Genome size, ploidy levels, and development of novel SSR primer to evaluate genetic diversity of *Corylopsis* Siebold & Zucc. germplasm collections

Jason D. Lattier · Harvey E. Ballard Jr. · Matthew Kramer · Margaret R. Pooler

Received: 26 September 2021 / Accepted: 6 March 2022 / Published online: 14 April 2022
This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply 2022

**Abstract** *Corylopsis* Siebold & Zucc. is a genus of flowering shrubs with few cultivars and little information on genome size, ploidy levels, and genetic relationships among taxa. The United States National Arboretum and collaborating gardens in North America maintain a diverse germplasm collection of *Corylopsis* including cultivated and wild-collected taxa. The objectives of this study were to determine genome size, ploidy, and genetic diversity in *Corylopsis*. Flow cytometry revealed a ploidy series of diploid, tetraploid, pentaploid, and hexaploid taxa in *Corylopsis* with 2C genome sizes ranging from 1.9 pg to 5.7 pg. Representative taxa were selected for root tip cytology, revealing diploid (*C. platypetala* Rehder & E.H. Wilson), tetraploid (*C. gotoana* Makino), pentaploid (unnamed hybrid selection), and hexaploid (*C. willmottiae* Rehder & E.H. Wilson) taxa. Seventeen genomic simple sequence repeat markers were developed and used to determine the genetic diversity of *Corylopsis*. These markers had an average of 7.4 alleles per locus with a range of three to 17 alleles. A dendrogram was generated by unweighted pair group method with arithmetic mean (UPGMA) cluster analysis using the Jaccard similarity coefficient to visualize genetic relationships and ploidy series within taxa. Information gained from this study will further breeding efforts and aid management of germplasm collections of *Corylopsis*.

**Keywords** Winter hazel · Ploidy · Genome size · Cytology · Microsatellite markers

**Introduction**

*Corylopsis* Siebold & Zucc. is a genus of woody ornamental shrubs in the Hamamelidaceae with a complex taxonomic history. Prior revisions have designated from seven to over 30 species, creating confusion for the nursery industry and public gardens (Bean and Aniško 2014; Hinkley 2009). Many herbarium vouchers and species descriptions of this complex group are incomplete (Morley 1977).
Corylopsis species are widely distributed from Northeast India across China and into Korea, Japan, and Taiwan, with species in the easternmost end of this distribution having the most distinct characters within the genus (Bean and Aniśko 2014; Morley and Chao 1977; Zhang et al. 2003). Morley and Chao (1977) recognize several species native to Japan, including C. pauciflora Siebold & Zucc. (with populations in Taiwan), C. spicata Siebold & Zucc., and C. glabrescens Franch. & Sav. (with populations in Southern Korea), C. gotoana Makino, and C. gotoana var. coreana (Uyeki) T. Yamaz. (Morley and Chao 1977; USDA 2021). Chinese species of Corylopsis are found in montane forests from Northeast India across China (Morley and Chao 1977; Bean and Aniśko 2014) and include C. henryi Hemsl., C. multiflora Hance, C. himalayana Griff. [C. himalayana var. himalayana and C. himalayana var. griffithii (Hemsl.) B.D. Morley & J.M. Chao], and C. sinensis Hemsl. [C. sinensis var. sinensis and C. sinensis var. calvescens Rehder & E.H. Wilson]. However, the latest treatment in the Flora of China recognizes 18 Chinese species (Zhang et al. 2003).

Corylopsis represents a group of ornamental shrubs with untapped breeding potential. Few improved cultivars have been developed, with most derived from chance nursery seedlings or field collections. However, diversity within Corylopsis offers opportunities for targeted breeding. Corylopsis has few pests and diseases, and most species perform best in moist soil under dappled afternoon shade (Bean and Aniško 2014). They vary from low, spreading vase-like forms to large upright plants (Bean and Aniško 2014), with bloom times from late winter to early spring. Species exhibit a range of cold hardiness. Corylopsis yunnanensis Diels is reported as the least cold hardy (Li 1962; Morley and Chao 1977) while some taxa (such as C. gotoana) are reported hardy in USDA Hardiness Zone 5 (Bean and Aniško 2014; Gapinski 2015). Flower colors vary from yellow-green in C. sinensis (and synonyms) to primrose yellow in species such as C. spicata (Bean and Aniško 2014). Anther color varies from green to red, with C. spicata ‘Red Eye’ selected for its compact growth and bright red anthers (Hop 2011). Inflorescences vary from 2 to 5 flowers on short racemes in C. pauciflora to chains of 12–20 flowers on long racemes in C. willmottiae Rehder & E.H. Wilson (Bean and Aniško 2014; Zhang et al. 2003). Flower fragrance is also variable across genotypes, ranging from unpleasant to a strong, sweet primrose fragrance (Bean and Aniško 2014). Foliage color offers further interest for Corylopsis breeding, as selections with blue-green or golden leaves have been reported (Bean and Aniško 2014). Corylopsis are also appreciated in the landscape for fall color and spring flush, providing several layers of seasonal interest (Bean and Aniško 2014; Hop 2011; Morley 1977).

Polyploidy is rare within the Hamamelidaceae, with only one other genus, Fothergilla Murr., reported to have a ploidy series (Ranney et al. 2007; Qi et al. 2015). Some variable morphology observed in Corylopsis may be due to the range of ploidy documented throughout the genus, which include diploid, tetraploid, pentaploid, and hexaploid taxa (Anderson and Sax 1935; Kim et al. 2008; Kono et al. 2008; Mehra and Khosla 1969; Oginuma 1991; Oginuma and Tobe 1991; Santamour 1965). In addition to ploidy, the position of the centromere has been examined to determine phylogenetic relationships in Corylopsis (Kono et al. 2008), and several molecular studies have provided further insight into relationships within Corylopsis (Li et al. 1997; Li 2008; Roh et al. 2007; Yamanaka et al. 2008); however, to date, microsatellites, or simple sequence repeats (SSRs) have not been developed specifically for Corylopsis.

The objectives of this study were to (1) determine genome size and ploidy among Corylopsis taxa and (2) develop novel microsatellite markers to assess genetic diversity of living collections to assist with collection management and facilitate breeding.

