



Article

Exploring Genetic Diversity in an *Ilex crenata* Breeding Germplasm

Esther Geukens ¹, Annelies Haegeman ¹, Jef Van Meulder ², Katrijn Van Laere ¹, Erik Smolders ³, Tom Ruttink ^{1,4} and Leen Leus ^{1,*}

- ¹ Plant Sciences Unit, Flanders Research Institute for Agricultural, Fisheries and Food Research (ILVO), Caritasstraat 39, 9090 Melle, Belgium; esther.geukens@ilvo.vlaanderen.be (E.G.); annelies.haegeman@ilvo.vlaanderen.be (A.H.); katrijn.vanlaere@ilvo.vlaanderen.be (K.V.L.); tom.ruttink@ilvo.vlaanderen.be or tom.ruttink@ugent.be (T.R.)
- ² Arboretum Provincial Domain Bokrijk, Bokrijk, 3600 Genk, Belgium
- ³ Division of Soil and Water Management, KU Leuven, Kasteelpark Arenberg 20, 3001 Leuven, Belgium; erik.smolders@kuleuven.be
- ⁴ Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 71, 9052 Ghent, Belgium
- * Correspondence: leen.leus@ilvo.vlaanderen.be

Abstract: Knowledge of genetic identity, genetic relationships, ploidy level, and chromosome numbers can enhance the efficiency of ornamental plant breeding programs. In the present study, genome sizes, chromosome numbers, and genetic fingerprints were determined for a collection of 94 *Ilex* accessions, including 69 *I. crenata*. The genome size of the entire collection ranged from 1.50 ± 0.03 to 8.01 ± 0.18 . Within the species of *I. crenata*, genome sizes varied (mean $\pm$ sd) between 1.98 ± 0.08 and 2.30 ± 0.04 , with three outliers: 3.06 ± 0.04 , 4.04 ± 0.09 , and 4.19 ± 0.08 . The chromosome counting results showed $2n = 40$ for *I. crenata* accessions and confirmed the outliers as one triploid and two tetraploids. A high intra-specific genetic diversity in *Ilex crenata* was found, after genetic fingerprinting using genotyping-by-sequencing (GBS). The species *I. crenata* was separated into three clades, which coincided with intraspecific differences in genome sizes (mean $\pm$ sd) of 2.09 ± 0.006 , 2.07 ± 0.05 , and 2.19 ± 0.06 pg/2C per clade as mean values for the diploids. Applying a principal coordinate analysis (PCoA) to the genetic fingerprinting data of all species in the collection revealed a wide genetic variation, which has not yet been commercially exploited. These findings could form the basis for selectively breeding parents, in order to create more genetic diversity via intra- and interspecific crosses.

Keywords: holly; genotyping-by-sequencing; 2C-value; chromosome number; genome size



Citation: Geukens, E.; Haegeman, A.; Van Meulder, J.; Van Laere, K.; Smolders, E.; Ruttink, T.; Leus, L. Exploring Genetic Diversity in an *Ilex crenata* Breeding Germplasm. *Horticulturae* **2023**, *9*, 485. <https://doi.org/10.3390/horticulturae9040485>

Academic Editor: Carlo Nicoletto

Received: 13 March 2023

Revised: 7 April 2023

Accepted: 11 April 2023

Published: 13 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The genus *Ilex* belongs to the monogeneric family Aquifoliaceae and is a large dioecious genus with 571 accepted species [1]. The genus can be found in regions of mesic growing conditions around the world. While most *Ilex* species originate from East Asia and South America, a multitude of species can also be found in Southeast Asia and North America. Only four species occur in Europe, one occurs in Africa, and another species occurs in Northern Australia. A few species can be found in Hawaii, the Caribbean, the Canary Islands, the Azores, Madeira, New Caledonia, and Fiji [2]. Previous phylogenetic studies have used fossil records [2–4] and nuclear and plastid markers to unravel the evolution of the *Ilex* genus, revealing a complex evolutionary history. Moreover, the taxonomy of the genus is not yet fully resolved. There are multiple heterotypic and homotypic synonyms within the genus, which occasionally cause confusion. For example, *I. maximowicziana* var. *kanehirae* (Yamamoto) Yamazaki and *I. crenata* var. *mutchagara* (Makino) Hara are two names given to the same plant [5]. The same applies to *I. leucoclada* (Maxim.) Makino and *I. integra*

var. *leucoclada* Maxim. [1]. Recently, the increasing interest in the genus and decreasing cost of DNA sequencing have led to reference genomes for *I. latifolia* [6], *I. polyneura* [4], and *I. asprella* [7]. In some parts of the world, *Ilex* species such as *I. opaca*, *I. vomitoria*, and *I. paraguariensis* are used for brewing teas, while *I. asprella* is used as a medicinal plant [8]. *I. aquifolia* (common holly, English holly, European holly, or Christmas holly) is well known as an ornamental garden plant, with typical wintergreen branches and red berries used as a Christmas decoration. Other hollies used for their ornamental properties include *I. polyneura* and *I. crenata* (box-leaved holly or Japanese holly). *I. crenata* resembles boxwood (*Buxus* L.) and is also used for hedges and topiary [9].

I. crenata originates from Japan, Korea, China, the Sakhalin Islands, the Kuril Islands, Taiwan, the Philippines, and the Himalayas. In Japan, the species naturally grows from sea level up to an altitude of 1000 m, with local populations exhibiting wide morphological variation [9]. The species was introduced to the garden of the Russian Czar Alexander II in 1864 by Russian botanist Carl Maximowicz. Starting from that introduction, distribution spread throughout Europe. In 1898, *I. crenata* was introduced to the United States. This evergreen, small-leaved shrub gained popularity in the 1930s as an ornamental plant. Since then, *I. crenata* has been introduced multiple times to other parts of the world, and the plant has been extensively used for breeding [5], leading to the release of numerous new cultivars. These cultivars were mainly lucky findings, selected based on phenotypic variation and subsequently vegetatively propagated [9,10]. With more targeted breeding, value could be added, by overcoming abiotic and biotic stress challenges, such as high soil pH and infection with black root rot caused by the soil-dwelling fungus *Berkeleyomyces basicola*. Knowledge about the genetic relations and variation in existing species and cultivars could guide breeders selecting breeding parents, thus paving the way for novel or improved, and thus commercially marketable, varieties. Moreover, knowledge of the ploidy level and chromosome number of genotypes can help to predict cross-compatibility. The genome sizes of *I. cornuta*, *I. mucronata*, and *I. verticillata* have previously been reported in the range of 1.31 pg/2C [11] to 4.13 pg/2C [12]. The Chromosome Count Database describes 41 *Ilex* species ranging from $34 = 2n$ for *I. crenata* and *I. leucoclada* [13,14] to $120 = 2n$ for *I. pedunculosa* [15,16]. For both *I. crenata* and *I. leucoclada*, Galle reports the chromosome number to be $40 = 2n$ [9]. Darlington and Wylie mention a basic number of $x = 9$ or 10 [17].

Some efforts have been made to clarify the phylogeny of the different species within the genus, but knowledge about the underlying genetic variation within the species *I. crenata* is still limited.

The present study included an investigation into the genetic diversity within the species *I. crenata* and a selection of other hardy *Ilex* species. The genetic identity of the plants in the collection was also examined in light of their taxonomy. Naming mistakes can occur for many reasons, including ex situ conservation over the years, selling plants under an incorrect name for economic reasons, erroneous determinations, etc. We assembled a collection of 94 accessions, comprising 69 accessions labeled as *I. crenata* and 28 accessions from 17 other species. The genomic constitution was characterized at three levels: (1) genome sizes were determined using flow cytometry, (2) the chromosome number and ploidy level were investigated through fluorescence microscopy, and (3) genotyping-by-sequencing (GBS) [18] was performed as a reduced-representation genome sequencing technique, which allowed for high-resolution genetic fingerprinting using thousands of genetic markers spread across the genome.

2. Materials and Methods

2.1. Collection of Plant Material for the *Ilex* Collection

A germplasm collection was compiled of 94 *Ilex* accessions, belonging to 18 species and hybrids including *I. x attenuata*, *I. aquifolium*, *I. x aquipernyi*, *I. crenata*, *I. glabra*, *I. integra*, *I. leucoclada*, *I. macrocarpa*, *I. x makinoi*, *I. maximowicziana*, *I. x meserveae*, *I. mutchagara*, *I. opaca*, *I. rugosa*, *I. serrata*, *I. sugerokii*, *I. vomitoria*, and *I. yunnanensis*. The collection also contains a total of 69 accessions labeled as *I. crenata* (Table 1). Plant material from *I. crenata*

‘Fastigiata’ was sampled multiple times from independent sources. Thirteen accessions were purchased as potted plants from Esveld Plant Garden (The Netherlands), 78 genotypes were collected as cuttings from the arboreta of Bokrijk (Genk, Belgium) and Het Leen (Eeklo, Belgium), and 3 genotypes were obtained as potted plants from the company Plant Select. The plants were maintained as potted plants in a container field at ILVO (51°00′ N, 3°78′ E, Melle, Belgium) (Table 1).

