Numbers 6629-22000-009-00D, 6629-22000-010-00D, and 6629-22000-011-00D.

Availability of data and material Not applicable.

Code availability Not applicable.

Declarations

Conflict of interest Not applicable.

References

Anisimova M, Gil M, Dufayard J-F, Dessimoz C, Gascuel O (2011) Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes. Syst Biol 60(5):685–699. https://doi.org/10.1093/sysbio/syr041

Blake ST (1954) Botanical contributions of the northern Australian regional survey. II. Studies on miscellaneous northern Australian plants. Aust J Bot 1:99–140. https://doi.org/10.1071/BT9540099

Blossey B, Häfliger P, Tewksbury L, Dávalos A, Casagrande R (2018) Host specificity and risk assessment of *Archanara geminipuncta* and *Archanara neurica*, two potential biocontrol agents for invasive *Phragmites australis* in North America. Biol Control 125:98–112. https://doi.org/10.1016/j.biocontrol.2018.1005.1019

Blossey B, Endriss SB, Casagrande R, Häfliger P, Hinz H, Dávalos A, Brown-Lima C, Tewksbury L, Bourchier RS (2020) When misconceptions impede best practices: evidence supports biological control of invasive *Phragmites*. Biol Invas 22:873–883. https://doi.org/10.1007/s10530-019-02166-8

Boutin C, Strandberg B, Carpenter D, Mathiassen SK, Thomas PJ (2014) Herbicide impact on non-target plant reproduction: What are the toxicological and ecological