Materials and methods

Plant materials

Fresh expanded leaves were sampled from living collections at botanical gardens and arboreta in the United States (Table 1). Samples represent material that originally came from plant hunting expeditions, botanic gardens, and Index Semina.

Flow cytometry

Approximately 2 cm² samples of fresh leaf tissue were chopped by razor blade in a polystyrene petri dish containing 500 μl of nuclei extraction buffer
Table 1 Accession number, origin, mean 2C value, and ploidy of *Corylopsis* accessions used in this study

| Species                  | Accession number | Origin                                      | Mean 2C value | Ploidy |
|--------------------------|------------------|---------------------------------------------|---------------|--------|
| *C. glabrescens* Franch. & Sav | AA 1030-90*B     | Kirishima Kinkowan National Park, Kyushu, Japan | 2.0           | 2x     |
|                          | AA 113-77*A      | Garden origin, Washington Park Arboretum    | 4.6           | 5x     |
|                          | LW 2002-0086*A   | Dawes Arboretum, OH                         | 5.5           | 6x     |
|                          | MS 1986-186*B    | Garden origin, Kamigamo ES, FSERC Kyoto U., Japan | 5.4 | 6x     |
|                          | MS 1987-115*C    | Garden origin, Kamigamo ES, FSERC Kyoto U., Japan | 5.5 | 6x     |
|                          | MS 1987-116*B    | Garden origin, Kamigamo ES, FSERC Kyoto U., Japan | 3.8 | 4x     |
|                          | MS 1989-057*A    | Garden origin, Chollipo Arboretum, Korea    | 4.7           | 5x     |
|                          | MS 1998-185*A    | Garden origin, Chollipo Arboretum, Korea    | 4.0           | 4x     |
|                          | MT 340-2006*1    | Arrowhead Alpines, MI                       |               |        |
|                          | NA 37204-H       | Kingsville Nursery, MD                      | 5.4           | 6x     |
|                          | NA 50804-J       | Mt. Kirishima, Kyushu, Japan                | 5.4           | 6x     |
|                          | NA 60236-H       | Camellia Forest Nursery, NC                 | 1.9           | 2x     |
|                          | NA 62376-H       | Mt. Karakuni, Kyushu, Japan (31°56’N 130°51’E, 1070 m) | 1.9 | 2x     |
|                          | NA 65652         | Gruga Park Botanischer Garten, Germany       | 5.4           | 6x     |
|                          | NA 65868         | Smith College Botanic Garden, MA            | 5.4           | 6x     |
|                          | NA 70027-H       | 'Longwood Chimes'—Seedling of NA53233       | 1.9           | 2x     |
| *C. glandulifera* Hems l | AA 557-88*B      | Changhua, Lin’an County, China              | 3.6           | 4x     |
|                          | AA 557-88*C      | Changhua, Lin’an County, China              | –             | –      |
|                          | AA 898-81*A      | EH Wilson, Japan, 1919                      | 1.9           | 2x     |
|                          | JC 100356        | Dr. Charles Keith and Muki Fairchild, NC    | 5.4           | 6x     |
|                          | MS 2016-248*#    | Wild-collected. Distributed by USENA. (possibly NA 59,542) | 5.6 | 6x     |
|                          | NA 57391-H       | Changhua, Lin’an County, China. Mixed forest, exposed rock (800 m) | 5.4 | 6x     |
|                          | NA 57391-J       | Changhua, Lin’an County, China. Mixed forest, exposed rock (800 m) | 5.3 | 6x     |
|                          | NA 57391-01C1    | Changhua, Lin’an County, China. Mixed forest, exposed rock (800 m) | 5.4 | 6x     |
| *C. gotoana* Makino      | AA 388-2008*A    | Nakatsugawa, Gifu, Japan                    | 1.9           | 2x     |
|                          | AA 388-2008*B    | Nakatsugawa, Gifu, Japan                    | 1.9           | 2x     |
|                          | JC 060155        | Garden origin, Chollipo Arboretum, Korea    | 4.6           | 5x     |
|                          | JC 940450        | 'March Jewel'—Camellia Forest Nursery, NC    | 3.5           | 4x     |
|                          | NA 36588-CH      | Hillier & Sons, England                     | 3.6           | 4x     |
|                          | NA 36588-CJ      | Hillier & Sons, England                     | 3.6           | 4x     |
|                          | NA 36588-CL      | Unknown origin, Hillier Nursery, England    | 3.5           | 4x     |
|                          | NA 36588-H       | Hillier & Sons, England                     | 3.6           | 4x     |
|                          | NA 67499-H       | Camellia Forest Nursery, NC                 | 3.7           | 4x     |
| *C. gotoana* var. coreana (Uyeki) T. Yamaz | AA 248-2006*A | Near Songgwangsa Temple, Suncheon, Korea | 3.4           | 4x     |
|                          | AA 248-2006*B    | Near Songgwangsa Temple, Suncheon, Korea    | 3.5           | 4x     |
|                          | AA 248-2006*C    | Near Songgwangsa Temple, Suncheon, Korea    | 3.4           | 4x     |
|                          | MT 26-2001*1     | Mt. Chiri, North Jeolla, Korea              | 3.5           | 4x     |
|                          | MT 26-2001*2     | Mt. Chiri, North Jeolla, Korea              | 3.4           | 4x     |
|                          | MT 26-2001*3     | Mt. Chiri, North Jeolla, Korea              | 3.4           | 4x     |
|                          | NA 56375-005     | Sahmyook University, Korea; seed from Hongneung Arboretum | 5.4 | 6x     |
|                          | NA 56375-02C3    | Sahmyook University, Korea; seed from Hongneung Arboretum | 3.5 | 4x     |
| Species* | Accession number* | Origin | Mean 2C value | Ploidy |
|----------|------------------|--------|---------------|--------|
| C. himalayana Griff | NA 5951 | Wild origin. Obtained from Mallet Court Nursery, England | 2.1 | 2x |
| C. multiflora Hance | NA 83146 | Cistus Nursery, OR | 5.4 | 6x |
| C. pauciflora Seibold & Zucc | JC 050021 | Highland Creek Nursery, NC | 1.9 | 2x |
| C. platypetala Rehder & | LW 1957-0994*B | Scott Arboretum, PA | 3.7 | 4x |
| C. sinensis var calvescens Rehder & Wilson | LS 1961-686*A | Kingsville Nursery, MD | 3.6 | 4x |
| C. veitchiana Bean | JC 042064 | Seedling of JCRA XX0560 | 5.4 | 6x |
| C. glaucescens Hand.-Mazz. | LW 2001-0993*A | Stichting Botanische Tuin Kerkrade, Netherlands | 3.7 | 4x |
| C. sinensis var sinensis | AA 311-2004*B | Sandiequan River Valley, Lushan, Jiangxi, China | 5.4 | 6x |
| | AA 642-81*C | Huangshan Mountain, Anhui, China | 5.4 | 6x |
| | JC 150435-17 | Bill Barnes | 4.6 | 5x |
| | LW 1957-0906*D | Tingle Nursery Company, MD | 5.5 | 6x |
| | LW 1989-0368*A | Winterthur, DE | 2.0 | 2x |
| | MS 1992-426*A | Garden origin. Kamigamo ES, FSERC Kyoto U., Japan | 5.3 | 6x |
| | MS 1995-081*B | Spring Purple’—Hillier & Sons, England | 2.0 | 2x |
| | MT 338-2000*1 | Botanischer Garten der Technischen Universität Dresden, Germany | 5.3 | 6x |
| | NA 49113-H | Shicao River, Shennongjia, Hubei (1150 m) | 5.6 | 6x |
| | NA 49113-J | Shicao River, Shennongjia, Hubei (1150 m) | 5.5 | 6x |
| | NA 49113-P | Shicao River, Shennongjia, Hubei (1150 m) | 5.6 | 6x |
| Species | Accession number | Origin | Mean 2C value | Ploidy |
|---------|-----------------|--------|---------------|--------|
| NA 49113-T | Shicao River, Shennongjia, Hubei, China (1150 m) | 5.6 | 6x |
| NA 49113-V | Shicao River, Shennongjia, Hubei, China (1150 m) | 5.5 | 6x |
| NA 49113-X | Shicao River, Shennongjia, Hubei, China (1150 m) | 5.5 | 6x |
| NA 60237-H | Camellia Forest Nursery, NC | 5.4 | 6x |
| NA 62528-H | Wild-origin. Lushan Botanic Garden, Jiangxi Province, China | 5.2 | 6x |
| NA 65594 | Arboretum SGGW, Poland | 5.3 | 6x |
| NA 70116 | Cultuurtuin voor tech gewassen, Tech Univer. Delft, Netherlands | 5.3 | 6x |
| NA 70116-J | Cultuurtuin voor tech gewassen, Tech Univer Delft, Netherlands | 5.5 | 6x |
| C. spicata | JC 042022 | 'Golden Spring'—Nurseries Caroliniana, SC | 5.3 | 6x |
| Seibold & Zucc | JC XX0560 | US National Arboretum, DC | 5.2 | 6x |
| MS 1932-1648*A | Morris Arboretum, PA | 5.3 | 6x |
| MT 130-2008*1 | Hokigamine Forest, Kochi, Japan | 5.3 | 6x |
| NA 33335-J | Kingsville Nursery, MD | 5.4 | 6x |
| NA 33335-SH | Kingsville Nursery, MD | 5.5 | 6x |
| NA 36592-CH | Hillier & Sons, England | 5.4 | 6x |
| NA 37900-H | Gulfstream Nursery, VA | 5.3 | 6x |
| NA 37900-L | Gulfstream Nursery, VA | 5.4 | 6x |
| NA 70117 | Cultuurtuin voor tech gewassen, Tech Univer. Delft, Netherlands | 5.3 | 6x |
| NA 77924 | Unknown origin (possibly NA40102) | 5.3 | 6x |
| NA 79490 | 'Golden Spring'—Gifu Prefecture, Japan (via Highland Crk Nrsy, NC) | 5.3 | 6x |
| NA 83145 | Garden origin, spontaneous seedling, JC Ralston Arboretum, NC | 5.4 | 6x |
| C. willmottiae | JC 920667 | Wells-Medina Nursery, WA | 5.6 | 6x |
| Rehder & | LW 1966-0037*A | Scott Arboretum, PA | 5.7 | 6x |
| E.H. Wilson | LW 2016-2002*A | Wild-origin. Arnold Arboretum, MA | 2.0 | 2x |
| NA 67192-H | Niba Mountain, Sichuan, China (1760 m) | 4.0 | 4x |
| NA 69317 | Cultuurtuin voor tech gewassen, Tech Univer. Delft, Netherlands | 5.4 | 6x |
| NA 83147 | Plant Delights Nursery, NC. (=Longwood 1966-0037*A) | 5.7 | 6x |
| NA 85065 | Songpan County, Sichuan, China (2171 m) | 2.0 | 2x |
| C. hybrid | NA 6-10 EH | Breeding program cross: 'March Jewel' x C. spicata | 4.7 | 5x |