Table 1. Overview of the *Ilex* collection. The table shows all the accessions and for every genotype: the accession number used at ILVO; the species and cultivar name; the collection of origin, along with the accession number from the arboretum; the mean genome size; and the chromosome number. Genetically identical accessions are indicated with the same superscript letter (a–d), accessions for which a naming problem was discovered are indicated with an asterisk.

ILVO Accession Number	Species and Cultivar Name	Collection of Origin	Arboretum Accession Number	Mean Genome Size (pg/2C ± sd)	Chromosome Number
IL001	<i>Ilex x aquipernyi</i>	Esveld		1.80 ± 0.09	40
IL002	<i>Ilex crenata</i> ‘Caroline Upright’	Esveld		2.11 ± 0.06	
IL003	<i>Ilex crenata</i> ‘Convexa’	Esveld		2.13 ± 0.05	
IL004	<i>Ilex crenata</i> ‘Dwarf Pagoda’	Esveld		2.30 ± 0.05	
IL005 ^a	<i>Ilex crenata</i> ‘Fastigiata’	Esveld		2.16 ± 0.02	
IL006	<i>Ilex crenata</i> ‘Glorie Dwarf’	Esveld		1.99 ± 0.07	
IL007	<i>Ilex crenata</i> ‘Green Hedge’	Esveld		2.10 ± 0.06	
IL008	<i>Ilex crenata</i> ‘Green Lustre’	Esveld		2.02 ± 0.14	
IL009	<i>Ilex crenata</i> ‘Microphylla’	Esveld		2.16 ± 0.05	
IL010	<i>Ilex crenata</i> ‘Stokes’	Esveld		2.10 ± 0.04	
IL011	<i>Ilex crenata</i> ‘Twiggy’	Esveld		2.13 ± 0.08	
IL012	<i>Ilex glabra</i>	Esveld		1.52 ± 0.05	40
IL013 ^{b,*}	<i>Ilex maximowicziana</i> var. <i>kanehirae</i>	Esveld		2.14 ± 0.09	40
IL014 ^c	<i>Ilex crenata</i> ‘Dark Green’	Plant Select		2.18 ± 0.05	
IL015 ^b	<i>Ilex crenata</i> ‘Blondie’	Plant Select		2.30 ± 0.04	
IL016	<i>Ilex crenata</i> ‘Braddock Heights’	Het Leen	00004670	1.98 ± 0.08	40
IL017	<i>Ilex crenata</i> ‘Cherokee’	Het Leen	00004020	4.19 ± 0.08	80
IL018	<i>Ilex crenata</i> ‘Convexa Gold’	Het Leen	00004672	2.05 ± 0.06	
IL019	<i>Ilex crenata</i> ‘Ellipta Gold’	Het Leen	00004017	2.25 ± 0.04	
IL020	<i>Ilex crenata</i> ‘Erecta’	Het Leen	00004021	2.12 ± 0.09	
IL021 ^a	<i>Ilex crenata</i> ‘Fastigiata’	Het Leen	00001571	2.09 ± 0.04	40
IL022 ^a	<i>Ilex crenata</i> ‘Fastigiata’	Het Leen	00001904	2.05 ± 0.04	
IL023	<i>Ilex crenata</i> ‘Golden Gem’	Het Leen	00001905	2.08 ± 0.08	40
IL024 ^d	<i>Ilex crenata</i> ‘Golden Helli’	Het Leen	00004014	2.09 ± 0.05	
IL025	<i>Ilex crenata</i> ‘Green Dragon’	Het Leen	00004013	2.13 ± 0.07	
IL026	<i>Ilex crenata</i> ‘Green Island’	Het Leen	00004090	2.12 ± 0.06	
IL027	<i>Ilex crenata</i> ‘Hetzii’	Het Leen	00004022	1.99 ± 0.05	40
IL028	<i>Ilex crenata</i> ‘Ivory Tower’	Het Leen	00004673	2.08 ± 0.05	
IL029 ^d	<i>Ilex crenata</i> ‘Lancaster Yellow’	Het Leen	00004024	2.03 ± 0.07	
IL030	<i>Ilex crenata</i> ‘Luteovariegata’	Het Leen	00005561	2.09 ± 0.05	
IL031	<i>Ilex crenata</i> ‘Mariesii’	Het Leen	00004010	2.22 ± 0.05	
IL032	<i>Ilex crenata</i> ‘Nakada’	Het Leen	00004016	2.23 ± 0.03	
IL033	<i>Ilex crenata</i> ‘Picas’	Het Leen	00004091	2.09 ± 0.06	
IL034	<i>Ilex crenata</i> ‘Piccolo’	Het Leen	00004004	2.13 ± 0.06	
IL035	<i>Ilex crenata</i> ‘Pride’s Tiny’	Het Leen	00004023	1.98 ± 0.06	40
IL036	<i>Ilex crenata</i> ‘Red Lion’	Het Leen	00004018	2.22 ± 0.03	
IL037	<i>Ilex crenata</i> ‘Robert Culpepper’	Het Leen	00004011	2.04 ± 0.03	
IL038	<i>Ilex crenata</i> ‘Rotundifolia’	Het Leen	00004671	2.03 ± 0.05	
IL039	<i>Ilex crenata</i> ‘Snowflakes’	Het Leen	00004669	2.25 ± 0.06	
IL040	<i>Ilex crenata</i> ‘Tee Dee’	Het Leen	00004025	2.26 ± 0.05	
IL042	<i>Ilex crenata</i> f. <i>watanabeata</i>	Het Leen	00005253	2.04 ± 0.04	
IL043	<i>Ilex crenata</i> ‘Wiesmoor Silber’	Het Leen	00004012	2.09 ± 0.07	
IL044	<i>Ilex crenata</i> ‘Yunnanensis’	Het Leen	00004019	2.18 ± 0.09	
IL045	<i>Ilex x aquipernyi</i>	Het Leen	00005002	1.78 ± 0.04	
IL046	<i>Ilex x meserveae</i> ‘Blue Girl’	Het Leen	00001829	1.93 ± 0.02	40
IL047 ^{b,*}	<i>Ilex mutchagara</i> Makinoi	Het Leen	00001901	2.17 ± 0.06	
IL048 ^{b,*}	<i>Ilex yunnanensis</i>	Het Leen	00004449	2.20 ± 0.03	
IL049	<i>Ilex aquifolium</i> ‘Angustifolia’	Het Leen	00001859	1.85 ± 0.04	
IL050	<i>Ilex aquifolium</i> ‘Crispa’	Het Leen	00001873	1.90 ± 0.07	
IL051 ^c	<i>Ilex x makinoi</i> Hara	Het Leen	00004998	2.00 ± 0.06	40
IL052	<i>Ilex x attenuata</i> ‘Foster’	Het Leen	00004994	1.89 ± 0.05	
IL053	<i>Ilex opaca</i> ‘Longwood Gardens’	Het Leen	00004993		
IL055	<i>Ilex crenata</i> ‘Black Beauty’	Bokrijk	19991318	3.06 ± 0.04	60
IL056	<i>Ilex crenata</i> ‘Border Gem’	Bokrijk	19991317	2.07 ± 0.09	
IL059	<i>Ilex crenata</i> ‘Cape Fear’	Bokrijk	19991320	2.09 ± 0.05	
IL060	<i>Ilex crenata</i> ‘Chesapeake’	Bokrijk	19900267	4.04 ± 0.09	80
IL061	<i>Ilex crenata</i> ‘Cole’s Hardy’	Bokrijk	1983ACQI175	2.14 ± 0.08	
IL062 ^a	<i>Ilex crenata</i> ‘Fastigiata’	Bokrijk	19900529	2.16 ± 0.08	
IL063	<i>Ilex crenata</i> ‘Goldstaub’	Bokrijk	19940813	2.16 ± 0.08	
IL065	<i>Ilex crenata</i> ‘Hatfieldii’	Bokrijk	19950491	2.06 ± 0.10	

Table 1. Cont.