implications? Environ Pollut 185:294–306. https://doi.org/10.1016/j.envpol.2013.10.009
Cabrera LI, Salazar GA, Chase MW, Mayo SJ, Bogner J, Davila P (2008) Phylogenetic relationships of aroids and duckweeds (Araceae) inferred from coding and noncoding plastid DNA. Am J Bot 95:1153–1165. https://doi.org/10.3732/ajb.0800073
Cabrera Walsh G, Maestro M (2014) Evaluation of intraguild interactions between two species of insect herbivores on *Pistia stratiotes*. Biol Control 76:74–78. https://doi.org/10.1016/j.biocontrol.2014.05.005
Canavan K, Canavan S, Harms NE, Lambertini C, Paterson ID, Thum R (2020) The potential for biological control on cryptic plant invasions. Biol Control 144:104243. https://doi.org/10.1016/j.biocontrol.2020.104243
Casagrande RA, Häfliger P, Hinz HL, Tewksbury L, Blossey B (2018) Grasses as appropriate targets in weed biocontrol: is the common reed, *Phragmites australis*, an anomaly? Biocontrol 63:391–403. https://doi.org/10.1007/s10526-018-9871-y
Center TD, Hill MP, Cordo HU, Julien MH (2002) Waterhyacinth. In: Van Drieche R, Bolssey B, Hoodle M, Lyon S, Reardon R (eds) Biological control of invasive plants in the eastern United States, USDA Forest Service Publication FHTET-2002-04, August 2002, pp 41–64
Chang J-M, Di Tommaso P, Notredame C (2014) TCS: A new multiple sequence alignment reliability measure to estimate alignment accuracy and improve phylogenetic tree reconstruction. Mol Biol Evol 31(6):1625–1637. https://doi.org/10.1093/molbev/msu117
Chernomor O, von Haeseler A, Minh BQ (2016) Terrace aware data structure for phylogenomic inference from supermatrices. Syst Biol 656:997–1008. https://doi.org/10.1093/sysbio/syw037
Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. Mol Ecol 9(10):1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x
Cusimano N, Bogner JS, Mayo J, Boyce PC, Wong SY, Hesse MW, Hetterscheid LA, Keating RC, French JC (2011) Relationships within the Araceae: comparison of morphological patterns with molecular phylogenies. Am J Bot 98:654–668. https://doi.org/10.3732/ajb.1000158
De Queiroz K (2007) Species concepts and species delimitation. Syst Biol 56(6):879–886. https://doi.org/10.1080/10635150701701083
Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic non-coding regions of mitochondrial and chloroplast DNA in plants. Mol Ecol 4:129–131. https://doi.org/10.1111/j.1365-294X.1995.tb00201.x
Dewald LB, Lounibos LP (1990) Seasonal growth of *Pistia stratiotes* L. in South Florida. Aquat Bot 36:263–275. https://doi.org/10.1016/0304-3770(90)90040-R
Dray Jr FA, Center TD (1992) Biological control of *Pistia stratiotes* L. (waterlettuce) using *Neohydronomus affinis* Hustache (Coleoptera: Curculionidae). Tech Report A-92–1. US Army Corps of Engineers, Waterways Experiment Station, Vicksburg. http://hdl.handle.net/11681/6398
Dray FA Jr, Center TD (1989) Seed production by *Pistia stratiotes* L. (water lettuce) in the United States. Aquat Bot 33:155–160. https://doi.org/10.1016/0304-770(89)90029-6
Dray FA Jr, Center TD (2002) Waterlettuce. In: Van Driesche R, Blossey B, Hoddle M, Lyon S, Reardon R (eds) Biological control of invasive plants in the Eastern United States, FHTET-2002-04. United States Department of Agriculture, Forest Service, Morgantown, West Virginia, USA, pp 65–78
Dray FA Jr, Center TD, Habeck DH, Thompson CR, Cofrancesco AF, Balciunas JK (1990) Release and establishment in the southeastern United States of *Neohydronomus affinis* (Coleoptera: Curculionidae), an herbivore of waterlettuce. Environ Entomol 19(3):799–802. https://doi.org/10.1093/ee/19.3.799
Dray FA Jr, Center TD, Wheeler GS (2001) Lessons from unsuccessful attempts to establish *Spodoptera pectinicornis* (Lepidoptera: Noctuidae), a biological control agent of waterlettuce. Biocontr Sci Technol 11:301–316. https://doi.org/10.1080/09583150120055718
Ellstrand N, Schierenbeck K (2006) Hybridization as a stimulus for the evolution of invasiveness in plants? Euphytica 148:35–46. https://doi.org/10.1007/s10681-006-5939-3
Evans J (2013) *Pistia stratiotes* L. in the Florida Peninsula: biogeographic evidence and conservation implications of native tenure for an 'invasive' aquatic plant. Conserv Soc 11:233–246. https://doi.org/10.4103/0972-4923.121026
Florida Department of Environmental Protection (FDEP) (2007) Status of the aquatic plant management program in Florida public waters: Annual report for fiscal year 2005–2006. Florida Department of Environmental Protection, Bureau of Invasive Plant Management, Tallahassee
Florida Fish & Wildlife Conservation Commission (FWCC) (2017) Status of the Aquatic Plant Maintenance Program in Florida Public Waters, Annual Report for fiscal year 2016–2017
Gillett JD, Dunlop CR, Miller IL (1988) Occurrence, origin, weed status, and control of water lettuce (*Pistia stratiotes* L.) in the Northern Territory. Plant Prot Quart 3:144–148
Godfrey RK, Wooten JW (1979) Aquatic and wetland plants of Southeastern United States: Monocotyledons. University of Georgia Press, Athens
Goode AB, Minteer CR, Foley JR, Tipping PW, Valmonte RJ, Knowles BK, Gettys L (2019) Host range of *Lepidelphax Pistiae* (Hemiptera: Delphacidae) and its potential impact on *Pistia stratiotes* L. (Araceae). Biocontrol Sci Technol 29(7):706–714. https://doi.org/10.1080/09583157.2019.1587738
Gould F, Brown ZS, Kuzma J (2018) Wicked evolution: can we address the sociobiological dilemma of pesticide resistance? Science 360(6390):728–732. https://doi.org/10.1126/science.aar3780
Grimm EC, Jacobson GL, Watts WA, Hansen BCS, Maasch KA (1993) A 50,000-year record of climate oscillations from Florida and its temporal correlation with the Heinrich events. Science 261:198–200. https://doi.org/10.1126/science.261.5118.198
Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O (2010) New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307–321. https://doi.org/10.1093/sysbio/syq010

Gutiérrez EE, Helgen KM, McDonough MM, Bauer F, Hawkins MTR, Escobedo-Morales LA, Patterson BD, Maldonado JE (2017) A gene-tree test of the traditional taxonomy of American deer: the importance of voucher specimens, geographic data, and dense sampling. ZooKeys 697:87–131. https://doi.org/10.3897/zookeys.697.15124

Hamilton MB (1999) Four primer pairs for the amplification of chloroplast intergenic regions with intraspecific variations. Mol Ecol 8:513–525

Hanley N, Roberts M (2019) The economic benefits of invasive species management. People nat 1(2):124–137. https://doi.org/10.1002/pan3.31

Harper F (1998) (ed) The travels of William Bartram: Naturalist's Edition. New Haven, Conn: Yale University Press, 1958. Reprint, University of Georgia Press, Athens, Georgia. https://doi.org/10.2307/2954623

Henderson L, Wilson JRU (2017) Changes in the composition and distribution of alien plants in South Africa: an update from the Southern African Plant Invaders Atlas. Bothalia 47(2):2172. https://doi.org/10.4102/abc.v47i2.2172