*Species and authorities are based on taxonomy designated by the USDA Germplasm Resources Information Network (GRIN; USDA 2021), except for C. glaucescens, which does not appear in GRIN taxonomy, so is based on checklists by Hassler (2021)

*Accessions in bold are documented wild-collected

*Accession prefix indicates where samples originated: AA Arnold Arboretum, Boston, MA; JCJC Raulston Arboretum, Raleigh NC, MSMorris Arboretum, Philadelphia, PA; MTMorton Arboretum, Lisle, IL; LWLongwood Gardens, Kennett Square, PA; NA US National Arboretum, Washington, DC

*Only 106 of these accessions were used for diversity analysis (Fig. 2), based on repeatable and clear SSR data
(Cystain PI Absolute P, Sysmex Partec GmbH, Görlitz, Germany). Included with each Corylopsis sample was chopped leaf tissue from the internal size standard *Pisum sativum* L. ‘Minerva Maple’ (2C = 9.56 pg, Johnston et al. 1999). The *Corylopsis-Pisum* nuclei suspension was passed through a 30-μm gauze filter (Sysmex Partec), and 2 ml of staining solution was added to the filtrate (CyStain PI Absolute P, Sysmex Partec). The stained nuclei suspension was processed using a flow cytometer (Cyflow Ploidy Analyzer, Sysmex Partec) to a minimum of 3000 nuclei per sample. For each accession, three samples were combined to represent a random sample of nuclei from each plant. Terminology describing genome size and ploidy levels follows Greilhuber et al. (2005): Relative 2C genome size = (DNA content of size standard) × (mean fluorescence value of sample)/(mean fluorescence value of standard). The monoploid (1C) value was calculated as the 2C value divided by the ploidy.