ILVO Accession Number	Species and Cultivar Name	Collection of Origin	Arboretum Accession Number	Mean Genome Size (pg/2C ± sd)	Chromosome Number
IL066	<i>Ilex crenata</i> ‘Helleri’	Bokrijk	19940347	2.25 ± 0.05	40
IL067	<i>Ilex crenata</i> ‘Highlander’	Bokrijk	19950296	2.12 ± 0.07	
IL069	<i>Ilex crenata</i> ‘Horizontalis’	Bokrijk	19910107	2.18 ± 0.08	
IL070	<i>Ilex crenata</i> ‘Ivory Hall’	Bokrijk	19991306	2.09 ± 0.08	
IL071	<i>Ilex crenata</i> ‘Kingsville Dwarf’	Bokrijk	19940329	2.29 ± 0.08	40
IL072	<i>Ilex crenata</i> ‘Midas Touch’	Bokrijk	19900222	2.13 ± 0.07	
IL073	<i>Ilex crenata</i> ‘Monmouth’	Bokrijk	19950499	2.18 ± 0.06	
IL075	<i>Ilex crenata</i> ‘Oleafera’	Bokrijk	19980688	2.04 ± 0.06	
IL076	<i>Ilex crenata</i> var. <i>paludosa</i>	Bokrijk	19991354	2.05 ± 0.10	40
IL077	<i>Ilex crenata</i> ‘Piedmont Pyramidal’	Bokrijk	19900271	2.16 ± 0.03	
IL078	<i>Ilex crenata</i> ‘Schwoebel’s Compact’	Bokrijk	19950345	2.15 ± 0.03	
IL079	<i>Ilex crenata</i> ‘Sentinel’	Bokrijk	1988ACQI162	2.01 ± 0.09	
IL082 ^a	<i>Ilex crenata</i> ‘Sky Pencil’	Bokrijk	20040562	2.15 ± 0.07	40
IL084	<i>Ilex crenata</i> ‘Tennyson’	Bokrijk	19900395	2.28 ± 0.06	
IL086	<i>Ilex crenata</i> ‘Valeria Rankin’	Bokrijk	19900274	2.08 ± 0.11	
IL087	<i>Ilex crenata</i> ‘Variegata’	Bokrijk	19980670	2.22 ± 0.04	
IL088	<i>Ilex crenata</i> ‘Fructu Lutea’	Bokrijk	19890583	2.01 ± 0.07	40
IL094	<i>Ilex vomitoria</i> ‘Folsom Weeping’	Bokrijk	20200779	1.57 ± 0.10	
IL095	<i>Ilex vomitoria</i> ‘Hoskins shadow’	Bokrijk	20110310	1.58 ± 0.06	
IL096	<i>Ilex sugerokii</i>	Bokrijk	19980060	1.78 ± 0.07	
IL097	<i>Ilex sugerokii</i> var. <i>longipedunculata</i>	Bokrijk	19900505	1.75 ± 0.02	40
IL098	<i>Ilex rugosa</i>	Bokrijk	20070853	1.97 ± 0.04	
IL099	<i>Ilex leuoclada</i>	Bokrijk	20020979	2.00 ± 0.05	
IL100	<i>Ilex opaca</i>	Bokrijk	19940225	1.94 ± 0.02	
IL101	<i>Ilex serrata</i> ‘Rakusogu’	Bokrijk	0000ACQI10	1.86 ± 0.07	160
IL102	<i>Ilex integra</i>	Bokrijk	20120839	1.81 ± 0.04	
IL103	<i>Ilex macrocarpa</i>	Bokrijk	19960445		
IL104	<i>Ilex yunnanensis</i> var. <i>gentilis</i>	Bokrijk	00004515	8.01 ± 0.18	
IL105	<i>Ilex crenata</i> ‘Jeran’ (Bokrijk select)	Bokrijk	20140097	2.15 ± 0.07	40
IL106	<i>Ilex crenata</i> ‘Convexa Variegata’	Bokrijk	19900459	2.17 ± 0.05	
IL107	<i>Ilex crenata</i> ‘Nummularia’	Bokrijk	19991312	2.23 ± 0.08	
IL108	<i>Ilex vomitoria</i> ‘Will Fleming’	Bokrijk	20080892	1.50 ± 0.03	
IL121	<i>Ilex leuoclada</i> seedling	Bokrijk			40
IL122	<i>Ilex crenata</i> ‘Luxus’	Plant Select		2.15 ± 0.08	

2.2. Genome Size of the Collected *Ilex* Accessions

The genome size of 91 accessions in the collection was measured using flow cytometry (Table 1), according to the methods described by Van Oost et al. 2021 [19]. Four accessions did not develop rooted cuttings and could not be used. The genome size is expressed as pg/2C. Terminology used is according to Greilhuber et al. [20]. All samples were analyzed using a Quantum P flow cytometer and CyPAD software (Quantum Analysis, Münster, Germany). For all genotypes, except IL104, the internal standard used was *Zea mays*, with a known genome size of 5.43 pg/2C [21]. *Pisum sativum* (9.09 pg/2C [22]) was used as an internal standard for IL104, because of its deviant genome size. Young leaf material of the sample and standard was mixed, co-chopped, and used for sample preparation with a CyStain PI kit (Sysmex, Münster, Germany), according to the manufacturer’s protocol with minor modifications: using 0.4 mL extraction buffer, and 1.2 mL staining buffer, to which 1% (w/w) PVP10 was added. The stained samples were incubated in the dark at 4 °C for at least 20 min. At least three replicates were analyzed for every genotype, preferably from three plants of this genotype and on three different days. The sample genome size was derived by calculating the ratio of the peak positions of the sample and internal standard in the histograms of both the FL2 and FL3 detectors. Mean values and standard deviations of the sample genome size were calculated as the mean of a minimum of six values, specifically the peak position ratio of both detectors for at least three replicates.

2.3. Chromosome Number of the Collected *Ilex* Accessions

Based on the flow cytometry results, a subset of 17 *I. crenata* genotypes (Table 1), covering the full range of genome sizes, was selected to count the chromosomes. The chromosomes of the following species were also counted: *I. x aquipernyi* (IL001), *I. glabra* (IL012), *I. maximowicziana* var. *kanehirae* (IL013), *I. x meserveae* (IL046), *I. x makinoi* (IL051), *I. rugosa* (IL098), and *I. leuoclada* (IL099) (Table 1). Young root tips of newly rooted cuttings were harvested and incubated in 0.1% colchicine for 3 h at room temperature, to arrest mitosis at the metaphase stage. Subsequently, the root tips were fixed in 3:1 ethanol:acetic

acid for 1 h at room temperature. Cell suspensions were made by digesting the root tips in 0.6% enzyme solution (0.6% cellulase, 0.6% pectolyase, and 0.6% cytohelicase in 0.1 M citrate buffer (0.1 M sodium citrate tribasic dehydrate, and 0.1 M citric acid)) at 37 °C for 45 min. Chromosome slides were prepared using the SteamDrop protocol [23], using 1:1 ethanol:acetic acid solution for both fixation steps. Subsequently, the chromosome slides were stained with 1% DAPI (100 g/mL) diluted in Vectashield mounting medium and analyzed with a fluorescence microscope (AxioImager M2, Carl Zeiss MicroImaging, Belgium) at 1000× magnification. Images were captured using an Axiocam MRm camera and ZEN software (Carl Zeiss MicroImaging, Belgium). Chromosomes were counted on at least five images of well-spread metaphase chromosomes per genotype using DRAWID v0.26 [24].

2.4. GBS Fingerprinting of the ILEX Collection

2.4.1. Library Preparation and Sequencing

DNA was extracted from all genotypes of the collection using a modified CTAB protocol [25]. Preparation of genotyping-by-sequencing (GBS) libraries and sequencing was performed by LGC Genomics (Berlin, Germany), using the following steps: (1) quality control (agarose gel check); (2) library preparation, including indexing and quality control using a double digest with *Pst*I and *Mse*I; and (3) PE-150 sequencing using an Illumina NextSeq 500/550 instrument.

2.4.2. GBS Data Analysis

All data analysis steps described below are also available, together with accompanying scripts, on https://gitlab.com/ilvo/GBS_Ilex (accessed on 25 May 2021) and on Zenodo with doi 10.5281/zenodo.7669149 under an MIT license.

Data preprocessing. The obtained GBS reads were processed using GBprocess v3.0.3 [26]. This software makes use of Cutadapt [27] and PEAR [28], and comprises the following steps: (1) trimming of adapters, barcodes, restriction site remnants from 5'- and 3'-ends, and discarding reads that are shorter than 30 bp using Cutadapt; (2) merging of forward and reverse reads with a minimum overlap of 10 bp and a minimal final size of 40 bp with PEAR [28]; (3) filtering of reads that have a mean Phred quality score lower than 25 or two consecutive bases with a Phred quality score below 20 and discarding reads containing Ns; and (4) filtering of reads with intact internal restriction enzyme recognition sites (due to partial restriction digest). The raw PE-150 read data, as well as the merged preprocessed read data, were submitted to NCBI's Sequence Read Archive (SRA) under the BioProject number PRJNA895194.

