

1 **Phylogeny of the Australian *Solanum dioicum* group using**
2 **seven nuclear genes: Testing Symon's fruit and seed**
3 **dispersal hypotheses**
4

5
6
7 Christopher T. Martine^{1*¶}, Ingrid E. Jordon-Thaden^{2¶}, Angela J. McDonnell^{1¶}, Jason T.

8 Cantley^{3¶}, Daniel S. Hayes^{1&}, Morgan D. Roche^{1,#a&}, Emma S. Frawley^{1,#b&}, Ian S. Gilman^{1,#c&},

9 David C. Tank^{4¶}

10
11 ¹Biology Department & Manning Herbarium, Bucknell University, Pennsylvania, United States
12 of America

13
14 ²Department of Botany, University of Wisconsin, Madison, Wisconsin, United States of America

15
16 ³Biology Department, San Francisco State University, San Francisco, California, United States
17 of America

18
19 ⁴Department of Biological Sciences & Stillinger Herbarium, University of Idaho, Moscow,
20 Idaho, United States of America

21
22 ^{#a}Current address: Department of Ecology and Evolutionary Biology, University of Tennessee,
23 Knoxville, Tennessee, United States of America

24
25 ^{#b}Current address: Biology Department, Saint Louis University, St. Louis, Missouri, United
26 States of America

27
28 ^{#c}Current address: Department of Ecology & Evolutionary Biology, Yale University, New
29 Haven, Connecticut, United States of America

30
31
32
33 *Corresponding Author

34
35 email: ctm015@bucknell.edu (CTM)

36
37
38 ¶These authors contributed equally to this work.

39
40 &These authors also contributed equally to this work.

41

42 **Abstract**

43 The dioecious and andromonoecious *Solanum* taxa (previously described as the “*S. dioicum*
44 group”) of the Australian Monsoon Tropics have been the subject of phylogenetic and taxonomic
45 study for decades, yet much of their basic biology is still unknown. This is especially true for
46 plant-animal interactions, including the influence of fruit form and calyx morphology on seed
47 dispersal. We combine field/greenhouse observations and specimen-based study with
48 phylogenetic analysis of seven nuclear regions obtained via a microfluidic PCR-based
49 enrichment strategy and high-throughput sequencing, and present the first intron-containing
50 nuclear gene dataset in the genus *Solanum* and the first species-tree hypothesis for the *S. dioicum*
51 group. Our results suggest that epizoochorous trample burr seed dispersal (strongly linked to
52 calyx accrescence) is far more common among Australian *Solanum* than previously thought and
53 support the hypothesis that the combination of large fleshy fruits and endozoochorous dispersal
54 represents a reversal in this study group. The general lack of direct evidence related to biotic
55 dispersal (epizoochorous or endozoochorous) may be a function of declines and/or extinctions of
56 vertebrate dispersers. Because of this, some taxa might now rely on secondary dispersal
57 mechanisms (e.g. shakers, tumbleweeds, rafting) as a means to maintain current populations and
58 establish new ones.

59 **Introduction**

61 The large and cosmopolitan plant genus *Solanum* L. consists of nearly 1,400 accepted
62 species [1], the majority of them exhibiting fleshy fruits linking them to biotic agents that
63 disperse seeds as a consequence of frugivory [2]. Symon, [3] focusing only on the ca. 90
64 *Solanum* species described for Australia at the time, defined a set of fruit morphologies that

65 correspond to hypothesized dispersal categories that can be summarized as follows: 1)
66 pulpy/fleshy berries of various colors dispersed following ingestion (67 species), 2) firm,
67 ultimately bony berries enclosed in a calyx with unclear dispersal (perhaps ingestion) (10
68 species), 3) smallish berries enclosed in a prickly calyx and ostensibly functioning as trample
69 burrs on the feet of mammals (6 species), and 4) a postulated set of “oddball” fruits/mechanisms
70 not matching those above (8 species).

71 A fair number of the species in categories 2, 3, and 4 belong to a recently-radiated [4,5]
72 group of spiny solanums (*Solanum* subgenus *Leptostemonum*) known as the “*Solanum dioicum*
73 group” [6,7], a set of species restricted to the Australian Monsoon Tropics (AMT) of northern
74 Western Australia, the northern portions of the Northern Territory, and far-western Queensland.
75 This species-group was first recognized by Whalen [7] as a diverse group of erect to spreading
76 shrubs, and is the only group of *Solanum* species that includes both a large number of
77 andromonoecious species and a set of cryptically dioecious species. Already considered unusual
78 among Australian solanums for their breeding systems [8,9], these species are also markedly
79 variable in morphology -- with many taxa easily recognized via differences in vegetative form,
80 habit, tomentum and armature, among other characters. Some of the greatest diversity, however,
81 is related to fruit form and the degree to which fruits are enclosed by armed calyces, characters
82 exhibiting particular influence on seed dispersal.

83 The 16 *S. dioicum* group taxa treated by Symon [3] include species that appear to
84 disperse seeds via fruit fracturing (*S. oedipus*, *S. heteropodium*), censer-like mechanisms (*S.*
85 *tudununggae*), or by catching in fur as trample burrs (*S. leopoldense*, *S. asymmetriphyllum*). The
86 dispersal mechanisms of three other taxa were considered unclear by Symon (*S. carduiforme*, *S.*
87 *petraeum*, *S. cataphractum*). The remaining species (9) were placed in the more typical fleshy-

88 fruited category, particularly in the subcategory defined by Symon as having large (2-4 cm
 89 diameter), firm, yellowish berries at maturity (Table 1; Fig 1). Symon also assigned the trample
 90 burr character to members of the *S. echinatum* group (sensu Bean [8]), a clade of species known
 91 to be closely allied to the *S. dioicum* group based on more recent phylogenetic work [4,5].

92

93 **Table 1. Species of the *S. dioicum* group and *S. echinatum* group sensu Bean [6] considered**
 94 **in this study.**

Species	Clade	Herbarium voucher w/ fruits	Symon's dispersal designation	Our dispersal designation	Notes
<i>S. apodophyllum</i> A.R. Bean	<i>S. clarkiae</i> + <i>S. melanospermum</i>	Not examined	not considered	ingested	Species described in 2016.
<i>Solanum asymmetriphyllum</i> Specht	Kakadu dioecious	CTM 3161 (BUPL), C!	trample burr	trample burr	Includes 'Deaf Adder' morphotype.
<i>Solanum beagleholei</i> Symon	Andromonoecious bush tomatoes	CTM 4050 (BUPL), C!	ingested	ingested	
<i>Solanum</i> sp. nov. 'Buchanan'	Andromonoecious bush tomatoes	CTM 4743 (BUPL), C!	not considered	ingested	Undescribed taxon.
<i>Solanum cataphractum</i> Benth.	Kimberley dioecious	Not examined	unclear	trample burr	
<i>Solanum carduiforme</i> F. Muell.	Kimberley dioecious	CTM 4229 (BUPL)	unclear	trample burr (sheet flow)	Includes multiple morphotypes.
<i>Solanum chippendalei</i> Symon	Andromonoecious bush tomatoes	CTM 4724 (BUPL), C!	ingested	ingested (shaker)	Mechanical dispersal witnessed (Martine).
<i>Solanum clarkiae</i> Symon	<i>S. clarkiae</i> + <i>S. melanospermum</i>	CTM 4728, C!	ingested	ingested	Evidence of fruit removal recorded (Martine).
<i>Solanum cowiei</i> Martine	Kimberley dioecious	K.G. Brennan 7274 (DNA)	not considered	trample burr	Species described in 2013. Includes multiple

					morphotypes.
<i>Solanum cunninghamii</i> Benth.	Kimberley dioecious	Not examined	ingested	trample burr	
<i>Solanum dioicum</i> W. Fitzg.	Kimberley dioecious	CTM 4040 (BUPL), C!	ingested	trample burr (ingested?)	Includes multiple known morphotypes. One observation of bustard handling fruits. (Martine)
<i>Solanum diversiflorum</i> F. Muell.	Andromonoecious bush tomatoes	CTM 4033 (BUPL), C!	ingested	ingested	Observations of mammal frugivory by Martu people and R. Bird.
<i>Solanum eburneum</i> Symon	Andromonoecious bush tomatoes	CTM 4007 (BUPL), C!	ingested	ingested (shaker)	Includes all morphotypes of the species sampled for phylogeny.
<i>Solanum echinatum</i> R. Br.	<i>S. echinatum</i> group	CTM 4206 (BUPL)	trample burr	trample burr (sheet flow)	
<i>Solanum</i> sp. nov. 'Fitzroy'	Andromonoecious bush tomatoes	CTM 4036 (BUPL), C!	not considered	ingested	Undescribed taxon.
<i>Solanum heteropodium</i> Symon	<i>S. oedipus</i> + <i>S. heteropodium</i>	AJM 375 (BUPL), C!	fracturing	trample burr (shaker)	
<i>Solanum leopoldense</i> Symon	Kimberley dioecious	AJM 374 (BUPL), C!	trample burr	trample burr	
<i>Solanum</i> sp. 'Longini Landing'	Kimberley dioecious	Not examined	not considered	trample burr?	Fruits unknown. Undescribed taxon.
<i>Solanum lucani</i> F. Muell.	<i>S. echinatum</i> group	CTM 4054 (BUPL)	trample burr	trample burr	
<i>Solanum medicagineum</i> A.R. Bean	<i>S. echinatum</i> group	CTM 4722 (BUPL), C!	not considered	trample burr	Species described in 2012.
<i>Solanum melanospermum</i> F. Muell.	<i>S. clarkiae</i> + <i>S. melanospermum</i>	CTM 4216 (BUPL), C!	ingested	ingested (tumbleweed)	
<i>Solanum ossicruentum</i> Martine & J.	Kimberley dioecious	CTM 4011 (BUPL), C!	not considered	trample burr (sheet flow)	Species described in 2016.