Cytology

Early and late season cuttings were taken from current year’s growth and rooted using 1000 ppm K-IBA (Sigma-Aldrich, St. Louis, MO) to produce root tips for cytology. Based on initial flow cytometry results, four taxa were chosen to represent each genome size group, corresponding to putative ploidy levels of diploid, tetraploid, pentaploid, and hexaploid. Cuttings were maintained under mist in a polyhouse in Beltsville, MD. Samples included *C. platypetala* (NA66131), *C. gotoana* var. *coreana* (NA56375-02C3), *C. willmottiae* (NA69317), and the *F*₁ interspecific hybrid *C. gotoana* ‘March Jewel’ × *C. spicata* (NA6-10EH). Root tips were digested and chromosomes were stained and spread following protocols of Lattier et al. (2017). Chromosomes were counted using a light microscope (Nikon Eclipse 50i; Tokyo, Japan). A minimum of 15 highly resolved cells were captured per taxon. Photomicrographs were prepared by focus stacking and blending multiple images to produce high-resolution chromosome images (Photoshop CC 2014; Adobe, San Jose, CA).

Microsatellite identification and primer design

Genomic DNA was extracted from *C. himalayana* (NA78641) using the methods described in the next section, and was subjected to shotgun genome sequencing on a Roche 454 high-throughput sequencer at the Ohio State University Plant–Microbe Genomics Facility. The resulting 65,000 reads were then analyzed using QDD-2 software (Meglécz et al. 2010) following the approach of Takayama et al. (2011) to identify potential microsatellites, build consensus sequences, and design optimal primer pairs. In using the software, all defaults were maintained. From the 1000+ optimal primer pairs identified by QDD software, we selected 100 for testing on *Corylopsis*, focusing especially on loci with 3 or longer repeat motifs as these have been shown to produce well separated peaks and less ambiguous scoring in an automated sequencing environment in comparison with dinucleotide repeat motifs. PCR primers were manufactured by Integrated DNA Technologies (Corvalle, IA). The 17 primers used for genetic diversity analysis (Table 2) contained the additional M13 (−21) universal sequence (TGTTAAACGACGGCC AGT) on the 5’ end of the forward primer to allow indirect fluorescent labeling of PCR products using a FAM (6-carboxyfluorescein)–labeled M13 primer (Schuelke 2000).

DNA extraction and microsatellite amplification

DNA was extracted from fresh leaf tissue using a FastPrep homogenizer (Savant Instruments, Farmingdale, NY) and DNeasy extraction and PowerClean Cleanup kits (Qiagen GmbH, Germany). DNA samples were quantified with a spectrophotometer (NanoDrop 1000 Spectrophotometer, Thermo Fisher, Wilmington, DE). Primers were screened for optimal annealing temperature and amplification quality by gradient PCR using two Corylopsis accessions—tetraploid *C. glandulifera* (AA557-88*B) and diploid *C. glabrescens* (AA1030-90*B) (Table 1). PCR profiles consisted of denaturation at 94 °C for 5 min followed by 30 cycles of 94 °C for 30 s, optimized annealing temperature (Table 2) for 45 s, and extension at 72 °C for 45 s. Next, eight cycles were run at 94 °C for 30 s, 53 °C for 45 s, and 72 °C for 45 s, followed by a final extension at 72 °C for 10 min. For analysis, 1 µl of PCR product was combined with 10 µl of formamide (Applied Biosystems, Foster City, CA) and 0.2 µL of GeneScan 500 LIZ Size Standard (Applied Biosystems), and samples were analyzed.
on a 3730xl DNA Analyzer (Applied Biosystems). A minimum of 10 repeated samples (positive controls) were run across each PCR plate. Two samples of nuclease-free water (negative control) were run with each PCR plate, and any samples that yielded null alleles were run twice. Allele scoring was performed with GeneMarker Version 2.6.3 (SoftGenetics, State College, PA).

Table 2  Characteristics of the 17 Corylopsis genomic microsatellite loci used in the diversity analysis of 106 taxa