Loci identification. Data were analyzed reference-free, using GIBPSs software [29]. Unless mentioned otherwise, standard settings were used and the recommended workflow of GIBPSs was followed. More detailed information can be found in the accompanying software repository, as mentioned above. Briefly, loci were identified in all merged reads using *indloc*, *poploc*, and *indpoploc*, and errors were corrected using *indloc*. Loci of lengths between 32 and 300 bp were selected using *data selector*. Loci containing indels were detected using *indelchecker* and discarded. Loci that were extremely deeply sequenced were identified using *depth analyzer* and discarded. The minimum value for the median depth percentage and the maximum value for the median scaled depth were set to 1 and 0.15, respectively. Additionally, split loci were removed from the dataset. Finally, a GIBPSs genotype call table was exported, containing the locus genetic information for all samples.

Determination of common loci and allele similarity. All samples were compared to each other using a custom python script available in the script repository [30], which uses the GIBPSs genotype call table as input. This script calculates (1) the number of loci per sample; (2) the number of common loci between every pair of samples and; (3) the allele similarity of the common loci of every pair of samples. The allele similarity is defined as the percentage of common loci that share at least one allele. An allele similarity of 100% means that all common loci have at least one shared allele (=a locus sequence with a combination

of SNPs) (see Supplementary Materials). The results are represented in a false-colored similarity matrix, displaying the total number of loci for every genotype on the diagonal, the number of loci in common between two genotypes in the upper half of the matrix, and the allele similarity in the lower half of the matrix.

Concatenated alignment and phylogeny. Using a custom perl script available in the script repository, all SNP-containing loci from the GIBPSs genotype call table were concatenated and saved as an alignment file. In addition, a text file that kept track of the original locus (and positions in that locus) in the concatenated alignment was saved. Next, again using a custom perl script, positions where more than 20% of the samples contained missing data were filtered out. Invariant positions were removed using the ‘ascbias.py’ script (available at https://github.com/btmartin721/raxml_ascbias) (accessed on 25 May 2021). The resulting filtered concatenated alignment was used for (1) principal coordinate analysis (PCoA); and (2) phylogenetic tree construction. The PCoA was performed by reading the multiple alignments in R v4.2.2 in R Studio Server v2022.07.2 build576, using the Biostrings package. Next, the pairwise distance between all individuals was calculated using the Hamming distance (Decipher package), while ignoring gap–letter matches and performing the PCoA using the ‘cmdscale’ function (stats package). The R scripts are available in the script repository. This analysis was performed for the entire collection of 94 individuals and two technical replicates, as well as for the 69 individuals in the *I. crenata* group only and one technical replicate. A phylogenetic tree was constructed containing the collection of all 94 individuals using the maximum likelihood method with RAXML v8.2.12, including a GTRCAT model without rate heterogeneity (-V), with ascertainment bias correction (–asc-corr=lewis), and including 100 bootstraps. *I. vomitoria* ‘Hoskins Shadow’ (IL095) was used as an outgroup. The tree was then visualized using iTOL v6 [31].

3. Results and Discussion

This study explored the genetic diversity within the genus of *Ilex*, more specifically within the species *I. crenata*. We assembled a large collection of 94 *Ilex* accessions from arboreta and commercial growers, 69 of which were labeled *I. crenata*, as well as three others with other names that were genetically identical to *I. crenata*. For the accessions in this collection, we delineated the genetic identity, investigated the genetic relationships, and studied the genetic diversity between the accessions of the collection.

3.1. Genetic Identity

GBS was performed on all 94 accessions of the *Ilex* collection. The number of preprocessed merged reads per sample (see Materials and Methods above) ranged from 1.8 M to 8.4 M, with a median of 3.6 M. Reference-free locus delineation was performed using GIBPSs and resulted in a total of 218,024 loci for all samples combined. Rarefaction analysis of a few selected samples revealed that the number of identified loci using GIBPSs reached a plateau at ca. 3–4 M reads, suggesting that most of the samples had been sequenced deep enough to reach locus saturation (data not shown). The number of identified loci per sample ranged from 9513 to 53,403, with a mean of 18,079. These loci had a mean length of 184 bp (range 40 to 270 bp), a mean read depth of 57, and a median read depth of 20. In total, 70% of the loci had a mean read depth per sample of at least 15. The GBS loci were compared between samples in a pairwise manner to estimate the number of common loci and the fraction of loci with shared alleles (for definitions of ‘common loci’ and ‘allele similarity’, see Materials and Methods and [30]). The heat map (Figure S1) shows clear clusters, indicating the genetic substructure in the *Ilex* collection. The lowest number of common loci between species was 2880 for *I. crenata* ‘Luxus’ (IL122) and *I. macrocarpa* (IL103); the highest number was 11,249 for *I. crenata* ‘Pride’s Tiny’ (IL035) and *I. sugerokii* var. *longipedunculata* (IL097). Within the species *I. crenata*, the lowest number of common loci between two accessions was 6340 for *I. crenata* ‘Luxus’ (IL122) and *I. crenata* ‘Nummularia’ (IL107); the highest was 15,219 for *I. crenata* ‘Braddock Heights’ (IL016) and *I. crenata* ‘Pride’s Tiny’

(IL035). The lowest allele similarity between species was 8.85% between *I. macrocaropa* (IL103) and *I. vomitoria* ‘Will Fleming’ (IL108); the highest allele similarity between species was 60.25% for *I. opaca* (IL100) and *I. integra* (IL102).

Next, we analyzed the allele similarity of replicates, to set a threshold for identifying genetically identical accessions across the *Ilex* collection. For two accessions (IL014 and IL051), we used two technical replicates to analyze the reproducibility of GBS fingerprinting. The total number of loci of *I. crenata* ‘Dark Green’ (IL014) replicates was similar (11,312 and 11,945), with 10,335 loci in common. Likewise, technical replicates of *I. x makinoi* ‘Hara’ (IL051) had 16,716 and 19,756 loci, with 15,195 loci in common. The allele similarity between these technical replicates was 99.88% and 99.77% for IL014 and IL051, respectively. Furthermore, we collected *I. crenata* ‘Fastigiata’, a common cultivar, from four sources as ‘biological’ replicates (IL005, IL021, IL022 and, IL062). As expected, this group of Fastigiata accessions displayed high allele similarities, ranging from 98.13% to 99.68%. Taken together, these results showed that the allele similarity method for pairwise comparison based on GBS genetic fingerprints was especially accurate. The number of common loci was less reliable for identifying similar genotypes, which was probably due to datasets that were not completely saturated. Consequently, we used a 98% threshold of minimal allele similarity to consider GBS genetic fingerprints as identical (=the same cultivar) and screened the *Ilex* collection for genetically identical accessions. For instance, *I. crenata* ‘Sky Pencil’ (IL082), morphologically identical to the Fastigiata genotypes with an upright phenotype, can also be considered genetically identical to the Fastigiata genotypes (allele similarity ranging from 98.19% with IL022 to 99.50% with IL005). Furthermore, the allele similarity of *I. crenata* ‘Golden Helleri’ (IL024) and *I. crenata* ‘Lancaster Yellow’ (IL029) was 99.89%, thus surpassing this threshold. Dudley and Eisenbeiss (1992) previously reported that a mutation of ‘Helleri’ was commercialized under the name ‘Lancaster Yellow’ [5], thus confirming that these accession names are genetically identical. Comparison of the allele similarities of other genotypes revealed that *I. yunnanensis* (IL048) is highly related to *I. mutchagara* Makino (IL047), *I. crenata* ‘Blondie’ (IL015), and *I. maximowicziana* var. *kanehirae* (IL013), with allele similarity values ranging from 98.83% to 99.76% (Figure S1). There are two reasons for these high allele similarities. The first is that *I. yunnanensis* (IL048) is actually an *I. crenata* genotype. Visual assessment of the original plants in the arboretum revealed that the mother plant of *I. yunnanensis* (IL048) looked like an *I. crenata* and not like the other *I. yunnanensis* in the collection, leading to the conclusion that the mother plant, and subsequently the *I. yunnanensis* (IL048) accession in our collection, was initially incorrectly labeled as *I. yunnanensis*. The other reason for the observed high allele similarities is that *I. maximowicziana* var. *kanehirae* (Yamam.) T. Yamaz. and *I. mutchagara* Makino are both synonyms of one another and are also synonyms of *I. crenata* var. *mutchagara* (Makino) Ohwi, as previously reported by Dudley and Eisenbeiss (1992) [5]. More recently, many more heterotypic synonyms have been added to the list of synonyms, including *I. crenata* var. *kanehirae* Yamam., *I. crenata* var. *scoriatum* Yamam., and *I. maximowicziana* Loes. [1]. Our data show that both *I. maximowicziana* var. *kanehirae* (IL013) and *I. mutchagara* Makino (IL047) are not more genetically different from *I. crenata* than the different *I. crenata* genotypes are from one another (more than 98% allele similarity) (Figure S1). Moreover, they are genetically almost identical to the *I. crenata* cultivar ‘Blondie’, with allele similarities only slightly lower than that of the technical replicates. In conclusion, our data confirm that *I. maximowicziana* var. *kanehirae* (Yamam.) T. Yamaz. and *I. mutchagara* Makino are synonyms and further show that they are also synonyms of *I. crenata* Thunb.