Cantley					
<i>Solanum oedipus</i> Symon	<i>S. oedipus</i> + <i>S. heteropodium</i>	CTM 814 (CONN)	fracturing	trample burr (shaker)	
<i>Solanum petraeum</i> Symon	Kimberley dioecious	CTM 833 (CONN)	unclear	trample burr	
<i>S. phlomoides</i> A. Cunn. ex. Benth.	Andromonoecious bush tomatoes	Albrecht 11246 (DNA), C!	ingested	ingested	Observations of mammal frugivory by Martu people and R. Bird.
<i>Solanum raphiotes</i> A.R. Bean	<i>S. echinatum</i> group	CTM 1703 (BUPL), C!	not considered	trample burr	Species described in 2012.
<i>Solanum sejunctum</i> K. Brennan, C. Martine, & Symon	Kakadu dioecious	CTM 1729 (BUPL), C!	not considered	trample burr	Species described in 2006.
<i>Solanum succosum</i> A.R. Bean & Albr.	Andromonoecious bush tomatoes	CTM 4273 (BUPL), C!	not considered	ingested	Species described in 2008.
<i>Solanum</i> sp. ‘Galvans Gorge’	Kimberley dioecious	Not examined	not considered	trample burr?	Fruits unknown. Undescribed taxon.
<i>Solanum tudununggae</i> Symon	Kimberley dioecious	CTM 823 (CONN)	censer	trample burr (censer/shaker)	Mechanical dispersal witnessed (Symon).
<i>Solanum ultraspinosum</i> A.R. Bean	<i>S. clarkiae</i> + <i>S. melanospermum</i>	CTM 4002 (BUPL), C!	not considered	ingested	Species described in 2016. Evidence of fruit removal recorded (Martine).
<i>Solanum vansittartense</i> C. Gardner	Kimberley dioecious	Not examined	ingestion	trample burr?	
<i>Solanum watneyi</i> Martine & Frawley	Andromonoecious bush tomatoes	CTM 4065 (BUPL), C!	not considered	ingested	Species described in 2016.

95
96 Associated fruiting herbarium vouchers (collector & number + herbarium acronym in
97 parentheses; taxa marked with “C!” have also been cultivated at Bucknell as living specimens)
98 along with dispersal methods as assigned by Symon [3] and the authors of this study (proposed
99 secondary dispersal method in parentheses). Taxa included in the phylogenetic analyses are in

100 bold. For taxa where fruiting specimens are unknown, dispersal designation is inferred based on
101 phylogeny. (See S1 File for complete herbarium voucher and GenBank accession details.)

102

103 **Figure 1. Fruit and calyx forms of selected AMT *Solanum* taxa.** Photos A-C: Putative
104 epizoochorous trample burr dispersal via accrescent prickly calyx (A: *S. carduiforme*, B: *S.*
105 *ossicruentum*, C: *S. asymmetriphyllum*). Photos D-F: Putative ingestion dispersal after reflexing
106 of accrescent calyx (D: *S. melanospermum*, E: *S. ultraspinosum* [fruit intact], F: *S. ultraspinosum*
107 [fruit removed by unknown frugivore]). Photos G-I: Putative ingestion dispersal, calyx not
108 enveloping fruit (G: *S. beaugleholei* [mature]; H: *S. diversiflorum* [immature, showing “cryptic
109 coloration”], I: *S. chippendalei* [post-mature fruits exhibiting “shaker” mechanism]). Photos by C.
110 Martine.

111

112 Calyx morphology is variable across the *S. dioicum* group, with calyces all prickly
113 (sometimes heavily so) and accrescent to varying degrees of fruit coverage (illustrated in Fig 1).
114 Fruits with calyces enveloping the fruit by half or less are assumed to fit into Symon’s broad
115 category of ingested fruits, while those that are fully enveloped by calyces are assumed to be
116 trample burrs. The exceptions to the latter condition recognized by Symon are 1) *S. tudumunggae*
117 and its censer mechanism, and 2) members of the *S. melanospermum* + *S. clarkiae* clade [5] in
118 which the accrescent calyx reflexes and presents the berries at maturity – representing a sort of
119 pre-dispersal defense mechanism to prevent the ingestion and distribution of immature seeds.
120 Likewise, Symon [3] suggested that even immature fruits without enveloping calyces are
121 cryptically colored (being green or striped green) and exhibit greater levels of alkaloids than
122 those that are mature, a character present across the fleshy-fruited taxa (see Fig 1-H).

123 Specific seed dispersal mechanisms have not been published for any of the spiny
124 *Solanum* species in the AMT. Seeds of some taxa can survive gut passage and are germinable
125 after ingestion by rats *ex situ* [10], but throughout nearly 20 years of AMT *Solanum* field
126 observations and inspection of thousands of wild plants by Martine and colleagues, frugivory
127 (whether by direct observation or removed fruits) has rarely been witnessed (see Table 1).
128 Indigenous knowledge of the biota of the AMT does, however, confirm that a few *Solanum*
129 species are eaten by native mammals, especially macropods. The Martu people of the western
130 desert report hill kangaroo (*Macropus robustus*) and burrowing bettong (*Bettongia leseuer*, now
131 locally extinct) frugivory on *S. diversiflorum* and *S. phlomoides* [11] (R. Bird, pers comm). Local
132 indigenous groups from the region of Kakadu National Park (a biodiversity hotspot in far-
133 northern Northern Territory) suggest that the fruits of a few regional endemics may also
134 occasionally be ingested by rock kangaroos [12], but only one species from the region's flora
135 (*Gardenia fucata*, Rubiaceae) has been found to be ingested and effectively (although rarely)
136 dispersed by these reclusive marsupials [13]. Most AMT *Solanum* populations we have observed
137 either retain the majority of their fruits well into the following season or, if the fruits are abscised
138 when mature, the berries lay in uneaten piles beneath the parent plants. Given the limited extant
139 evidence for frugivory/dispersal, one is left to ponder whether the production of large, fleshy
140 berries by some AMT solanums represents an anachronism related to the mass extinction of
141 nearly all large-bodied Australian animals over the last 400,000 years; or, perhaps, whether the
142 production of such fruits is a relict of ancestral morphology and not related to biotic dispersal in
143 Australia, at all.

144 Since Symon's revision [14], many new species have been described from the region [15-
145 22]. We here revisit seed dispersal strategies among the AMT *Solanum* taxa based on the present

146 taxonomic understanding of the group and two decades of additional observations, primarily to
147 test the hypothesis presented by Symon [3] that biotic dispersal by ingestion is the most common
148 dispersal mechanism (with a few exceptions). Through new observations coupled with
149 phylogenetic data we here show that not only are large, fleshy fruits less common than burr fruits
150 in our study group, but also that they appear to represent a recent reversal in AMT *Solanum*.

151

152 **Methods**

153 **Taxon sampling**

154 Seventy-six individuals representing 50 species of *Solanum* were sampled from field
155 collected dried tissue and/or fragments removed from pressed, dried herbarium specimens. This
156 included 35 ingroup taxa to thoroughly sample clades and species-groups identified by Martine
157 et al. [4,5] and to sample morphological diversity identified via fieldwork and specimen
158 examination. We also sampled nine spiny *Solanum* outgroups, including three species from the
159 closely-allied *S. echinatum* group sensu Bean [6] and six non-spiny *Solanum* outgroup species.
160 Seven putative single-copy, intron-containing nuclear loci were newly identified and sequenced
161 for all accessions (Jordon-Thaden, in prep; Table 2).

162

163 **Table 2: Primer sequences (5'-3') used to amplify loci used in this study.**

164

Primer pair name	Locus short name	Forward primer sequence	Reverse primer sequence
S_2G01490_1254_1752_3	SOL_3	CAA CAT GCA CTG GCT TGT G	GCC CAT TTG CAG CCA TTA
S_4G00740_1029_1529_8	SOL_8	TGA GGA CCG CCC TTG ATA	ATC TTG CTT GGG CCA CTG
S_4G09750_777_1310_9	SOL_9	GGA GAG GCT GCT CTT ACC AA	GGA GAG GCT GCT CTT ACC AA

S_5G10460_979_1452_12	SOL_12	TTG GCA ACT TGT GGT GCA	GGC TTC AGT TCC GTG AGC
S_5G42520_564_1106_14	SOL_14	CAT GGA GCG GGA TAA TGC	GGG ACT GGC ATG GTT GTT
S_5G46800_295_804_15	SOL_15	TCA AAG CCA GCC TAC TCC A	ATC CAT TGG CCC TGC ATA
S_5G54080_500_1155_16	SOL_16	GCC TGT CGA GAT CCC AGA	AGC AGC CAG ACC ATT TGC