| Primer | Repeat motif | Sequence (5’–3’) | Ta (°C) | Allele size range (bp) | Total no. of alleles | PIC | H_o | H_e |
|--------|--------------|------------------|--------|------------------------|---------------------|-----|-----|-----|
| P1     | AGCT         | FAM-TCCTTAGTTGCAGTACC R-GGGATGACAGATAAATAACATGTAGTAAAGCC | 57     | 89–103                 | 3                   | 0.2022 | 0.0943 | 0.2172 |
| P2     | AAC          | FAM-GTCAAAAGACAGTATTGGGCA R-AGTTCTGCTCGAAACCCTTC | 57     | 147–202               | 14                  | 0.8530 | 0.8679 | 0.8667 |
| P3     | AGCT         | FAM-GTGCAATGGCGGCTACTAAC R-GAGCCCTTCTGATCCGTTTTG | 57     | 119–132               | 3                   | 0.6310 | 0.6698 | 0.6768 |
| P4     | ACAT         | FAM-GAAAGAGGACGGCTTGGGT R-TTTCCTCAAGAAATGTTGGGA | 57     | 142–161               | 5                   | 0.3869 | 0.0377 | 0.4203 |
| P21    | AAG          | FAM-CATGAAATGGTAGCTTGGGAG R-GGTGAACATCTTTGATGCTTTG | 57     | 107–158               | 17                  | 0.8319 | 0.8962 | 0.8472 |
| P27    | AGT          | FAM-TGCGAGACTATGTGGCAGTT R-AAAAAGAGACCCAAATGTAAGA | 57     | 158–206               | 12                  | 0.7024 | 0.6415 | 0.7270 |
| P30    | AAG          | FAM-ACCTTAACTGTAAGGAACCTCTTTGACACCA | 57     | 133–169               | 9                   | 0.5688 | 0.5943 | 0.6200 |
| P31    | AAGGAG       | FAM-ATGAGAAGCCGCTTACAGCAG R-TTGATGGAAAGCAGATTGTGCCTTT | 59     | 178–226               | 9                   | 0.7551 | 0.7925 | 0.7823 |
| P32    | AAAAT        | FAM-AAATGCTAAATTCTTGTCAATG T-AGCCAAAGCCCTTTCCTT | 59     | 169–188               | 4                   | 0.4143 | 0.2547 | 0.4432 |
| P35    | AAAT         | FAM-TGAAAGAGGCGCTCAAGAGG R-CTGTGCTGTTGATGGGCGCA | 59     | 171–187               | 5                   | 0.6186 | 0.5566 | 0.6564 |
| P40    | ACG          | FAM-AAGGCATGGAAGTCGCATAAA R-ATTTGCTGTAAGCTCAAGAAG | 57     | 133–161               | 8                   | 0.5312 | 0.5094 | 0.5719 |
| P54    | AAG          | FAM-TAGGCCAGATCGCTTCTATG R-ACCTACGCCACCTCATG | 57     | 152–158               | 3                   | 0.1969 | 0.0660 | 0.2214 |
| P60    | ACC          | FAM-GGCCGACACTCTTATAATGAA R-AGCGCGGCACCTCTCCTC | 57     | 155–183               | 7                   | 0.7193 | 0.2075 | 0.7582 |
| P62    | ACC          | FAM-TCTGTCATACCTTCCAGCTCC R-GGAAATGTCAGGCGGTATG | 57     | 133–155               | 8                   | 0.6273 | 0.9340 | 0.6815 |
| P64    | AAG          | FAM-CCCTCACATAGCCGGTTGAG R-GAGGAGAAGCCCAAAGGAA | 57     | 153–180               | 8                   | 0.7421 | 0.8208 | 0.7747 |
| P65    | CCG          | FAM-TGCGGTCTCTCTCTATCTTCC R-ACTGTGCTGCCCATGAAATG | 57     | 180–195               | 5                   | 0.4350 | 0.3208 | 0.4622 |
| P66    | ACC          | FAM-CCGCGGTGAAATGTTGATA R-GGCAATGAAAGTTGACG | 57     | 137–157               | 3                   | 0.3557 | 0.0472 | 0.4628 |

T_a  Annealing temperature; PIC Polymorphic information content; H_o  Observed heterozygosity; H_e  Expected heterozygosity
Data analysis

Allele scores were converted to a binary matrix with “0” for absence of an allele and “1” for presence of an allele. The R ‘pvclust’ package, version 1.3-2, (Suzuki et al. 2017) was used to construct the dendrogram, with genetic distance based on the Jaccard similarity coefficient (“binary”) and clustering using the ‘unweighted pair group method with arithmetic mean’ (UPGMA) algorithm. Confidence levels for branches of the dendrogram were calculated using approximately unbiased (AU) p-values with multiscale bootstrap resampling based on 20,000 replications, as implemented in the R ‘pvclust’ package (Suzuki et al. 2017). Primer statistics (PIC, Ho, and He) were calculated according to standard algorithms (Botstein et al. 1980; Nei 1987) coded in the R language.

Results

Cytology

Root tip cytology of a single sample of *C. platypetala* (NA66131) confirmed that it was diploid with 24 chromosomes (Fig. 1A) contributing to its 2C genome size of 1.9 pg (Table 1). Similarly, root tip cytology of a sample of *C. gotoana* var. *coreana* (NA56375-02C3) confirmed that it was tetraploid with 48 chromosomes (Fig. 1B) contributing to its 2C genome size of 3.5 pg. An interploid hybrid (NA6-10EH) from the USNA breeding program was included in the study,

![Photomicrographs of Corylopsis metaphase chromosomes viewed at 1000× magnification. Scalebar at 0.01 µm. A C. platypetala (NA66131) 2n=2x=24; B C. gotoana (NA56375-02C3) 2n=4x=48 C F1 hybrid NA 6-10EH (C. gotoana ‘March Jewel’× C. spicata) 2n=5x=60; D C. willmottiae (NA69317) 2n=6x=72](image)
representing a hybrid between the tetraploid _C. gotoana_ ‘March Jewel’ and hexaploid _C. spicata_. Root tip chromosome counts confirmed that this selection is indeed pentaploid (2n = 5x = 60), as predicted based on the interploid cross (Fig. 1C). Root tip chromosome counts also confirmed a hexaploid sample, _C. willmottiae_ (NA69317), with 2n = 6x = 72 (Fig. 1D).

Flow cytometry

Flow cytometry revealed relative 2C genome sizes ranging from 1.9 to 5.7 pg (Table 1). Based on confirmed chromosome counts of taxa representing the entire ploidy series (2x, 4x, 5x, 6x; Fig. 1), we could confidently assign ploidy to all accessions tested based on the genome size determined by flow cytometry data (Table 1). While ploidy differences between species were expected based on previous reports, we found that several species had varying ploidy among accessions, including _C. glabrescens_ (2x, 4x, 5x, 6x), _C. glandulifera_ (2x, 4x, 5x, 6x), _C. gotoana_ (2x, 4x, 5x, 6x), _C. gotoana_ var. _coreana_ (4x, 6x), _C. platypetala_ (2x, 4x), _C. sinensis_ (2x, 5x, 6x), and _C. willmottiae_ (2x, 4x, 6x).