3.2. Genetic Relationships among *Ilex* Species

Pairwise comparison of allele similarities across the collection showed blocks of samples that are more closely related to each other. These blocks were consistent with the clades of the phylogenetic tree (Figure 1) and the clusters of the PCoA (Figure 2). Four clusters were distinguished using PCoA analysis of the GBS data of the *Ilex* collection. Cluster 1, containing *I. crenata* accessions, was separated on the PC1 axis. Cluster 2, containing

I. vomitoria accessions, was separated on the PC2 axis. Cluster 3, containing *I. opaca* accessions, was separated from Cluster 4, containing multiple other species. The phylogenetic tree generally agreed with this genetic structure. The species *I. vomitoria* was used as an outgroup, as it became clear during preliminary analyses that this species was phylogenetically less related to the other species than the other species were related to one another. Besides *I. vomitoria*, *I. opaca* was also less related to the other species than the remaining species were related to each other. The phylogenetic analysis generally confirmed the species boundaries, as the different species were clearly grouped in separate clusters in the phylogenetic tree, with high bootstrap support (100% for species).

In previously published phylogenetic trees based on a limited number of nuclear markers (Cuénoud et al. [2] and Manen et al. [3]), *I. sugerokii* and *I. yunnanensis* were closely related. However, in phylogenetic trees based on plastid markers, *I. sugerokii* is more closely related to *I. crenata* [2,3]. In our study, *I. sugerokii* and *I. yunnanensis* are more closely related. In the studies of Cuénoud et al. [2] and Manen et al. [3], *I. opaca* and *I. x attenuata* are placed in the same group in the phylogenetic trees resulting from both nuclear and plastid markers. In addition, *I. rugosa*, *I. aquifolium*, *I. leucoclada*, *I. x makinoi*, and *I. integra* were placed in the same group. Those results are consistent with our findings. Furthermore, our phylogenetic tree (Figure 1) was consistent with the traditional classifications of Galle [9], but also revealed some deviations. First, similar to Galle's classification, *I. integra*, *I. leucoclada*, and *I. aquifolium* were also closely related in our data. According to Galle, they belong together in subgenus *Aquifolium*, section *Aquifolium*, series *Aquifoliodes*. Second, *I. x makinoi*, *I. x meserveae*, and *I. rugosa* were closely related in our study. Galle [9] places them under subgenus *Aquifolium*, section *Aquifolium*, series *Hookerianae*. Third, in our study, *I. vomitoria* was not closely related to any other species in our collection (the highest allele similarity is 17.45% with *I. x meserveae*). This species is also classified separately in Galle's classification, under subgenus *Aquifolium*, series *Vomitoriae*, along with one other species, *I. fuertisiana* (not in our collection). Our GBS data also revealed some discrepancies with previous classifications. Galle [9] places *I. x attenuata*, *I. glabra*, *I. opaca*, *I. sugerokii*, and *I. yunnanensis* together in the subgenus *Prinos*, section *Paltoria*, series *Cassinoides*. However, in the phylogenetic tree shown above, *I. x attenuata* and *I. opaca* were placed separately from the others. Furthermore, *I. serrata* and *I. macrocarpa*, both belonging to different sections, were placed close to *I. glabra*, *I. yunnanensis*, and *I. sugerokii*.

The genome size of *I. aquifolium* (included in our collection) is previously reported to be 2.30 pg/2C [32] as measured by Feulgen densitometry, an older and less accurate technique. The results for *I. aquifolium* in our study were lower, with 1.85 ± 0.04 pg/2C for the cultivar 'Angustifolia' and 1.90 ± 0.07 for the cultivar 'Crispa'. The genome sizes of other accessions in our collection have not been previously reported. We determined the chromosome numbers of several species in our collection: *I. aquifolium*, *I. x attenuata*, *I. crenata*, *I. glabra*, *I. integra*, *I. opaca*, *I. vomitoria*, *I. rugosa*, *I. serrata*, and *I. leucoclada* [9,16]. Fluorescence microscopy images of stained nuclei are shown in Figure 3. The genome sizes and chromosome numbers for all analyzed accessions in the collection are presented in Table 1. For most species, our data confirmed the results of previous studies. All diploid species studied in this research had $2n = 40$ chromosomes. In addition, we observed $2n = 40$ for *I. crenata* and *I. leucoclada* (confirmed by Galle, 1997 for *I. crenata* [9]), which does not agree with Sugiura [13] and Fedorov [14], who reported a chromosome number of $2n = 34$ for both species (found in the Chromosome Count Database (CCDB)) [16]. The largest reported chromosome number in the Chromosome Count Database is that of *I. pedunculosa* Miq. ($2n = 120$). In our collection, a larger chromosome number of $2n = 160$ was observed for *I. yunnanensis* var. *gentilis* (IL104). This species also had the largest genome size in our collection, at 8.01 ± 0.18 pg/2C.

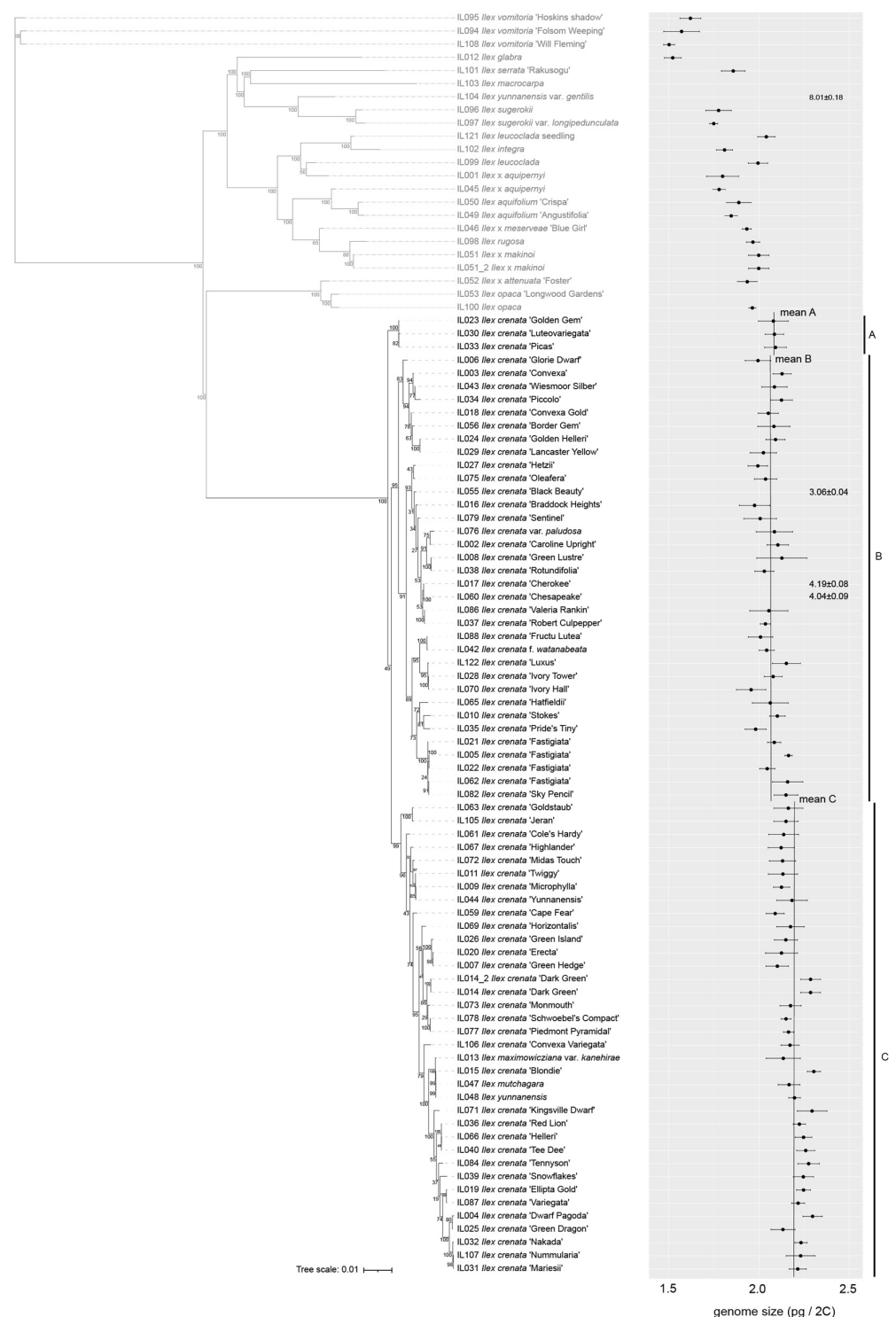


Figure 1. Phylogenetic tree of the *Ilex* collection along with the genome size (pg/2C-value). The clades containing non-*Ilex crenata* spp. are depicted in grey and the clades containing *I. crenata* accessions in black. The three clades within *I. crenata* are indicated on the right (Clade A, Clade B, and Clade C). For every genotype a dot represents the mean genome size ($n = 6$) and the horizontal bar the standard deviation, as measured using flow cytometry. Vertical lines labeled mean A, mean B, and mean C indicate the mean genome sizes per clade (polyploids not included) of 2.09 ± 0.006 ; 2.07 ± 0.05 , and 2.19 ± 0.06 pg/2C $\pm$ sd for clade A, B, and C, respectively. The genome size of genotypes with a genome size exceeding the scale are depicted on the right.