165
166

167 **Primer design, validation, & sequencing**

168 Primers for this study were designed using MarkerMiner v.1.0 [23]. First, transcriptome
169 assemblies for *Solanum dulcamara* L., *S. ptycanthum* Dunal ex DC., *S. cheesemanii* Geras., *S.*
170 *sisymbriifolium* Lam., *S. lasiophyllum* Humb. & Bonpl. ex Dunal, *S. xanthocarpum* Schrad. &
171 J.C. Wendl., were obtained from the 1 KP Project (www.onekp.org) and mapped to the
172 *Arabidopsis thaliana* (L.) Heynh. genome. MarkerMiner v.1.0 was used to mine the resulting
173 alignments for transcripts of single copy nuclear orthologs by filtering transcripts with scaffold
174 length at least 900 bp and BLAST similarity at least 70% using both BLASTX and TBLASTN.
175 The output was used to design primer pairs for the Fluidigm Access Array System (California,
176 USA) using the Primer3 plugin [24] in Geneious (Biomatters Limited, New Zealand). Primers
177 were selected that flanked predicted intronic regions with a size of 400-700 bp. Primer design
178 and validation followed Uribe-Convers et al. [25] and the Access Array System protocol
179 (Fluidigm, San Francisco, California, USA). Target-specific portions of primer pairs had a length
180 of 20-25 bp, a melting temperature between 59 and 61° C, and contained no more than 3bp of
181 homopolymer sequence for any nucleotide. Forty-eight primer pairs were selected for *Solanum*.
182 To provide an annealing site for the Illumina sequencing adapters and sample-specific barcodes,
183 a conserved sequence (CS) tag was added to the 5' end of the forward and reverse primers at the

184 time of oligonucleotide synthesis (CS1 for forward primers and CS2 for reverse primers [25];
185 purchased from Operon, Eurofins Scientific, Luxembourg).

186 To validate the primers, PCR amplification was conducted for 24 of the 48 primer pairs;
187 validation reactions simulate the four-primer reaction of the Fluidigm microfluidic PCR system
188 using a standard thermocycler. Reaction were carried out in 10 μ l volumes each with final
189 concentrations of reagents as follows: 1ng/ μ L gDNA (placed in master mix), 400 μ M of each
190 dNTP, 2x Fast Start High Fidelity reaction buffer (Roche), 9mM MgCl₂ (Roche), 10% DMSO,
191 0.1U/ μ L FastStart High Fidelity Enzyme Blend (Roche), 2X AccessArray Loading Reagent
192 (Fluidigm), 800nM AccessArray Barcode Primers for Illumina (designed by University of Idaho
193 Institute for Bioinformatics and Evolutionary Studies Genomics Resources Core facility,
194 iBEST), 200nM target specific primer mix (Operon), and water. PCR was conducted using
195 following program: one round of 50°C for 2 minutes, 70°C for 20 minutes, 95°C for 10 minutes;
196 ten rounds of 95°C for 15 seconds, 60°C for 30 seconds, 72°C for 1 minute; two rounds of 95°C
197 for 15 seconds, 80°C for 30 seconds, 60°C for 30 seconds, and 72°C for 1 minute; eight rounds of
198 95°C for 15 seconds, 60°C for 30 seconds, 72°C for 1 minute; two rounds of 95°C for 15
199 seconds, 80°C for 30 seconds, 60°C for 30 seconds, 72°C for 1 minute; six rounds of 95°C for 15
200 seconds, 60°C for 30 seconds, 72°C for 1 minute; five rounds of 95°C for 15 seconds, 80°C for
201 30 seconds, 60°C for 30 seconds, 72°C for 1 minute. Each of the primer pairs were tested on
202 three *Solanum* species (*S. asymmetriphyllum*, *S. eburneum*, and *S. elaeagnifolium*) and PCR
203 products were visualized on 1.5% agarose gels run at 80V for 90 minutes. Validation was
204 considered successful when, in all three taxa, amplification was observed as a single band that
205 was not present in the negative and was within the appropriate size range [25].

206 Microfluidic PCR was carried out for 96 *Solanum* samples in the Fluidigm Access Array
207 system (Fluidigm, San Francisco, California, USA) following the manufacturer's protocols. Each
208 of the 96 samples were amplified with the 24 primer pairs two times to completely fill the 48-
209 well Fluidigm chip. To remove unused reagents and/or undetected primer dimers smaller than
210 ~250 bp, each pool was purified with 0.6X AMPure XP beads (Agencourt, Beverly,
211 Massachusetts, USA). PCR pools were analyzed in a Bioanalyzer High-Sensitivity Chip (Agilent
212 Technologies, Santa Clara, California, USA) and standardized to 13 pM using the KAPA qPCR
213 kit (KK4835; Kapa Biosystems, Woburn, Massachusetts, USA) on an ABI StepOnePlus Real-
214 Time PCR System (Life Technologies, Grand Island, New York, USA). The resulting pools were
215 multiplexed and sequenced in an Illumina MiSeq (San Diego, California, USA) to obtain 300 bp
216 paired-end reads. Microfluidic PCR, downstream quality control and assurance, and sequencing
217 were carried out at iBEST.

218

219 **Data processing**

220 Raw reads were filtered, trimmed, and demultiplexed by barcode and target-specific
221 primer using dbcAmplicons (<https://github.com/msettles/dbcAmplicons> [25]) and merged using
222 FLASH [26]. Consensus sequences for each sample in all amplicons were generated using the
223 reduce_amplicons R script
224 (https://github.com/msettles/dbcAmplicons/blob/master/scripts/R/reduce_amplicons.R [25]).

225

226 **Phylogenetic Analyses**

227 Individual loci were aligned using MUSCLE v3.8.5 [27] and adjusted manually.
228 Phylogenetic topologies were estimated from the seven loci (summarized in Table 3) using three

229 methods. First, gene trees for each locus were estimated using IQ-TREE version 1.5.5 [28,29]
230 with the optimal substitution models selected by ModelFinder [30]. Clade support was
231 determined by nonparametric bootstrapping using the ultrafast bootstrap with 1000 replicates
232 [31]. Next, the resulting gene trees were used as input to ASTRAL-III [32] to estimate a species
233 tree and multi-locus bootstrapping to calculate local posterior probability values for the species
234 tree (100 replicates). Finally, we used Bayesian inference to estimate a phylogeny of a
235 concatenated dataset using MrBayes 1.6.4 in parallel [33-35]. We applied the GTR+I+G model
236 but allowed substitution rates for each gene to vary. The topology, branch lengths, shape, and
237 state frequencies were unlinked. We sampled trees from two runs using eight chains (two hot, six
238 cold) that ran for 5×10^8 generations with trees and model parameters sampled every 5000
239 generations. Convergence of the runs and stationarity were assessed for all parameters with the
240 assistance of Tracer v.1.6.0 [36]. The Bayesian maximum clade credibility topology with
241 posterior probabilities for each bipartition was summarized after discarding 25% of the sampled
242 trees as burn-in using TreeAnnotator v.2.4.1 [37].

243

244 **Ancestral state reconstruction**

245 We constructed an unordered morphological character matrix in Mesquite v.3.5 [38] for
246 calyx form at fruit maturity, using the following character states: 1) reflexed, 2)
247 subtended/typical, and 3) enclosed/trample burr. Character states were obtained for each taxon
248 from Symon's monograph [14] and/or field observations by the authors. We mapped calyx form
249 onto the ASTRAL-III species tree estimate using unordered parsimony in Mesquite. The number
250 and locations of transitions between character states was noted to determine whether or not calyx

251 form is homoplasious and to better understand the associations between shifts in morphology
252 within and among clades.

253

254 **Results**

255 **Primer design and validation**

256 MarkerMiner v.1.0 allowed for efficient design of primer pairs targeting putatively
257 single-copy nuclear regions. We designed 48 pairs of primers for *Solanum*, of which 32 pairs
258 were successfully validated and 24 were chosen for final amplification and sequencing in all
259 samples (S1 File).

260

261 **Nuclear gene sequencing**

262 Seven of the original 48 primer pairs we attempted obtained enough sequence data and
263 coverage for downstream use (Table 2). In total, 503 new *Solanum* sequences were generated
264 and analyzed. Data from this study are available in GenBank (S2 File).

265

266 **Phylogenetic analyses**

267 Characteristics of the individual alignments used for phylogenetic analysis are
268 summarized in Table 3. Individual gene trees estimated with IQ-TREE reveal weakly supported
269 relationships among outgroup and ingroup clades and low overall resolution within sampled
270 Australian *Solanum* (S1 Fig, trees available on Treebase xyz). No individual topologies lend
271 support to the monophyly of the Australian species but do support some relationships within

272 species and among close relatives. The relationships between many species and within species
273 that are represented by multiple individuals are unresolved by the individual genes.

274

275

276 **Table 3: Alignment characteristics by locus.**

277

Locus	# seqs in alignment	Ungapped length (bp)	Aligned length (bp)	Conserved sites	Variable sites	Parsimony informative sites	Missing data (%)
SOL_3	76	283-476	483	353	126	44	1.6
SOL_8	80	458-465	466	313	130	47	1.1
SOL_9	76	561	676	284	290	137	0
SOL_1 2	84	267	267	161	106	65	0
SOL_1 4	77	432-444	444	319	125	60	0
SOL_1 5	64	565	673	377	228	120	0
SOL_1 6	46	566	580	328	241	98	0

278

279

280 The ASTRAL-III species tree analysis was informed by the seven gene trees and
281 attributed conflicts in the data to incomplete lineage sorting. This analysis uncovered a
282 moderately well supported topology that is aligned with previous studies [4,5] (Fig 2). The
283 ASTRAL-III topology shows Australian *Solanum* is divided into four lineages: the Kimberley
284 dioecious clade, the Andromonoecious bush tomatoes, the Kakadu dioecious clade, and the *S.*
285 *echinatum* group. The earliest diverging of these was the Kimberley dioecious clade (local
286 posterior probability = 0.83). The Andromonoecious bush tomatoes were found to be

287 polyphyletic in this study due to the position of *S. oedipus*, which occurred as a sister both to the
288 Kakadu dioecious clade and to the rest of the Andromonoecious bush tomatoes. However, the
289 remaining species of Andromonoecious bush tomatoes form a clade (local posterior probability
290 <0.70). The Kakadu dioecious clade is well supported (local posterior probability = 0.94), and
291 the *S. echinatum* group is moderately supported (local posterior probability = 0.79).