Microsatellite analysis

After we screened microsatellite primers developed for other genera in the Hamamelidaceae (Hatmaker et al. 2015) and found inconsistent results in _Corylopsis_, we developed primers specifically for _Corylopsis_. Of the original 100 primer pairs tested, we chose 17 that produced consistent, polymorphic, easily scorable amplification products across all species tested. These loci amplified an average of 7.3 alleles (range 3–17) across all species tested (Table 2). These 17 microsatellite primer pairs generated 120 scorable fragments ranging in size from 89 to 226 bp across the 106 _Corylopsis_ accessions tested. Polymorphic Information Content (PIC) ranged from 0.1969 for P54 to 0.8530 for P2; observed heterozygosity (H_o) ranged from 0.0660 for P54 to 0.9340 for P62; and expected heterozygosity (H_e) ranged from 0.217 for P1 to 0.866 for P2 (Table 2).

UPGMA cluster analysis revealed high levels of genetic diversity among accessions, even among some accessions collected from the same area (Fig. 2). Clustering of accessions roughly followed species designations, although exceptions were common, and likely represented misidentification or possibly undetected hybridizations. While some clusters were supported by high bootstrap values (> 80%, Fig. 2), some of the lower-order clusters were not robust.

Discussion

Ploidy

Our finding of ploidy series within and between species of _Corylopsis_ supports prior studies, but also represents previously unreported ploidy levels in some species. In _C. glabrescens_, we found four diploid samples, including the cultivar ‘Longwood Chimes’ (Table 1), consistent with previous reports of diploids in this species by Kono et al. (2008), who investigated somatic chromosomes in meristematic cells of young leaves. Two samples of _C. glabrescens_ (MS1987-116*B and MS1998-185*A) were identified as tetraploids, consistent with a previous report by Santamour (1965), based on cytological studies. Pentaploid and hexaploid _C. glabrescens_ were also revealed with our study, representing the first report of pentaploids and hexaploids in this species.

Two samples of _C. gotoana_ were diploid with a genome size of 1.9 pg, representing the first report of a diploid. One sample was pentaploid (JC060155), representing the first report of a pentaploid. Most samples were tetraploids. This is consistent with reports of tetraploids in _C. gotoana_ found in Japan and Korea (Kim et al. 2008; Kono et al. 2008). Likewise, most samples of _C. gotoana_ var. _coreana_ were also tetraploid, except for one hexaploid accession (NA56375-005). This is the first reported assessment of ploidy in _C. gotoana_ var. _coreana_.

One sample of _C. glandulifera_ (AA898-81*A) was found to be diploid with a genome size of 1.9 pg, while all other samples were hexaploid (Table 1). This is the first reported assessment of ploidy in _C. glandulifera_. The single sample of _C. himalayana_ included in our study was diploid with a genome size of 2.1 pg, consistent with previous cytological studies (Mehra and Khosla 1969). The two samples of _C. multiflora_ were both hexaploid, representing the first assessment of ploidy for this species.

All samples of _C. pauciflora_ were diploids, ranging from 1.9 to 2.1 pg, confirming previous reports
based on cytology (Anderson and Sax 1935; Ogimura and Tobe 1991; Santamour 1965). Corylopsis × ‘Winterthur’ was also diploid, which does not support a previous report of its origin as an interploid hybrid of C. spicata (6x) and C. pauciflora (2x) (Bean and Anisko 2014).

We found one diploid C. platypetala, confirming a report by Santamour based on cytology (1965). However, two samples of C. platypetala (LWL-2769*A and LW1957-0994*B) were tetraploids, which represents the first evidence of polyplody within this species. Two samples of C. sinensis were found to be diploid including one cultivar (‘Spring Purple’). These results confirm a previous report of diploid individuals of C. sinensis (Santamour 1965). However, except for one pentaploid accession (JC150435-17), and two tetraploids in C. sinensis var. calvescens, all the other C. sinensis accessions we tested were hexaploid (Table 1). This represents the first report of polyploids in C. sinensis.

Two samples of C. willmottiae (LW2016-2002*A and NA85065) were diploid with a genome size of 2.0 pg, conflicting with a previous report of tetraploid C. willmottiae based on cytology (Santamour 1965). This represents the first report of diploid individuals in C. willmottiae. One sample of C. willmottiae (NA67192-H) was found to be tetraploid, and four samples were hexaploid (Table 1). All samples of C. veitchiana and C. spicata that we tested were hexaploid. Similar reports have come from studies of the collections of the Arnold Arboretum (Anderson and Sax 1935; Ogimura 1991) and Morris Arboretum (Santamour 1965). Hexaploid C. spicata have also been confirmed in cultivated material from Kōchi, Shikoku, Japan (Oginuma 1991).

Although several studies exist for chromosome counts in Corylopsis, only one other study has reported on genome size. Using propidium iodide-based flow cytometry, Garcia et al. (2010) reported a genome size for C. coreana of 2C = 3.93 pg (using Pisum L as an internal standard) and for C. sinensis of 2C = 4.62 pg (using Petunia Juss. as an internal standard). Although genome size and ploidy level were not calibrated using cytology (chromosome counts), it appears that these 2C genome sizes are larger than those calculated in our study. In prior research on other ornamental taxa, a combination of genome size estimates (from flow cytometry) and chromosome counts have revealed substantial species-level variation in 1Cx values (Gilloy and Ranney 2015; Parris et al. 2010; Ranney et al. 2018; Shearer and Ranney 2013). Regardless of the absolute genome size, it is evident from our study that the range of ploidy levels in Corylopsis is greater than previously reported. The ploidy series within species is especially noteworthy as it presents heretofore untapped opportunities for interploid breeding.

Genetic diversity

The primer statistics (PIC, H_o and H_e) can be useful in determining the genetic diversity in a population for the alleles at that locus. Because we looked at individual accessions, and not plants from one or more populations in Hardy–Weinberg equilibrium, the primer statistics have little meaning from a population genetics standpoint. However, these properties of the markers may be useful to future population studies in Corylopsis, especially for examining diversity or population structure within a species.