Hinz HL, Winston RL, Schwarzlander M (2020) A global review of target impact and direct nontarget effects of classical weed biological control. Curr Opin Insect Sci 38:48–54. https://doi.org/10.1016/j.cois.2019.11.006

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 35:518–522. https://doi.org/10.1093/molbev/msx281

Holm LG, Plucknett DL, Pancho JV, Herberger JP (1977) The World's worst weeds: distribution and biology. University Press of Hawaii, Honolulu

Katoh K, Misawa K, Kuma K, Miyata T (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucl Acids Res 30(14):3059–3066. https://doi.org/10.1093/nar/gkf436

Kiewnick S, Holterman M, van den Elsen S, van Megen H, Frey JE, Helder J (2014) Comparison of two short DNA barcoding loci (COI and COII) and two longer ribosomal DNA genes (SSU & LSU rRNA) for specimen identification among quarantine root-knot nematodes (*Meloidogyne* spp.) and their close relatives. Eur J Plant Pathol 140:97–110. https://doi.org/10.1007/s10658-014-0446-1

Kiviat E, Meyerson LA, Mozdzer TJ, Allen WJ, Baldwin AH, Bhattarai GP, Brix H, Caplan JS, Kettenring KM, Lambertini C (2019) Evidence does not support the targeting of cryptic invaders at the subspecies level using classical biological control: the example of *Phragmites*. Biol Invasions 21(8):2529–2541. https://doi.org/10.1007/s10530-019-02014-9

Liebman M, Baraibar B, Buckley Y, Childs D, Christensen S, Cousens R, Eizenburg H, Heijting S, Loddo D, Merotto A, Renton M (2016) Ecologically sustainable weed management: How do we get from proof-of-concept to adoption? Ecol Appl 26(5):1352–1369. https://doi.org/10.1002/15-0995

Masters BC, Fan V, Ross HA (2011) Species delimitation—a Geneious plugin for the exploration of species boundaries. Mol Ecol Resour 11:154–157. https://doi.org/10.1111/j.1755-0998.2010.02896.x

Mayden RL (1997) A hierarchy of species concepts: The denouement in the saga of the species problem. In: Claridge MF, Dawah HA, Wilson MR (eds) Species: the units of biodiversity. Chapman and Hall, London, pp 381–424

Messing RH, Wright MG (2006) Biological control of invasive species: solution or pollution? Front Ecol Environ 4(3):132–40. https://doi.org/10.1890/1540-9295(2006)004[0132:BCOISS]2.0.CO;2

Morais P, Reichard M (2018) Cryptic invasions: a review. Sci Total Environ 613–14:1438–1448. https://doi.org/10.1016/j.scitotenv.2017.06.133

Myers JH, Bazely DR (2003) Ecology and control of introduced plants. Cambridge University Press. 313 pp. ISBN:9780511067440,0511067445

Nauheimer L, Metzler D, Renner SS (2012) Global history of the ancient monocot family Araceae inferred with models accounting for past continental positions and previous ranges based on fossils. New Phytol 195:938–950. https://doi.org/10.1111/j.1469-8137.2012.04220.x

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300

Pemberton RW (2000) Predictable risk to native plants in weed biocontrol. Oecologia 125:489–494. https://doi.org/10.1007/s004420000477

Pickering C (1879) Chronological history of plants. Little, Brown, and Co., Boston

Quillen AK (2009) Diatom-based paleolimnological reconstruction of quaternary environments in a Florida Sinkhole Lake. FIU electronic theses and Dissertations, 211. http://digitalcommons.fiu.edu/etd/211

Quillen AK, Gaiser EE, Grimm EC (2013) Diatom-based paleolimnological reconstruction of regional climate and local land-use change from a protected sinkhole lake in southern Florida, USA. J Paleolimnol 49:15–30. https://doi.org/10.1007/s10933-011-9558-1

Ralser M, Querfurth R, Warnatz H-J, Lehrach H, Yaspo M-L, Krobitsch S (2006) An efficient and economic enhancer mix for PCR. Biochem Biophys Res Commun 347:747–751. https://doi.org/10.1016/j.bbrc.2006.06.151

Renner SS, Zhang L-B (2004) Biogeography of the *Pistia* clade (Araceae): Based on chloroplast and mitochondrial DNA sequences and Bayesian divergence time inference. Syst Biol 53:422–432. https://doi.org/10.1080/10635150490445904

Rolfe J, Windle J (2014) Public preferences for controlling an invasive species in public and private spaces. Land Use Policy 41:1–10. https://doi.org/10.1016/j.landusepol.2014.04.013