292

293 **Figure 2: ASTRAL-III species tree generated from ML gene trees estimated in IQ-TREE.**

294 Values at nodes reflect local posterior probabilities of .50 or greater. Clade labels follow
295 Martine, et al. (4,5). The *S. echinatum* group is identified sensu Bean [6].

296

297 An identical overall topology was recovered by the Bayesian maximum clade credibility
298 tree (Fig 3) which was derived from a concatenated partitioned analysis of all loci, although the
299 posterior probabilities (pp) from this analysis vary. The Kimberley dioecious clade was less well
300 supported (pp < 0.70), while the clade of Andromonoecious bush tomatoes (excluding *S.*
301 *oedipus*), the Kakadu dioecious clade, and the *S. echinatum* group were more well supported (pp
302 of 0.99, 1.0, and 1.0 respectively). Also worth noting is the nonmonophyly of many species for
303 which multiple accessions were included, a pattern also apparent in the individual gene trees (S1
304 Fig).

305

306 **Figure 3: Maximum clade credibility topology inferred by Bayesian inference from**
307 **concatenated partitioned loci.** Red nodes reflect posterior probabilities of <0.9 and black nodes
308 reflect posterior probabilities of 0.9-1. Clade labels follow Martine, et al. (4,5). The *S. echinatum*
309 group is identified sensu Bean [6].

310

311 Overall, the phylogenetic analyses support a close relationship between the *S. dioicum*
312 group sensu Whalen [7] and the *S. echinatum* group sensu Bean [6], and recover with some
313 confidence three of the five clades within the *S. dioicum* group as identified by Martine et al. [5]:
314 the Kimberley dioecious clade, the Kakadu dioecious clade, and the andromonoecious bush
315 tomatoes. Of the two remaining clades not recovered here, the close relationship of the *S.*
316 *melanospermum* + *S. clarkiae* clade to the bush tomatoes is inferred by the placement of the
317 single species from the former clade in our dataset, *S. ultraspinosum*. The placement of *S.*
318 *oedipus*, the single species included representing the *S. oedipus* + *S. heteropodium* clade, renders
319 the bush tomatoes paraphyletic. Likewise, placement of the multiple morphotypes of *S.*
320 *eburneum*, *S. carduiforme*, *S. dioicum*, and *S. cowiei* suggests that these species are not
321 monophyletic as currently circumscribed and in need of revisionary study.

322

323 **Ancestral state reconstruction**

324 The character state reconstruction (Fig 4) supports a typical or subtended calyx as
325 ancestral in these *Solanum* species, including the spiny *Solanum* outgroup taxa. Within the
326 Australian *S. dioicum* group, there is a transition to an enclosed or trample burr calyx
327 morphology, which is the ancestral form in the *S. dioicum* group. There is equivocal support for
328 a transition to reflexed calyx morphology or a reversal to the typical or subtended calyx
329 morphology within most of the Andromonoecious bush tomatoes. *Solanum oedipus*, the
330 andromonoecious species sister to the rest of the bush tomatoes plus the Kakadu dioecious taxa,
331 retains the enclosed (trample burr) calyx form.

332

333 **Figure 4: Most parsimonious ancestral state reconstruction of calyx morphology mapped**
334 **onto the ASTRAL-III species tree.** Clade labels follow Martine, et al. (4,5). The *S. echinatum*
335 group is based on Bean [6].

336
337

338 **Discussion**

339 We present a phylogeny of the *Solanum dioicum* group derived from nuclear data mined
340 with MarkerMiner v.1.0 using 1KP transcriptomes with sequence data generated via microfluidic
341 PCR-based amplicon enrichment and Illumina sequencing. This is the first study of the group to
342 include intron-containing nuclear genes for phylogenetic analyses in the genus. These data
343 support a monophyletic assemblage that includes the *S. dioicum* group species and reveals a
344 close alliance with the *S. echinatum* group taxa. Future phylogenomic study of these plants will
345 target members of both groups to test the close relationship uncovered here (McDonnell &
346 Martine, in prep). The species tree hypothesis for these groups was used to investigate the
347 evolution of calyx form within the lineage in an effort to better understand morphological
348 diversity and natural history of these plants.

349 Calyx form and dispersal type are somewhat closely correlated with breeding systems
350 among the species in what now might be considered the “*S. dioicum* + *S. echinatum* Group.”
351 Species with large, fleshy fruits and mostly non-acrescent calyces are andromonoecious, while
352 those with putative trample burr dispersal are dioecious or hermaphrodite. The phylogeny
353 inferred by our analyses suggests that the andromonoecious breeding system (and thus their fruit
354 type) is derived in our study group (having arisen from either dioecy or hermaphroditism), a
355 hypothetical sequence that runs counter to the hermaphroditism-andromonoecy-dioecy sequence
356 proposed by previous authors [3-5,8,9,14]. While this would upend conventional understanding

357 of breeding system evolution in this group, we acknowledge that the limitations of our data force
358 us to maintain skepticism pending additional study. The phylogeny presented here suffers from
359 poor resolution which is a natural result from our use of a small number of short loci. We believe
360 primer design challenges, such as the initial limitations of comparing *Solanum* transcripts to
361 *Arabidopsis* genomes with MarkerMiner 1.0, could be a cause for the failure of many primer
362 pairs to successfully amplify loci. However, we are confident that our current species-tree
363 hypothesis for the group is plausible and that the combination of our phylogeny with many years
364 of field observations offers worthwhile insight into the role of fruit and calyx form in seed
365 dispersal among Australian *Solanum*.

366

367 **Fruit/calyx form and seed dispersal**

368

369 Despite decades of study, the natural history of solanums in the AMT is still quite poorly
370 understood – particularly with regard to plant-animal interactions including pollination [39] and
371 seed dispersal [3]. Nearly all of the AMT *Solanum* species produce fleshy berries; and although
372 traditional knowledge suggests that some animals do eat *Solanum* fruits [11,12] and may disperse
373 seeds, no effective biotic seed dispersal interaction has yet been identified for a single species
374 within the group. Still, one might infer that biotic dispersal following frugivory is a likely means
375 by which many AMT solanums colonize new areas. Symon [3] suggested that birds, mammals,
376 and lizards all play a role in endozoochorous dispersal of *Solanum* seeds in Australia;
377 meanwhile, epizoochorous and non-biotic modes of dispersal (including wind, water, and
378 mechanical modes) were hypothesized to play a secondary role.

379 As Symon pointed out, the interaction of fruit form and calyx morphology in seed
380 dispersal appears to be significant. Based on a combination of field and greenhouse observations,
381 herbarium study, experimental work, description of new species, and molecular phylogenetics
382 over the last 15+ years, we suggest that for the *S. dioicum* group, where many species exhibit
383 accrescent fruiting calyces for all or part of fruit/seed development, calyx form may be even
384 more important than previously assumed. Specifically, current understanding of diversity in this
385 group (including the many new species described since Symon’s 1981 monograph [14]) suggests
386 that burr-fruited species (with accrescent calyces that remain closed after fruit maturity) are more
387 common than calyces that only partly cover fleshy fruits. Whereas Symon hypothesized that six
388 species across all of *Solanum* in Australia dispersed by trample burrs, we suggest that this
389 number is closer to 18 in the *S. dioicum* group sensu stricto alone (Table 1). Likewise, the
390 phylogenetic analyses infer that the *S. echinatum* group sensu Bean [6] (here represented by *S.*
391 *echinatum*, *S. raphiotes*, and *S. lucani*) is nested in the *S. dioicum* group with which it shares the
392 enclosed trample burr morphology (or, at least, an accrescent calyx). Recent work by Bean [40]
393 suggests that species diversity in the *S. echinatum* group is greater than currently described and,
394 therefore, so is the presence of trample burr dispersal in that clade.

395 For the species with accrescent calyces, the calyx appears to play two significant roles.
396 One function, for most species, is in fruit and seed dispersal. The other function is likely that of a
397 “pre-dispersal defense,” or the prevention of unripened fruits from being removed and ingested
398 before the seeds are mature and germinable. In the case of some species, such as *S. ossicruentum*,
399 readiness of seeds appears to correspond with whole fruits (with enveloping calyx) dropping off
400 the plants entirely [21]. In perhaps the most elegant example of pre-dispersal defense, the
401 members of the *S. melanospermum* + *S. clarkiae* Clade (those two species plus the recently-

402 described *S. ultraspinosum* and *S. apodophyllum* [16]) enclose their unripened fruits in heavily-
403 armed calyces before presenting them, still attached, with reflexed calyces at fruit maturity (Fig
404 1-E). This phenomenon is unusual in the family Solanaceae, where accrescent calyces are
405 common but calyces reflexing to expose ripened berries is known from just two other genera [2]
406 (*Jaltomata* and *Chamaesaracha*). Pre-dispersal defense via secondary chemistry among
407 Australian species with non-acrescent calyces was proposed by Symon [3], with green striping
408 on fruits produced by most of the non-acrescent species presumed to be both a warning signal
409 and a camouflage via cryptic coloration against green foliage (Fig 1-H).