The SSR-based dendrogram (Fig. 2) provides insight into relationships between species, as well as highlights possible misidentification of accessions, both of which are valuable for collections management and breeding. While some accessions of a species grouped together as expected, many grouped in different clusters, particularly those with different ploidy. It is clear that ploidy played at least as large a role in clustering as species and geographic origin (Fig. 3). While it is possible that some accessions were mislabeled, which would lead to inconsistent clustering, it is also possible that some of the seedling material or accessions from nurseries or gardens may actually be interspecific hybrids resulting from outcrossing with other cultivated accessions. Wild-collected accessions (highlighted in bold in Table 1 and Fig. 2) are often deemed the most “authentic”
representatives of the species. However, based on our results, it is clear that there is substantial genetic and ploidy variability within species that cannot be explained by outcrossing in a garden setting. The SSR markers developed in this study may be useful for confirming some cases of potential interspecific hybrids in *Corylopsis*; otherwise, it may be necessary to use SNPs or direct sequencing to determine outcrossing and hybridization in the genus.

Among the accessions we studied, the most distantly related taxa based on the UPGMA algorithm occur in Cluster A (Fig. 2), which contains the only sample of *C. himalayana*, as well as two diploid and one tetraploid accession of *C. willmottiae*. Additional accessions of hexaploid *C. willmottiae* grouped in clusters I and J. The clustering of these species does not appear to be closely related to geographical origin of these accession, which all originated in the westernmost areas of collection (Fig. 3). The remaining accessions are grouped in Cluster B as a well-supported and distinct group that is resolved into two additional clusters, C and D, both with bootstrap values of 98%. Cluster C contains diploid accessions of *C. gotoana*, *C. glabrescens*, *C. pauciflora*, and *C. glandulifera*. Interestingly, accessions of tetraploid *C. gotoana* and hexaploid *C. glabrescens* did not cluster with the diploid accessions. This could be due to additional allelic diversity in the higher ploidy taxa, or possibly interspecific hybridization, or an allopolyploid origin. Previous karyomorphological studies of the centromere location (Kono et al. 2008) indicate possible divergence between Japanese diploids *C. pauciflora* and *C. glabrescens*, although our study did not indicate this.

Cluster D is further resolved into Clusters E and F. Cluster E contains only tetraploid *C. gotoana* var. *coreana* with high bootstrap support (96%). Previous work by Roh et al. (2007) using nrRNA-ITS sequence showed a tighter relationship among *C. gotoana* var. *coreana*, *C. glabrescens*, *C. gotoana*, and *C. pauciflora* than we saw in our study. This is likely due to different accessions and different types of markers used. Cluster F contains the remaining accessions, further divided into Clusters G and H. Cluster G also has high bootstrap support, and contains primarily tetraploid accessions of *C. glandulifera*, *C. sinensis* var. *calvescens*, *C. platypetala*, and *C. gotoana*. Previous molecular studies placed *C. sinensis* var *calvescens* in a group by itself (Roh et al. 2007) and grouped *C. gotoana* and *C. spicata* more closely than our data indicate (Yamanaka et al. 2008). Also in contrast to our study, studies by Kono et al. (2008) revealed similar karyotypes in tetraploid *C. gotoana* and hexaploid *C. spicata*, implying possible shared ancestries.

Cluster H is well-resolved (100%), and contains the final clusters, I and J. These clusters contain most of the hexaploid taxa, and are not as well-supported by bootstrap values, at least at the higher clustering levels. All accessions of *C. spicata* grouped in Cluster F, and most accessions of *C. sinensis*, hexaploid *C. glabrescens*, and hexaploid *C. veitchiana* grouped in Cluster E. A study performed on Hamamelidaceae
by Li (2008) using rDNA ITS sequence data grouped C. sinensis (and subspecies) and C. spicata together. While our study placed these two species in distinct clusters, the difference likely reflects the higher order taxa that were used in Li’s study, rather than contrasting results.

We used UPGMA to analyze and visualize the data instead of other methods such as STRUCTURE or Principal Coordinate Analysis (PCoA) because we were looking at individual plants (cultivars or unrelated accessions) and not at populations under selection. The microsatellite markers can be considered random or neutral in terms of selection or fitness, so a phenetic approach (UPGMA) is appropriate. Our primary objective was to examine relationships among accessions from a collections management perspective, rather than to ascertain evolutionary or phylogenetic histories. Indeed, upon visualizing the data in the UPGMA phenogram, we immediately found several accessions that were mislabeled in our collection or that led us to seek additional information on seed source or provenance. It is likely that different clustering algorithms, different plant materials, and different markers would result in different clusters of many accessions and taxa. For example, the study by Roh et al. (2007) based on RAPD markers or ITS sequence data yielded different dendrograms from each other and from the dendrogram presented here. We also performed PCoA on the data using the Jaccard distance coefficient, but found that this method was not useful at explaining variation—it took eleven dimensions to account for 50% of the explainable variation (data not shown). We are therefore confident that the methods we used to analyze the data are appropriate for our markers, plant material, and objective of assessing genetic diversity of living collections to assist with collection management and facilitate breeding.

**Conclusion**

This study determined ploidy and molecular genetic relationships among 106 Corylopsis accessions representing 14 taxa from botanic garden collections in the U.S. In addition to providing valuable ploidy data that was previously missing from many of these accessions, this study also developed a set of robust Corylopsis-specific SSR markers that will be useful across the genus to make decisions about species identity, collection management, or breeding targets.

**Acknowledgements** The authors thank Chris Benson (Ohio University), and Ron Beck, Richard T. Olsen, and Chandra Thammina (USDA US National Arboretum) for preliminary analysis using AFLP markers and testing these microsatellite markers.

**Funding** This study was funded by USDA-ARS base funds to Project 8020-21000-071-00D.

**Availability of data and material** SSR sequence data may be obtained from the corresponding author.

**Declarations**

**Conflict of interest** The authors declare that they have no conflict of interest.

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

**References**

Anderson E, Sax K (1935) Chromosome numbers in the Hamamelidaceae and their phylogenetic significance. J Arnold Arb 16:210–215

Bean E, Aniško T (2014) Captivating Corylopsis. Am Nursery (February):8–14

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am J Hum Genet 32:314–331

Gapinski A (2015) Hamamelidaceae, part 2: exploring the witch-hazel relatives of the Arnold Arboretum. Arnoldia 72(4):20–35

Garcia S, Garnatje T, Hildalgo O, Mas de Xaxars G, Pellicer J, Sanchez-Jimenez D, Vitales D, Valles J (2010) First genome size estimations for some eudicot families and genera. Collect Bot 29:7–16