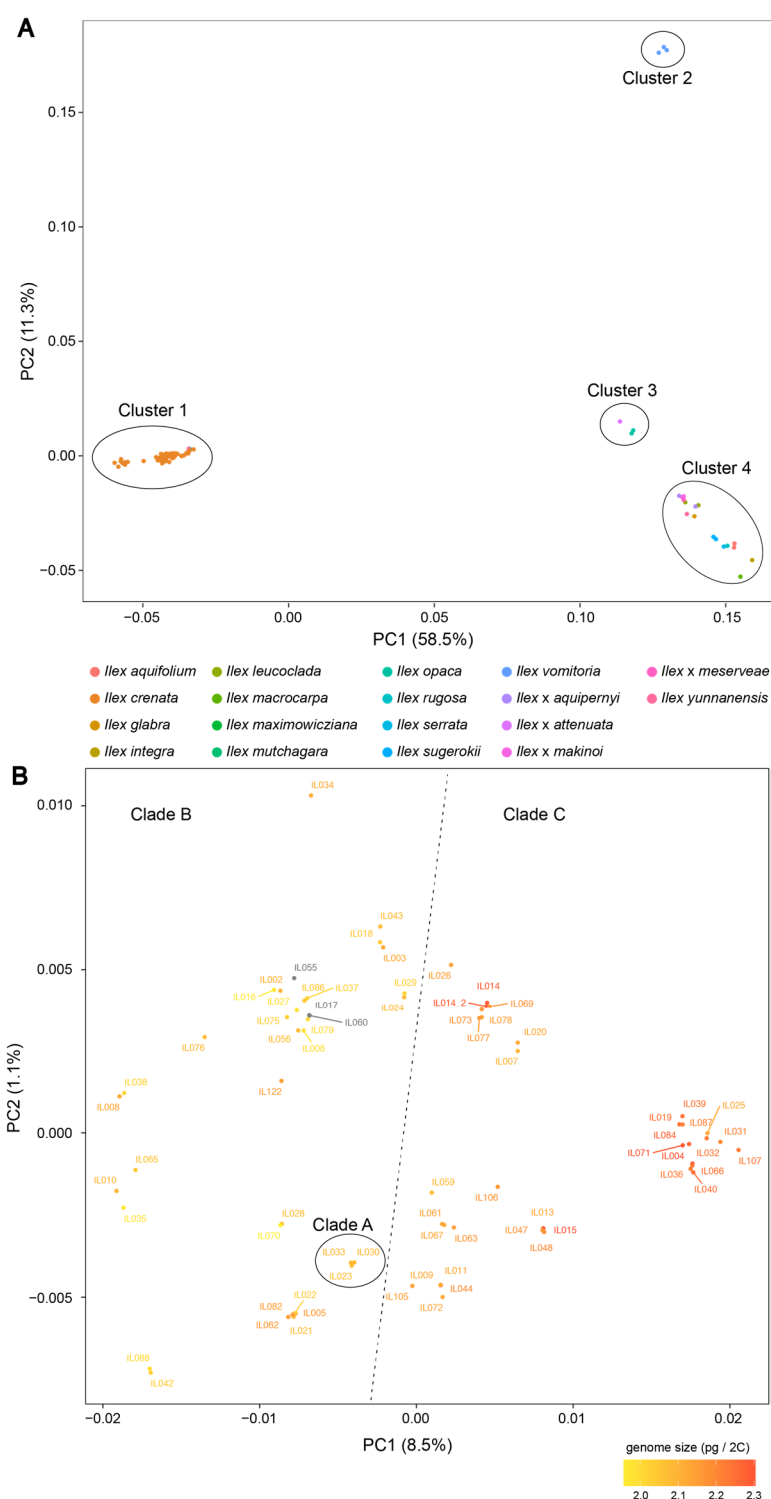


Figure 2. PCoA of the 94 accessions of the *Ilex* collection (A) and the subset of 72 *I. crenata* accessions (B). (A) Samples are colored by species and grouped in four clusters. (B) Samples are false-colored according to their genome size (range 1.98–2.30 pg/2C), while the polyploids are pictured in gray. Accessions of Clade A are grouped using a circle. The dotted line separates accessions of Clade B and Clade C.

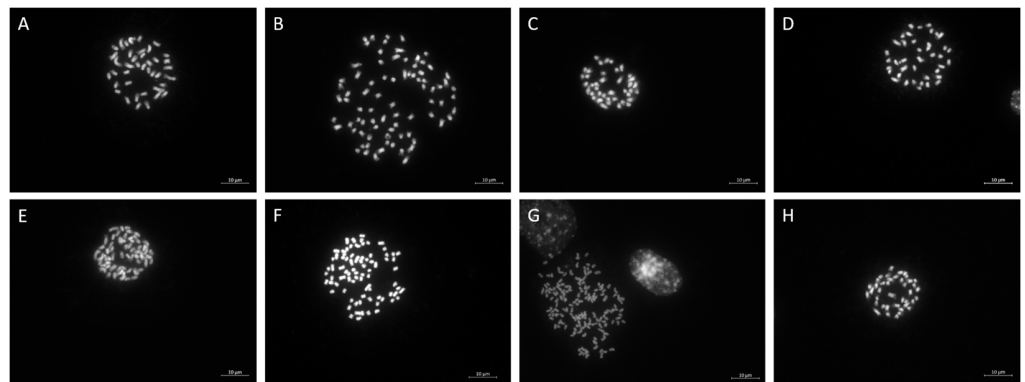


Figure 3. Fluorescence microscopy images of stained nuclei reveal chromosome numbers of *Ilex* spp. (A) *I. crenata* 'Braddock Heights', $2n = 40$; (B) *I. crenata* 'Cherokee', $2n = 80$; (C) *I. crenata* 'Fastigiata', $2n = 40$; (D) *I. x makinoi* 'Hara', $2n = 40$; (E) *I. crenata* 'Black Beauty', $2n = 60$; (F) *I. crenata* 'Chesapeake', $2n = 80$; (G) *I. yunnanensis* var. *gentilis*, $2n = 160$; and (H) *I. leucoclada*, $2n = 40$.

3.3. Genetic Relationships of *Ilex Crenata* Cultivars

When studying the phylogenetic tree (Figure 1), combined with the pairwise allele similarities of the different *I. crenata* accessions (Supplementary Materials), a segregation into three groups emerged within the species of *I. crenata*, which was also observed based on the number of common loci and based on the allele similarity. Clade A consists of *I. crenata* 'Golden Gem' (IL023), *I. crenata* 'Luteovariegata' (IL030), and *I. crenata* 'Picas' (IL033) (bootstrap value of 100%). Clade B contains 34 *I. crenata* genotypes (bootstrap value of 95%), and Clade C contains 36 *I. crenata* genotypes. Clade B has three highly supported subgroups (bootstrap values of 83%, 93% and 99%), while Clade C has some highly supported subgroups but also some genotypes (such as *I. crenata* 'Monmouth' (IL073) and *I. crenata* 'Tee Dee' (IL040)) with less clear phylogenetic positions. The PCoA of *I. crenata* (Figure 2B) roughly confirmed this structure, with clear separation of clusters A and B from C. Contrary to clusters B and C, there was only very little variation within Clade A. In Figure 1, the color gradient shows the distribution of the genome size of *I. crenata*. For all analyzed samples, the mean genome size and standard deviation are shown in Table 1. The genome size of the *I. crenata* genotypes varied from 1.98 pg/2C to 2.30 pg/2C (all diploids $2n = 2x$, data shown in Table 1), with three outliers: *I. crenata* 'Black Beauty' (3.06 ± 0.04 pg/2C, triploid: $2n = 3x = 60$), *I. crenata* 'Chesapeake' (4.04 ± 0.09 pg/2C, tetraploid: $2n = 4x = 80$), and *I. crenata* 'Cherokee' (4.19 ± 0.08 pg/2C, tetraploid: $2n = 4x = 80$). The observed genome size variation in *I. crenata* (1.16-fold difference) was larger than generally expected within a species. Interestingly, the genome sizes were more concise when calculated in accordance with the observed genetic clade structure. The phylogenetic clades had a significant difference in genome size, as depicted in the right panel of Figure 1. Clade A had a mean genome size of 2.09 pg/2C ± 0.006 , Clade B had a mean genome size of 2.07 pg/2C ± 0.05 , and Clade C contains the *I. crenata* samples with the largest genomes (mean of 2.19 pg/2C ± 0.06). All three clades showed little variation within them and have low standard deviations, which is the normal variation expected within a species. There was, however, a significant difference in the mean genome size of Clade B and Clade C (p -value = 7.44×10^{-13}). As Clade A only comprises three genotypes, statistical differences were not calculated for this clade.