410 In addition to an assumed greater role for trample burr dispersal, our observations have
411 also revealed that some of the secondary abiotic dispersal strategies proposed by Symon [3] are
412 more common than previously imagined. Intriguingly, some of these seem to function as
413 dispersal insurance for when fleshy fruits are not removed by frugivores and, instead, are left to
414 dry out and eventually break apart. The mechanical “censer” dispersal Symon [3] ascribed to a
415 single Australian species, *S. tudununggae*, might be better broadly described as a “shaker”
416 mechanism inclusive of other species in which this has been observed. A prime example is *S.*
417 *chippendalei*, where hundreds of individuals (near Taylor Creek, Northern Territory) were
418 witnessed bearing the previous years’ fruits which, now dry and brittle, were dropping loose
419 seeds from open fruit wall cavities (Fig 1-J) with the slightest touch or breeze (Martine, pers.
420 obs.). Even in cultivation, *S. chippendalei* fruits age differentially – with some areas of the fruit
421 wall breaking down much quicker than others as fruits pass beyond maturity. Similar shaker
422 dispersal has also been seen in *S. eburneum* and may apply to what Symon³ described as
423 “fracturing” in *S. oedipus* and *S. heteropodium*.

424 Symon [3] suggested that *S. pugiunculiferum*, a species found in tidal mudflats that is not
425 a member of our study group, might be described as using tumbleweed dispersal – with whole
426 plants (with fruits retained) breaking off and tumbling during the AMT dry season or rafting on
427 water during the wet season. Our recent observations of *S. melanospermum* in the Northern
428 Territory’s highly-dynamic Robinson and MacArthur River systems suggest that this species
429 moves about in a similar fashion, even though it is also an obviously reflexed-calyx fruit-
430 presenter (Fig 1-D). Shrubs holding onto hundreds of ripened and post-mature fruits have been
431 found tossed along the deep sandy banks of these rivers like so much flotsam and jetsam. Water
432 might likewise disperse some fruits that mature as bony and hollow, particularly when enclosed
433 in calyces providing additional surface area among the prickles. We have observed the enclosed
434 fruits of *S. ossicruentum*, *S. carduiforme*, and *S. echinatum* piled up in debris lines deposited by
435 sheet wash from the previous season’s monsoon rains, a phenomenon not uncommon in arid
436 zones [41-43].

437 Still, the broad observation that we have made is that, regardless of the “intended” mode
438 of dispersal, precious few fruits appear to be carried off (whether internally or externally) by a
439 biotic disperser and have thus become dependent on (seemingly) secondary abiotic forms of
440 dispersal. Our work growing thousands of plants from seed in cultivation has shown that
441 untreated seeds are able to germinate (even though rates are much higher following gibberellic
442 acid treatments), suggesting that gut passage is not a dispersal requirement – even if it might
443 offer some obvious benefits. Martine and Anderson [9] proposed that AMT *Solanum* might be
444 engaged in a type of secondary dispersal by which macropods provide short-distance seed
445 dispersal through a three-step process: fruit ingestion, “fecal seed storage” (after seeds are moved
446 with animals to their denning sites and defecated), and “seasonal redispersal” via wet season

447 rains. The general concept still applies to plant species producing trample burrs and may be even
448 more likely than a scenario dependent on ingestion, given the hypothesized prevalence of
449 epizoochory in the group.

450 The notion that the extinction of many of Australia's medium- to large-bodied vertebrates
451 in the last 400,000 years [44] may have rendered some primary dispersal mechanisms
452 anachronistic is worth considering. One might imagine that the lack of current data related to
453 effective seed dispersal is a simple reflection of the dearth of potential dispersers in the AMT, a
454 factor made limiting by past megafaunal extinctions in Australia. However, no fossil record
455 exists for large-scale megafauna extinctions over much of what is now the AMT [45,46].
456 Meanwhile, fossil records provide ample evidence for extinctions in southwestern [47] and
457 southeastern [48] Australia; the assumed range of megafauna reached northward primarily along
458 the western and eastern coasts [47,48], perhaps because of inconsistent availability of water in
459 central and north-central regions.

460 In the areas of the continent where evidence is lacking for past megafauna abundance,
461 relatively few extant plant species produce fruits fitting the typical profile of the megafauna
462 dispersal system (large size, tough endocarp, and soft pulp [49]). Over much of interior and
463 northern Australia, Pleistocene seed dispersal may have been left to macropods and other
464 browsers (both extinct and extant), who may have been infrequently and opportunistically
465 frugivorous – but who, largely, may have carried fruits/seeds in their fur. Trample burr
466 fruit/calyx morphology may thus have been an advantage in an ecosystem lacking abundant large
467 frugivores – and continues to be an advantage today.

468 Even where the megafauna syndrome is present in Australia it appears to not parallel the
469 large-fruited big-tree version of Janzen and Martin [49]. Instead, plants exhibit a suite of growth

470 forms that would have allowed shrubbier plants to resist browsing by large herbivorous animals
471 – including physical and chemical defenses – while also producing fruits and seeds adapted for
472 vertebrate dispersal [46], including trample burrs. In the present day, many of the remaining
473 small- and medium-sized mammals (as well as frugivorous birds) throughout the AMT have
474 recently experienced and continue to face declines [50-52]. This likely places many biotically-
475 dispersed plants at risk and may put those species with potential secondary mechanisms (e.g.
476 shakers or tumbleweeds) at a short-term advantage. In the long-run, however, local extirpations
477 and wholesale extinctions of biotic dispersers are likely to have consequences for the plants
478 linked to them. The observations of the Martu people [11] and R. Bird (pers.comm) at the
479 boundary of the Little Sandy and Great Sandy Deserts are seemingly telling in this regard: one
480 confirmed *Solanum* frugivore, the burrowing bettong (*Bettongia leseuer*), is now locally extinct;
481 and the other, the likely-opportunistic hill kangaroo (*Macropus robustus*), handles the fruits of
482 two species in such a way that the bitter seeds are avoided entirely. Demographic analyses
483 (Cantley, et al., in prep) infer that several AMT *Solanum* species have undergone recent genetic
484 bottlenecks, a trend that would likely continue should their biotic dispersers also go the way of
485 the burrowing bettong.

486 The evolutionary development of calyx morphology is one of many exciting areas of
487 current research in Solanaceae [53,54], especially in the context of seed dispersal biology.
488 Among our *Solanum* study species, calyx form exhibits particular lability -- with broad
489 occurrence of derived accrescence/enclosure plus a reversal to the *Solanum* ancestral state of a
490 non-enveloping calyx associated with fleshy berries. Further research employing larger
491 molecular datasets (McDonnell and Martine, in prep) will shed greater light on the evolution of

492 this and other fascinating elements associated with the reproductive biology of AMT *Solanum*
493 species.

494
495

496 **Acknowledgements**

497 This work would not have been possible without generous assistance from staff of the Northern
498 Territory Herbarium, Palmerston (DNA) and the George Safford Torrey Herbarium, University
499 of Connecticut (CONN), as well as assistance with collecting permits from Parks Australia,
500 Northern Territory Parks and Wildlife Commission, and Western Australia Department of Parks
501 and Wildlife. We gratefully acknowledge the Traditional Owners of the lands where many of our
502 collections were made. T. Caton, M. Garanich, M. Spiro, W. Boop and others provided
503 greenhouse support at Bucknell. Numerous individuals assisted with collection of specimens in
504 the field during trips between 2004-2018, including R. Martine, I. Martine, J. Martine, E.
505 Sullivan, D. Symon, P. Jobson, B. Barker, R. Barker, K. Brennan, E. Capaldi, D. Vogt, B.
506 Figley, G. Lionheart, and B. Lavoie. Thank you to X reviewers for their valuable comments on
507 earlier drafts of this paper.

508

509 **References**

510

- 511 1. Knapp S, Sagona E, Carbonell A, Chiarini F. A revision of the *Solanum elaeagnifolium*
512 clade (Elaeagnifolium clade; subgenus Leptostemonum, Solanaceae). *PhytoKeys*.
513 2017;84: 1-104. doi:10.3897/phytokeys.84.12695
- 514 2. Hunziker AT. *Genera Solanacearum: the genera of Solanaceae illustrated, arranged*
515 *according to a new system*. Königstein: Koeltz Scientific Books; 2001.