Gillooly DA, Ranney TG (2015) Genome sizes and ploidy levels in the genus Kalmia. HortScience 50:1426–1428
Greilhuber J, Doležel J, Lysák MA, Bennett MD (2005) The origin, evolution, and proposed stabilization of the terms ‘genome size’ and ‘C-value’ to describe nuclear DNA contents. Ann Bot 95:255–260

Hassler M (2021) Synonymic checklists of the vascular plants of the world. In: Bánki O et al (eds) Catalogue of life checklist (version 2021–08–06). https://doi.org/10.48580/d4sw-3dd

Hatmaker EA, Wadl PA, Mantonth K, Scheffler BE, Ownley BH, Trigiano RN (2015) Development of microsatellites from *Fothergilla × intermedia* (Hamamelidaceae) and cross transfer to four other genera within Hamamelidaceae. Appl Plant Sci 3:1400123

Hinkley DJ (2009) The explorer’s garden: rare and unusual perennials. Timber Press, Portland, OR

Hop M (2011) An assessment of *Corylopsis*. Plantsman 10(1):10–15

Johnston JS, Bennett MD, Rayburn AL, Galbraith DW, Price HJ (1999) Reference standards for determination of DNA content of plant nuclei. Am J Bot 86:609–613

Kim SY, Kim CS, Kim GR, Kim JK, Park SH, Jang TS, Lee WK, Lee JK (2008) Chromosome numbers and karyotype analyses for 33 taxa of medicinal plants in Korea. Korean J Med Crop Sci 16:161–167

Kono Y, Okada M, Setoguchi J, Oginuma K (2008) Karyomorphology of *Corylopsis glabrescens* and *C. gotoana* endemic to Japan (Hamamelidaceae; Hamamelidoideae). Chromosom Bot 3:27–29

Lattier JD, Chen HC, Contreras RN (2017) Improved method of enzyme digestion for root tip cytology. HortScience 52:1029–1032

Li HL (1962) The cultivated *Corylopsis*. Morris Arbor Bull 13:63–68

Li J (2008) Molecular phylogenetics of Hamamelidaceae: evidence from DNA sequences of nuclear and chloroplast genomes. In: Sharma AK, Sharma A (eds) Plant genome: biodiversity and evolution, vol 1. part E. Science Publishers, Enfield, NH, pp 227–250

Li J, Bogle AL, Klein AS (1997) Phylogenetic relationships in the *Corylopsis* complex (Hamamelidaceae): evidence from sequences of the internal transcribed spacers of nuclear ribosomal DNA and morphology. Rhodora 99:302–318

Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics 26:403–404

Mehra PN, Khosla PK (1969) IOPB chromosome number reports XX. Taxon 18:105–107

Morley BD, Chao JM (1977) A Review of *Corylopsis* (Hamamelidaceae). J Arnold Arbor 58:382–415

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, NY, USA

Oginuma K (1991) Karyomorphology of *Corylopsis spicata* (Hamamelidaceae; Hamamelidoideae). Acta Phytotax Geobot 42:19–22

Oginuma K, Tobe H (1991) Karyomorphology and evolution in some Hamamelidaceae and Platanaceae (Hamamelidaceae; Hamamelidales). Bot Mag Tokyo 104:115–135

Parris JK, Ranney TG, Knapp HT, Baird VW (2010) Ploidy levels, relative genome sizes, and base pair composition in Magnolia. J Amer Soc Hort Sci 135:533–547

Qi ZC, Yu Y, Liu X, Pais A, Ranney T, Whetten R, Xiang QY (2015) Phylogenomics of polyploid *Fothergilla* (Hamamelidaceae) by RAD-tag based GBS – insights into species origin and effects of software pipelines. J Syst Evol 53:432–447

Ranney TG, Lynch NP, Fantz PR, Capiello P (2007) Clarifying taxonomy and nomenclature of *Fothergilla* (Hamamelidaceae) cultivars and hybrids. HortScience 42:470–473

Ranney TG, Thomasson TH, Neill K, Lynch NP, Weatherington M (2018) Ploidy, relative genome size, and inheritance of spotted foliage in *Aucuba* species (Garryaceae). HortScience 53:1271–1274

Roh MS, Lee AK, Choi IY, Kim JY, Joung YH, Lee SH, Suh JK (2007) Characterization of *Corylopsis coreana* using molecular markers. Hort Environ Biotechnol 48:176–187

Santamour FS Jr (1965) Chromosome number in *Corylopsis*. Morris Arb Bull 16:7

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. Nat Biotechnol 18:233–234

Shearer K, Ranney TG (2013) Ploidy levels and relative genome sizes of species, hybrids, and cultivars of dogwood (*Cornus* spp.). HortScience 48:825–830

Suzuki R, Terada Y, Shimodaia H (2017) *Pvclust*: hierarchical clustering with *p* values via multiscale bootstrap resampling. R package version 1.3–2. http://CRAN.R-project.org/package=pvclust. Accessed 22 Aug 2017

Takayama K, López PS, König C, Kohl G, Novak J, Stuessy TF (2011) A simple and cost-effective approach for microsatellite isolation in non-model plant species using small-scale 454 pyrosequencing. Taxon 60:1442–1449

USDA, Agricultural Research Service, National Plant Germplasm System (2021) Germplasm resources information network (GRIN-Taxonomy). National germplasm resources laboratory, Beltsville, Maryland. URL: https://npgsw.ars-grin.gov/gringlobal/taxon/taxonomysimple. Accessed 22 Aug 2021

Yamamaka M, Kobayashi S, Setoguchi H (2008) Distinct geographical structure across species units evidenced by chloroplast DNA haplotypes and nuclear ribosomal ITS genotypes of *Corylopsis* (Hamamelidaceae) in the Japanese islands. Bot J Linnean Soc 157:501–518

Zhang Z, Zhang H, Endress PK (2003) Hamamelidaceae. In: Wu ZY, Raven PH, Hong DY (eds) *Corylopsis*: flora of China, (Pittosporaceae through Connaraceae), vol 9. Science Press, Beijing, China, and Missouri Botanical Garden Press, St. Louis, Missouri, USA, pp 35–40

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