Intraspecific genome size variation should be interpreted with caution, as illustrated by Noirot et al. (2000) [33] and Greilhuber (2005) [34]. Noirot et al. showed that the use of an internal standard can reduce, but not eliminate, measurement errors [35]. As *Ilex* is a dioecious plant and sex chromosomes can differ in size, a possible link between genome size and sex was investigated but not found, indicating that the observed variations in genome size and grouping in clades were not due to the sex chromosomes. A more plausible explanation is that the genome size differences are linked to the evolutionary or

breeding history of the plants. However, information about the geographic origin, breeding parents, and year of discovery of the different cultivars in this study is scarce and did not lead to unraveling the origin of the observed clades. In the species *I. polyneura*, the emergence of two groups and some in between genotypes was described by Yao et al. using RADseq population genetic analyses [4]. A link between genome size and evolution has been proposed in maize [36,37]. Bilinsk et al. observed small differences in genome size in maize and described a negative correlation between genome size and cell production rate, as well as a negative correlation between cell production rate and flowering time. Their research suggested that a mechanistic relationship between genome size, cell production, and developmental rate may lead to differences in optimal flowering times across altitudes, thus affecting genome size [36].

3.4. Genetic Diversity via Crossing and Hybridization

In addition to identifying very closely related genotypes, GBS-derived calculations of pairwise common loci and allele similarities also make it possible to identify hybrids, while only considering a phylogenetic tree often does not [30]. In this study, we were able to identify IL051 *I. x makinoi* as the result of a spontaneous hybridization event between *I. leucoclada* (IL099) and *I. rugosa* (IL098), and *I. x meserveae* ‘Blue Girl’ (IL046) as the result of a manual cross between *I. rugosa* (IL098) and *I. aquifolium* (IL050) (Figure 4). Furthermore, the table with common loci and allele similarities in the supplemental materials (Figure S1) shows the known hybrid origin of *I. x aquipernyi* (IL045 and IL001) as being the result of a manual cross of *I. aquifolium* and *I. pernyi* (not in dataset) as parents for *I. x aquipernyi* (IL045). This accession has many loci in common with *I. aquifolium* (IL049 and IL050). For *I. x aquipernyi* (IL001), however, the number of common loci with *I. aquifolium* (IL049 and IL050) is not as high as for *I. x aquipernyi* (IL045). It is possible that *I. x aquipernyi* (IL001) has a higher similarity to its other parent, *I. pernyi* (not included in the data).

	IL050 <i>Ilex aquifolium</i> ‘Crispa’	IL046 <i>Ilex x meserveae</i> ‘Blue Girl’	IL098 <i>Ilex rugosa</i>	IL051 <i>Ilex x makinoi</i>	IL099 <i>Ilex leucoclada</i>
IL050	100	85.34	34.08	36.30	35.22
IL046	85.34	100	82.79	73.35	41.57
IL098	34.08	82.79	100	82.86	41.59
IL051	36.30	73.35	82.86	100	67.90
IL099	35.22	41.57	41.59	67.90	100

Figure 4. Parental relations of hybrids *I. x meserveae* (IL046) (*I. aquifolium* × *I. rugosa*) and *I. x makinoi* (IL051) (*I. leucoclada* × *I. rugosa*). The high allele similarity of GBS fingerprints reflects the relationship of *I. x meserveae* (IL046) to *I. aquifolium* ‘Crispa’ (IL050) (85.34%) and *I. rugosa* (IL098) (82.79%), whereas *I. aquifolium* ‘Crispa’ and *I. rugosa* have an allele similarity of only 34.08%. Similarly, the relation of *I. x makinoi* (IL051) to *I. leucoclada* (IL099) and *I. rugosa* (IL098) can be derived from their high allele similarities of 82.86% and 67.90%, respectively, whereas *I. leucoclada* and *I. rugosa* have a much lower allele similarity of 41.59%.

I. crenata ‘Pride’s Tiny’ (IL035) appears to have been derived from a spontaneous or manual hybridization event involving *I. sugerokii* var. *longipedunculata* (IL097) in its ancestral lineage. The allele similarities within *I. crenata* reflect the clades of the phylogenetic tree. Note the position of *I. crenata* ‘Dark Green’ (IL014), *I. crenata* ‘Monmouth’ (IL073), *I. crenata* ‘Horizontalis’ (IL069), *I. crenata* ‘Schwoebel’s Compact’ (IL078), and *I. crenata* ‘Piedmont

Pyramidal' (IL077). They cluster with genotypes in Clade C of the phylogenetic tree, but they also had a high allele similarity with *I. crenata* 'Convexa' (IL003), *I. crenata* 'Convexa Gold' (IL018), and *I. crenata* 'Wiesmoor Silber' (IL043). The same pattern is observed for *I. crenata* 'Twiggy' (IL011), *I. crenata* 'Midas Touch' (IL072), *I. crenata* 'Microphylla' (IL009), and *I. crenata* 'Yunnanensis' (IL044), all of which had high allele similarities with *I. crenata* f. *watanabeana* Mak. (IL042) and *I. crenata* 'Fructu Lutea' (IL088). Furthermore, *I. crenata* 'Green Island' (IL026) is part of Clade C, but has a high allele similarity with *I. crenata* 'Golden Helleri' (IL024) and *I. crenata* 'Lancaster Yellow' (IL029). These 'in between' genotypes could be the result of past intraspecific hybridization events between genotypes of Clade B and Clade C.

The variety of commercial *I. crenata* cultivars is rather limited, comprised mostly of a small selection of cultivars that are multiplied. The PCoA analyses revealed the breadth of currently unexploited genetic diversity. There were no genotypes in our collection that are positioned in the spaces between clusters, indicating a great potential for crosses between clusters. Interspecific hybridization in *Ilex* is usually more complicated than intraspecific crossing, and chromosome numbers can be used as a basis for estimating the rate of success [9]. Nevertheless, hybridization has led to many commercially important cultivars of different *Ilex* species in the past [9], such as *I. x meserveae* cultivars. Our data even confirm the possibility of inter-species hybridization. To obtain cultivars that are more robust to challenges such as high soil pH and black root rot, the knowledge from this paper needs to be combined with knowledge about phenotypic traits such as tolerance and resistance. In the future, both intra- and interspecific hybridization represent promising approaches to creating novel genetic diversity.

4. Conclusions

Despite the popularity of *I. crenata* as an ornamental shrub, a handful of cultivars are commercially dominant. The aim of this study was to investigate the genetic diversity within the species *I. crenata*, and between *I. crenata* and a few other hardy *Ilex* species. Genome size measurements, chromosome numbers, and genetic fingerprinting were used to unravel the genetic identity, genetic relationships, and genetic diversity of 72 *I. crenata* accessions and 22 accessions of other species. Our data showed that *I. crenata* has a high intra-specific genetic diversity. This was reflected in the differences in genome size of diploid *I. crenata* genotypes related to three phylogenetic clades identified through GBS genetic fingerprinting. Furthermore, the genetic fingerprinting data showed that large parts of this genetic variation might be interesting for commercial use. Interspecific hybridization or intraspecific crossings are a possible strategy to further enlarge the genetic diversity in the commercial assortment of *I. crenata*. This research is valuable for breeding programs, as it allows deliberately selecting for genetically diverse parental material, creating novel genetic diversity, and screening for novel morphological traits.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/horticulturae9040485/s1>, Figure S1: Common loci and allele similarities of the complete collection.

Author Contributions: Conceptualization, E.G., L.L., A.H. and T.R.; methodology, E.G., L.L., A.H., T.R. and K.V.L.; formal analysis, E.G. and A.H.; investigation, E.G.; resources, J.V.M.; data curation, E.G. and A.H.; writing—original draft preparation, E.G.; writing—review and editing, L.L., A.H., T.R. and K.V.L.; visualization, A.H. and E.G.; supervision, L.L. and E.S.; funding acquisition, E.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is part of a Baekeland mandate (HBC.2020.2847) awarded by Flanders Innovation and Entrepreneurship (VLAIO).