- 516 3. Symon DE. Fruit diversity and dispersal in *Solanum* in Australia. J Adel Bot Gard.
517 1979;1: 321–311. <https://www.jstor.org/stable/23872220>.
- 518 4. Martine CT, Anderson GJ, Les DH. Gender-bending aubergines: Molecular
519 phylogenetics of cryptically dioecious *Solanum* in Australia. Aust Syst Bot. 2009;22:
520 107-120. doi: 10.1071/SB07039.
- 521 5. Martine CT, Vanderpool D, Anderson GJ, Les DH. Phylogenetic relationships of
522 andromonoecious and dioecious Australian species of *Solanum* subgenus *Leptostemonum*
523 section *Melongena*: inferences from ITS sequence data. Syst Bot. 2006;31: 410–420.
524 <https://www.jstor.org/stable/25064163>.
- 525 6. Bean AR. The taxonomy and ecology of *Solanum* subg. *Leptostemonum* (Dunal) Bitter
526 (Solanaceae) in Queensland and far north-eastern New South Wales, Australia.
527 Austrobaileya. 2004;6: 639-816. <http://www.jstor.org/stable/41739063>.
- 528 7. Whalen MD Conspectus of species groups in *Solanum* subgenus *Leptostemonum*. Gentes
529 Herbarum. 1984;12: 179–282.
- 530 8. Anderson GJ, Symon DE. Functional dioecy and andromonoecy in *Solanum*. Evolution.
531 1989;43: 204-219. doi: 10.1111/j.1558-5646.1989.tb04218.x.
- 532 9. Martine CT, Anderson GJ. Dioecy, pollination, and seed dispersal in Australian spiny
533 *Solanum*. Acta Hortic. 2007;745: 269-283.
- 534 10. Lacey LM, Capaldi E, Jordon-Thaden I, Martine CT. Exploring the potential for *Solanum*
535 fruit ingestion and seed dispersal by rock-dwelling mammals in the Australian Monsoon
536 Tropics. Poster presented at: Botany Conference; 2015 Jul 25-29; Edmonton, Alberta.
537 <http://2015.botanyconference.org/engine/search/index.php?func=detail&aid=408>.

- 538 11. Coddling BF, Bliege Bird R, Kauhanen PG, Bird DW. Conservation or co-evolution?
539 Intermediate levels of aboriginal burning and hunting have positive effects on kangaroo
540 populations in Western Australia. *Hum Ecol.* 2014;42: 659–669. doi: 10.1007/s10745-
541 014-9682-4.
- 542 12. Telfer WR, Garde MJ. Indigenous knowledge of rock kangaroo ecology in western
543 Arnhem Land, Australia. *Hum Ecol.* 2006;34: 379-406. [https://doi.org/10.1007/s10745-](https://doi.org/10.1007/s10745-006-9023-3)
544 006-9023-3
- 545 13. Martine CT, Boni AJ, Capaldi EA, Lionheart G, Jordon-Thaden IE. Evidence of rock-
546 dwelling macropod seed dispersal in a tropical monsoon community. *North Territ Nat.*
547 2016;27: 68-77.
- 548 14. Symon DE. A revision of genus *Solanum* in Australia. *J Adel Bot Gard.* 1981;4: 1–367.
- 549 15. Barrett RL. *Solanum zoeae* (Solanaceae), a new species of bush tomato from the North
550 Kimberley, Western Australia. *Nuytsia.* 2013;23: 5-21.
- 551 16. Bean AR. Two new species of *Solanum* (Solanaceae) from the Northern Territory,
552 Australia. *Austrobaileya* 2016;9: 524-533.
- 553 17. Bean AR, Albrecht DE. *Solanum succosum* A.R.Bean & Albr. (Solanaceae), a new
554 species allied to *S. chippendalei* Symon. *Austrobaileya.* 2008;7: 669-675.
555 <http://www.jstor.org/stable/41739087>.
- 556 18. Brennan K, Martine CT, Symon DE. *Solanum sejunctum* (Solanaceae), a new
557 functionally dioecious species from Kakadu National Park, Northern Territory, Australia.
558 The Beagle, Records of the Museums and Art Galleries of the Northern Territory.
559 2006;22: 1-7.

- 560 19. Lacey LM, Cantley JT, Martine CT. *Solanum jobsonii*, a novel andromonoecious bush
561 tomato from a new Australian national park. *PhytoKeys*. 2017;82: 1-13. doi:
562 10.3897/phytokeys.82.12106
- 563 20. Martine CT, Frawley ES, Cantley JC, Jordon-Thaden IE. *Solanum watneyi*, a new bush
564 tomato species from the Northern Territory, Australia named for Mark Watney of the
565 book and film “The Martian”. *PhytoKeys*. 2016a;61: 1-13. doi:
566 10.3897/phytokeys.61.6995
- 567 21. Martine CT, Cantley JC, Frawley ES, Butler A, Jordon-Thaden IE. New functionally
568 dioecious bush tomato from northwestern Australia, *Solanum ossicruentum*, may utilize
569 “trample burr” dispersal. *PhytoKeys*. 2016b;63: 19-29. doi: 10.3897/phytokeys.637743
- 570 22. Martine CT, Symon DE, Capaldi Evans E. A new cryptically dioecious species of bush
571 tomato (*Solanum*) from the Northern Territory, Australia. *PhytoKeys*. 2013;30: 23-31.
572 doi: 10.3897/phytokeys.30.6003
- 573 23. Chamala S, García N, Godden GT, Krishnakumar V, Jordon-Thaden IE, De Smet R, et
574 al. MarkerMiner 1.0: A new application for phylogenetic marker development using
575 angiosperm transcriptomes. *Applications in Plant Sciences*, 2015;3(4): 1400115.
576 doi:10.3732/apps.1400115
- 577 24. Untergasser A, Cutcutache I, Koressaa T, Ye J, Faircloth BC, et al. Primer3—new
578 capabilities and interfaces. *Nucleic Acids Res*. 2012;40(15): e115. doi:
579 10.1093/nar/gks596.
- 580 25. Uribe-Convers S, Settles ML, Tank DC. A Phylogenomic Approach Based on PCR
581 Target Enrichment and High Throughput Sequencing: Resolving the Diversity within the

- 582 South American Species of *Bartsia* L. (Orobanchaceae). PLoS ONE. 2016;11(2):
583 e0148203. doi:10.1371/journal.pone.0148203
- 584 26. Magoc T, Salzberg S. FLASH: Fast length adjustment of short reads to improve genome
585 assemblies. Bioinformatics. 2011;27: 2957-63. doi: 10.1093/bioinformatics/btr507.
- 586 27. Edgar R. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
587 throughput. Nucleic Acids Research. 2004;32: 1792–1797. doi: 10.1093/nar/gkh340.
- 588 28. Chernomor O, von Haeseler A, Minh BQ. Terrace aware data structure for phylogenomic
589 inference from supermatrices. Syst Biol. 2016;65: 997–1008. doi:
590 10.1093/sysbio/syw037.
- 591 29. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective
592 stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol.
593 2015;32(1): 268-274. doi:10.1093/molbev/msu300
- 594 30. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder:
595 fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14: 587–
596 589. doi: 10.1038/nmeth.4285.
- 597 31. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic
598 bootstrap. Mol Biol Evol. 2013;30: 1188–1195. doi: 10.1093/molbev/mst024.
- 599 32. Zhang C, Rabiee M, Sayyari E, Mirarab S. ASTRAL-III: polynomial time species tree
600 reconstruction from partially resolved gene trees. BMC Bioinformatics. 2018;19(6): 153.
601 doi:10.1186/s12859-018-2129-y
- 602 33. Altekar G, Dwarkadas A, Huelsenbeck JP, Ronquist F. Parallel Metropolis-coupled
603 Markov chain Monte Carlo for Bayesian phylogenetic inference. Bioinformatics.
604 2004;20: 407–415. doi: 10.1093/bioinformatics/btg427.

- 605 34. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogeny.
606 Bioinformatics. 2001;17: 754–755.
- 607 35. Ronquist F, Huelsenbeck JP. MRBAYES 3: Bayesian phylogenetic inference under
608 mixed models. Bioinformatics. 2003;19: 1572–1574.
- 609 36. Rambaut A, Suchard MA, Xie D, Drummond AJ. Tracer. Version 1.6. 2014. Available
610 from: <http://tree.bio.ed.ac.uk/software/tracer>.
- 611 37. Rambaut A, Drummond AJ. TreeAnnotator. Version 2.4.1. 2016. Available from:
612 <http://beast.bio.ed.ac.uk/treeannotator>.
- 613 38. Maddison WP, Maddison DR. Mesquite: a modular system for evolutionary analysis.
614 Version 3.51. 2018. Available from: <http://www.mesquiteproject.org>.
- 615 39. Anderson GJ, Symon DE. Insect foragers on *Solanum* flowers in Australia. Ann Mo Bot
616 Gard. 1988; 75: 842-852. doi: 10.2307/2409175
- 617 40. Bean AR. A taxonomic revision of the *Solanum echinatum* group (Solanaceae).
618 Phytotaxa. 2012;57: 33-50. doi:10.11646/phytotaxa.57.1.6.
- 619 41. Mott JJ, McComb AJ. Patterns in annual vegetation and microrelief in an arid region of
620 Western Australia. J Ecol. 1974;62: 115-26. doi: 10.2307/2258884.
- 621 42. Reichman JO. Spatial and Temporal Variation of Seed Distributions in Sonoran Desert
622 Soils. J Biogeogr. 1984;11: 1-11. doi:10.2307/2844771
- 623 43. Venable DL, Flores-Martinez A, Muller-Landau HC, Barron-Gafford G, Becerra JX.
624 Seed dispersal of desert annuals. Ecology. 2008;89(8): 2218-2227.
625 <https://www.jstor.org/stable/27650747>.