Rosenberg NA (2007) Statistical tests for taxonomic distinctiveness from observations of monophyly. Evolution 61:317–323. https://doi.org/10.1111/j.1558-5646.2007.00023.x

Ross HA, Murugan S, Li WLS (2008) Testing the reliability of genetic methods of species identification via simulation. Syst Biol 57:216–230. https://doi.org/10.1080/10635150802032990

Saltonstall K (2002) Cryptic invasion by a non-native genotype of common reed, *Phragmites australis*, into North

America. PNAS 99:2445–2449. https://doi.org/10.1073/pnas.032477999

Saltonstall K (2003) Genetic variation among North American populations of *Phragmites australis*: implications for management. Estuaries 26:444–451. https://doi.org/10.1007/BF02823721

Schmitz DC, Schardt JD, Leslie AJ, Dray FA Jr, Osborne JA, Nelson BV (1993) The ecological impact and management history of three invasive alien aquatic plant species in Florida. In: McKnight WN (ed) Biological pollution: the control and impact of invasive exotic species. Indiana Academy of Science, Indianapolis, pp 173–194

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. Am J Bot 94:275–288. https://doi.org/10.3732/ajb.94.3.275

Smith L, Cristofaro M, Bon MC, De Biase A, Petanović R, Vidović B (2017) The importance of cryptic species and subspecific populations in classic biological control of weeds: a North American perspective. Biocontrol 63(3):417–425. https://doi.org/10.1007/s10526-017-9859-z

Stoddard AA III (1989) The phytogeography and paleofloristics of *Pistia stratiotes* L. Aquatics 11:21–24

Stuckey RL, Les DH (1984) *Pistia stratiotes* (water lettuce) recorded from Florida in Bartrams' Travels, 1765–1774. Aquaphyte 4:6

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109. https://doi.org/10.1007/bf00037152

Templeton AR, Crandall KA, Sing CF (1992) A cladistics analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III Cladogram estimation. Genetics 132:619–633. https://doi.org/10.1093/genetics/132.2.619

Theophrastus (1916) Enquiry into plants, volume I: Books 1–5. Translated by Arthur F. Hort. Loeb Classical Library 70. Harvard University Press, Cambridge

Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucl Acids Res 44:W232–W235. https://doi.org/10.1093/nar/gkw256

U.S. Army Corps of Engineers (USACE) (1977) Corps Report and Supplement to the Final Environmental Impact Statement, Aquatic Plant Control, State of Florida. U.S. Army Engineer District, Jacksonville; Jacksonville, Florida. (unpublished report dated November 1977)

Wallace IM, O'Sullivan O, Higgins DG, Notredame C (2006) M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucl Acids Res 34(6):1692–1699. https://doi.org/10.1093/nar/gkl091

Watts WA, Hansen BCS (1994) Pre-Holocene and Holocene pollen records of vegetation history from the Florida peninsula and their climatic implications. Paleogeo Paleoclim Paleoeco 109:163–176. https://doi.org/10.1016/0031-0182(94)90174-0

Watts WA (1969) A pollen diagram from Mud Lake, Marion County, north-central Florida. Geol Soc Am Bull 80(4):631–642. https://doi.org/10.1130/0016-7606(1969)80[631:APDFML]2.0.CO;2

Watts WA (1975) A late Quaternary record of vegetation from Lake Annie, south-central Florida. Geology 3(6):344–346.https://doi.org/10.1130/0091-7613(1975)3%3C344:ALQROV%3E2.0.CO;2

Weldon LW, Blackburn RD, Harrison DS (1969) Common aquatic weeds, agricultural handbook No. 352, US Gov Printing Office, Washington, DC

Witt A, Beale T, van Wilgen BW (2018) An assessment of the distribution and potential ecological impacts of invasive alien plant species in eastern Africa. Trans R Soc S Afr 73:217–236. https://doi.org/10.1080/0035919X.2018.1529003

Xu X, Liu FX, Ono H, Chen J, Kuntner M, Li D (2017) Targeted sampling in Ryukyus facilitates species delimitation of the primitively segmented spider genus *Ryuthela* (Araneae: Mesothelae: Liphistiidae). Zool J Linn Soc 181(4):867–909. https://doi.org/10.1093/zoolinnean/zlx024

Zhou X, Shen X, Hittinger CT, Rokas A (2018) Evaluating fast maximum likelihood-based phylogenetic programs using empirical phylogenomic data sets. Mol Biol Evol 35:486–503. https://doi.org/10.1093/molbev/msx302