- 626 44. Wroe S, Field JH, Archer M, Grayson DK, Price GJ, Louys J, et al. Climate change
627 frames debate over the extinction of megafauna in Sahul (Pleistocene Australia - New
628 Guinea). PNAS; 110: 8777-8781. doi: 10.1073/pnas.1302698110.
- 629 45. Horton DR. Red kangaroos: last of the Australian megafauna. In: Martin PS, Klein RG,
630 eds. Quaternary extinctions: a prehistoric revolution. Tucson: University of Arizona
631 Press; 1984, pp. 639-690.
- 632 46. Johnson C. Australia's mammal extinctions: a 50,000 year history. Cambridge:
633 Cambridge University Press; 2006.
- 634 47. Merilees D. Comings and goings of late Quaternary mammals in extreme southwestern
635 Australia. In: Martin PS, Klein RG, eds. Quaternary extinctions: a prehistoric revolution.
636 Tucson: University of Arizona Press; 1984, pp 629-638.
- 637 48. Murray P. Extinctions downunder: a bestiary of extinct Australian late Pleistocene
638 monotremes and marsupials. In: Martin PS, Klein RG, eds. *Quaternary extinctions: a*
639 *prehistoric revolution*. Tucson: University of Arizona Press; 1984.
- 640 49. Janzen DH, Martin PS. Neotropical anachronisms: the fruits the gomphotheres ate.
641 Science. 1982;215: 19-27. doi: 10.1126/science.215.4528.19.
- 642 50. Woinarski J, Milne DJ, Wanganeen G. Changes in mammal populations in relatively
643 intact landscapes of Kakadu National Park, Northern Territory, Australia. Austral Ecol.
644 2001;26: 360-370. doi: 10.1046/j.1442-9993.2001.01121.x.
- 645 51. Woinarski J, Fisher A. Threatened terrestrial animals of Kakadu National Park: Which
646 species?; How are they faring? And what needs to be done for them? In: Winderlich S,
647 Woinarski J, eds. Kakadu National Park Landscape Symposia Series. Symposium 7:

648 Conservation of threatened species. Internal Report 623. Darwin: Supervising Scientist;
649 2013;93-104.

650 52. Ziembicki MR, Woinarski JCZ, Mackey B. Evaluating the status of species using
651 Indigenous knowledge: Novel evidence for major native mammal declines in northern
652 Australia. *Biol Conserv.* 2013;157: 78-92. doi: 10.1016/j.biocon.2012.07.004.

653 53. Deanna R, Larter MD, Barboza GE, Smith SD. Repeated evolution of a morphological
654 novelty: a phylogenetic analysis of inflated fruiting calyx in the Physalidae tribe
655 (Solanaceae). 2018; bioRxiv 425991; doi: 10.1101/425991.

656 54. Knapp S. Tobacco to tomatoes: a phylogenetic perspective on fruit diversity in the
657 Solanaceae. *J. Exp. Bot.* 2002;53: 2001-2022. doi: 10.1093/jxb/erf068.

658
659
660

661 **Supporting information**

662

663 **S1 Fig. Individual gene trees for each of the seven loci as estimated by IQ-TREE.** Values at
664 nodes reflect bootstrap support.

665

666 **S1 File. List of specimen vouchers and GenBank numbers for sequences used in this study.**

667 For each accession, information is as follows: collector and collector number, locality, date of
668 collection, acronym for herbarium (in parentheses) where voucher is held, and gene regions
669 recovered for that accession. GenBank accession numbers for gene regions used in this study are
670 listed for each taxon in the following order, with a dash (--) inserted where a region was not
671 recovered: SOL3, SOL8, SOL9, SOL12, SOL14, SOL15, SOL16.

672

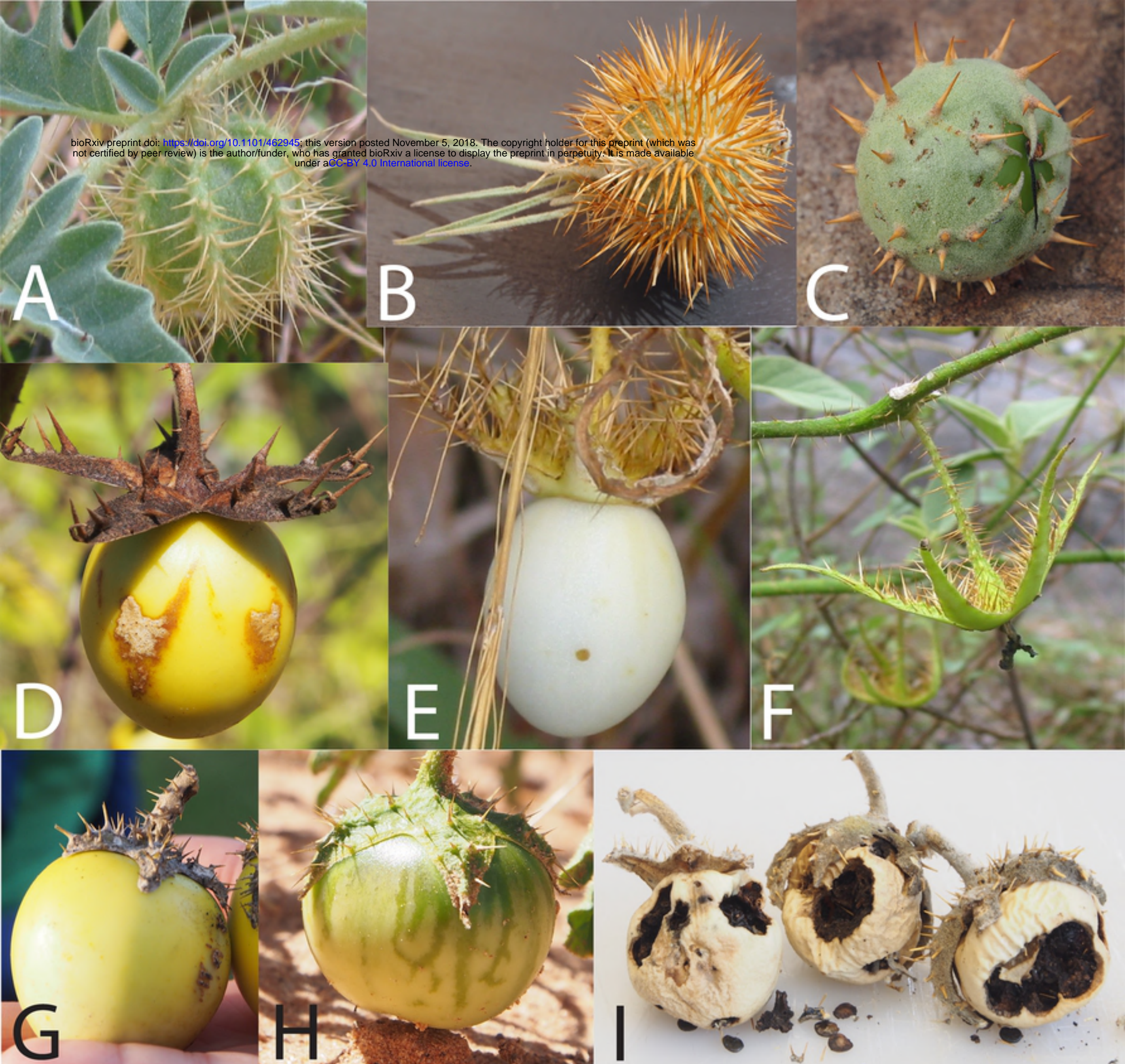


Fig 1. Photo plate

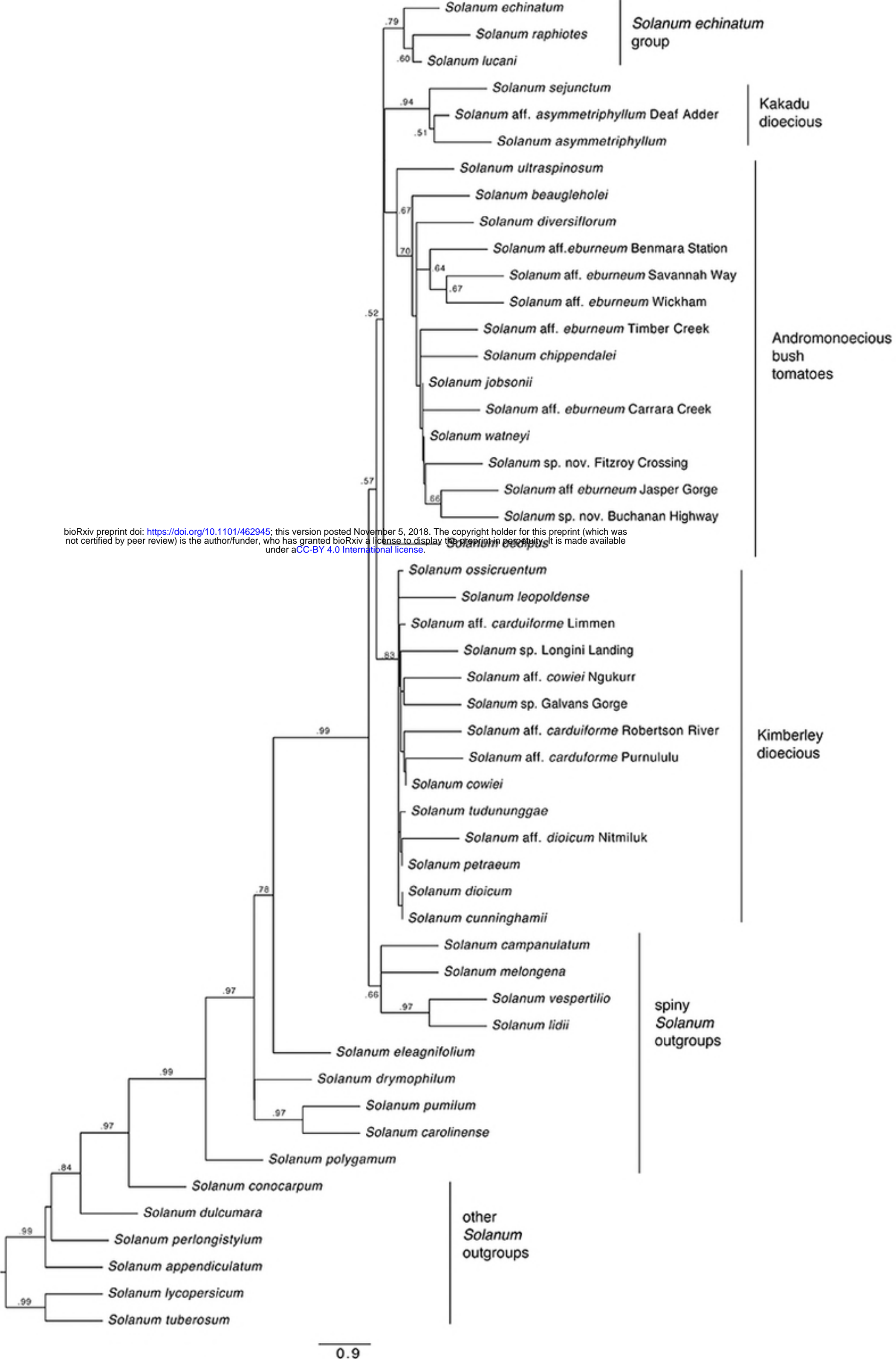


Fig. 2 Astral III species tree

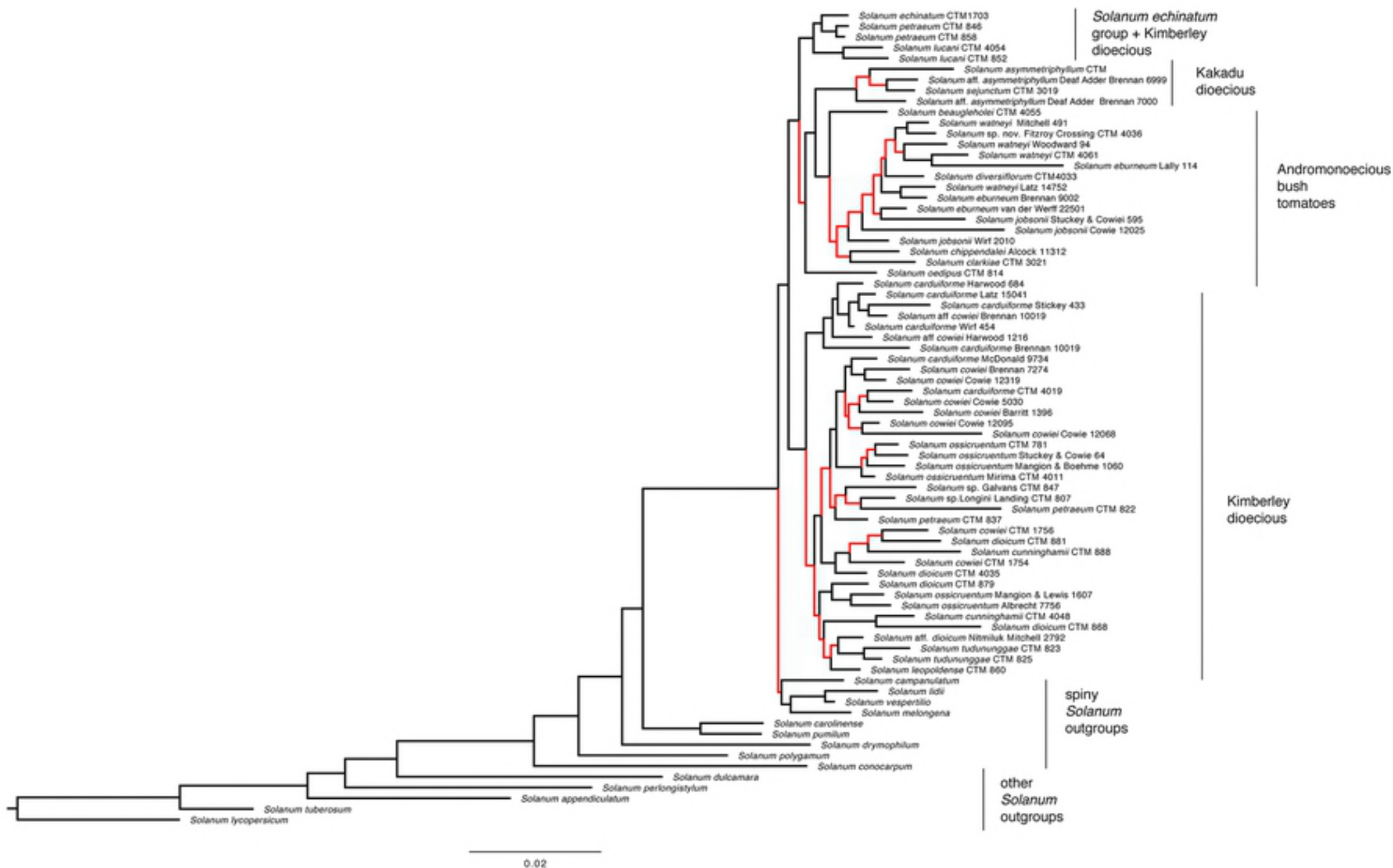


Fig. 3 Bayesian topology

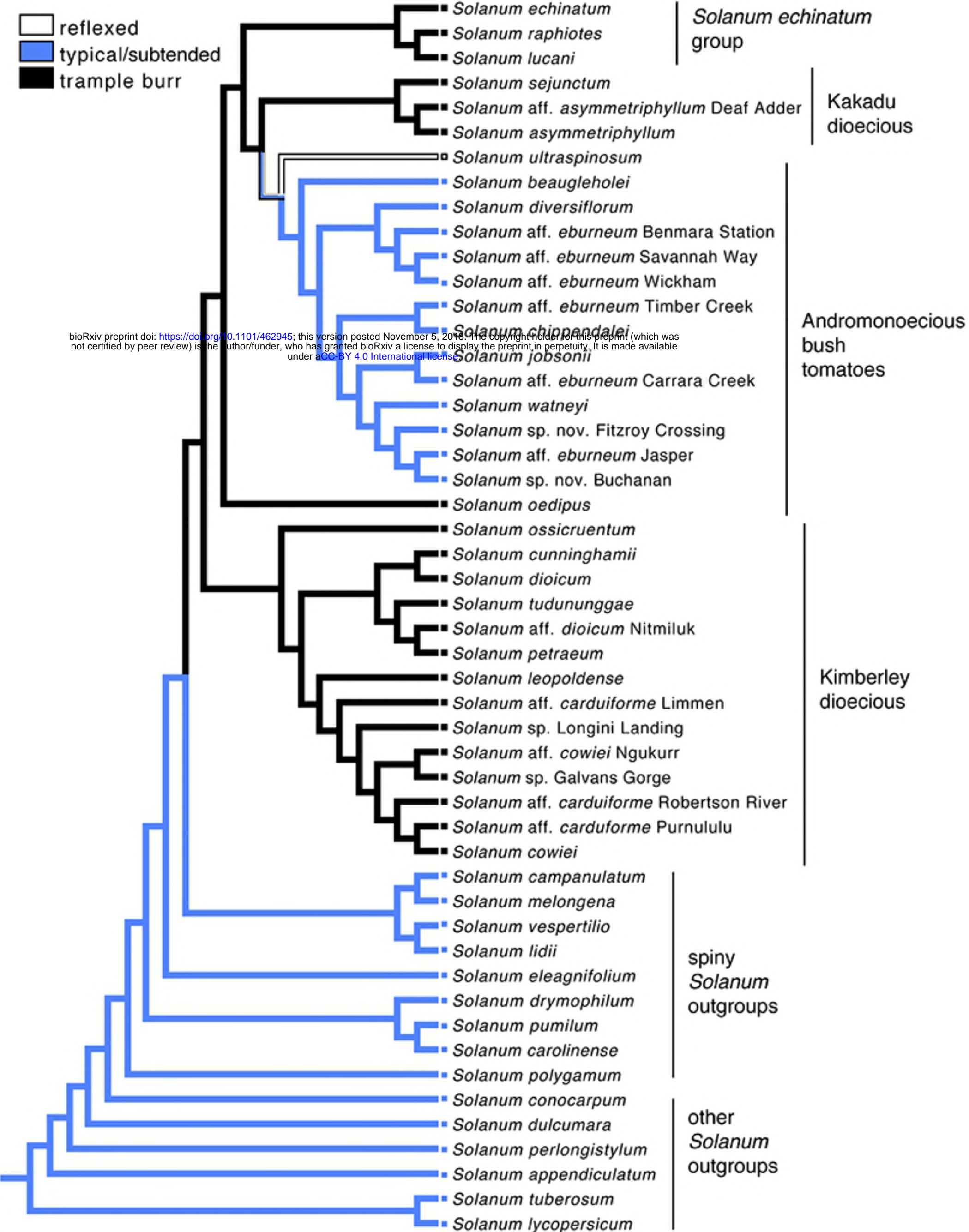


Fig. 4 Calyx reconstruction