Data Availability Statement: The raw PE-150 read data, as well as the merged preprocessed read data were submitted to NCBI's Sequence Read Archive (SRA) under BioProject number PRJNA895194.

Acknowledgments: The authors would like to sincerely thank the arboreta of Bokrijk (Genk, Belgium) and Het Leen (Eeklo, Belgium) for providing plant material to start our collection. Special thanks to Reine Audenaert for DNA extractions, and Magali Losschaert and Laura Schelfhout for chromosome counting. We would also like to thank Roger Dobbelaere and Rina Vanhaecke for their excellent management of the plants in the ILVO collection, and Miriam Levenson for English-language editing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. POWO. Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. 2022. Available online: <https://powo.science.kew.org/> (accessed on 27 February 2023).
2. Cuénoud, P.; Del Pero Martinez, M.A.; Loizeau, P.-A.; Spichiger, R.; Andrews, S.; Manen, J.-F. Molecular phylogeny and biogeography of the genus *Ilex* L. (Aquifoliaceae). *Ann. Bot.* **2000**, *85*, 111–122. [\[CrossRef\]](#)
3. Manen, J.-F.; Barriera, G.; Loizeau, P.-A.; Naciri, Y. The history of extant *Ilex* species (Aquifoliaceae): Evidence of hybridisation within a Miocene radiation. *Mol. Phylogenetics Evol.* **2010**, *57*, 961–977. [\[CrossRef\]](#)
4. Yao, X.; Lu, Z.; Song, Y.; Hu, X.; Corlett, R.T. A chromosome-scale genome assembly for the holly (*Ilex polyneura*) provides insights into genomic adaptations to elevation in Southwest China. *Hortic. Res.* **2022**, *9*, uhab049. [\[CrossRef\]](#)
5. Dudley, T.R.; Eisenbeiss, G.R. *International Checklist of Cultivated Ilex: Part 2, Ilex crenata Thunberg ex. J.A. Murray*; U.S. Department of Agriculture: Washington, DC, USA; U.S. National Arboretum Contribution: Washington, DC, USA, 1992; Volume 6, p. 91.
6. Xu, K.-W.; Wei, X.-F.; Lin, C.-X.; Zhang, M.; Zhang, Q.; Zhou, P.; Fang, Y.-M.; Xue, J.-Y.; Duan, Y.-F. The chromosome-level holly (*Ilex latifolia*) genome reveals key enzymes in triterpenoid saponin biosynthesis and fruit color change. *Front. Plant Sci.* **2022**, *13*, 982323. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Kong, B.L.-H.; Nong, W.; Wong, K.-H.; Law, S.T.-S.; So, W.-L.; Chan, J.J.-S.; Zhang, J.; Lau, T.-W.D.; Hui, J.H.-L.; Shaw, P.-C. Chromosomal level genome of *Ilex asprella* and insight into antiviral triterpenoid pathway. *Genomics* **2022**, *114*, 110366. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Yao, X.; Zhang, F.; Corlett, R.T. Utilization of the Hollies (*Ilex* L. spp.): A Review. *Forests* **2022**, *13*, 94. [\[CrossRef\]](#)
9. Galle, F.C. *Hollies: The genus Ilex*; Timber Press, Inc.: Portland, OR, USA, 1997; p. 573.
10. Molnar, T.; Capik, J.M. The Rutgers University woody ornamentals breeding program: Past, present, and future. *Acta Hort.* **2013**, *990*, 11. [\[CrossRef\]](#)
11. Zhang, L.; Cao, B.; Bai, C. New reports of nuclear DNA content for 66 traditional Chinese medicinal plant taxa in China. *Caryologia* **2013**, *66*, 375–383. [\[CrossRef\]](#)
12. Bai, C.; Alverson, W.S.; Follansbee, A.; Waller, D.M. New reports of nuclear DNA content for 407 U.S. plant species. *Ann. Bot.* **2012**, *110*, 1623–1629. [\[CrossRef\]](#)
13. Sugiura, T. A list of chromosome numbers in angiospermous plants. *Shokubutsugaku Zasshi* **1931**, *45*, 353–355. [\[CrossRef\]](#)
14. Fedorov, A. *Chromosome Numbers of Flowering Plants*; Reprint 1974; Academy of Sciences of the USSR: Moscow, Russia; Komarov Botanical Institute: Leningrad, Russia, 1969.
15. Goldblatt, P. New or noteworthy chromosome records in the angiosperms. *Ann. Mo. Bot. Gard.* **1976**, *63*, 889–895. [\[CrossRef\]](#)
16. Rice, A.; Glick, L.; Abadi, S.; Einhorn, M.; Kopelman, N.M.; Salman-Minkov, A.; Mayzel, J.; Chay, O.; Mayrose, I. The Chromosome Counts Database (CCDB)—A community resource of plant chromosome numbers. *New Phytol.* **2014**, *206*, 7. [\[CrossRef\]](#)
17. Darlington, C.D.; Wylie, A.P. *Chromosome Atlas of Flowering Plants*; Allen & Unwin: Crows Nest, Australia, 1955.
18. Poland, J.A.; Rife, T.W. Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* **2012**, *5*, 92–102. [\[CrossRef\]](#)
19. Van Oost, E.; Leus, L.; De Rybel, B.; Van Laere, K. Determination of genetic distance, genome size and chromosome numbers to support breeding in ornamental *Lavandula* species. *Agronomy* **2021**, *11*, 2173. [\[CrossRef\]](#)
20. Greilhuber, J.; Dolezel, J.; Lysák, M.A.; Bennet, M.D. The origin, evolution and proposed stabilization of the terms ‘genome size’ and ‘C-value’ to describe nuclear DNA contents. *Ann. Bot.* **2005**, *95*, 255–260. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Doležel, J.; Greilhuber, J.; Lucretti, S.; Meister, A.; Lysák, M.A.; Nardi, L.; Obermayer, R. Plant genome size estimation by flow cytometry: Inter-laboratory comparison. *Ann. Bot.* **1998**, *82*, 17–26. [\[CrossRef\]](#)
22. Lysak, M.A.; Dolezel, J. Estimation of nuclear DNA content in *Sesleria* (Poaceae). *Caryologia* **1998**, *51*, 123–132. [\[CrossRef\]](#)
23. Kirov, I.; Divashuk, M.; Van Laere, K.; Soloviev, A.; Khrustaleva, L. An easy “SteamDrop” method for high quality plant chromosome preparation. *Mol. Cytogenet.* **2014**, *7*, 21. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Kirov, I.; Khrustaleva, L.; Van Laere, K.; Meeus, S.; Romanov, D.; Soloviev, A.; Fesenko, I. DRAWID: User-friendly java software for chromosome measurements and idiogram drawing. *Comp. Cytogenet.* **2017**, *11*, 10. [\[CrossRef\]](#)
25. Doyle, J.J.; Doyle, J.L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **1987**, *19*, 11–15.
26. Schaumont, D. GBprocess: Genotyping-by-Sequencing Data Processing Toolkit. Available online: <https://gitlab.com/ilvo/GBprocess>. (accessed on 27 May 2021).
27. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10–12. [\[CrossRef\]](#)
28. Zhang, J.; Kobert, K.; Flouri, T.; Stamatakis, A. PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **2013**, *30*, 614–620. [\[CrossRef\]](#) [\[PubMed\]](#)

29. Hapke, A.; Thiele, D. GIBPSs: A toolkit for fast and accurate analyses of genotyping-by-sequencing data without a reference genome. *Mol. Ecol. Resour.* **2016**, *16*, 11. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Van Poucke, K.; Haegeman, A.; Goedefroit, T.; Focquet, F.; Leus, L.; Jung, M.H.; Nave, C.; Redondo, M.A.; Husson, C.; Kostov, K.; et al. Unravelling hybridization in *Phytophthora* using phylogenomics and genome size estimation. *IMA Fungus* **2021**, *12*, 16. [\[CrossRef\]](#)
31. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [\[CrossRef\]](#)
32. Mowforth, M.A. *Variation in Nuclear DNA Amounts in Flowering Plants: An Ecological Analysis*; University of Sheffield: Sheffield, UK, 1986.
33. Noirot, M.; Barre, P.; Louarn, J.; Duperray, C.; Hamon, S. Nucleus-cytosol interactions—A source of stoichiometric error in flow cytometric estimation of nuclear DNA content in plants. *Ann. Bot.* **2000**, *86*, 309–316. [\[CrossRef\]](#)
34. Greilhuber, J. Intraspecific variation in genome size in angiosperms: Identifying its existence. *Ann. Bot.* **2005**, *95*, 91–98. [\[CrossRef\]](#)
35. Noirot, M.; Poncet, V.; Barre, P.; Hamon, P.; Hamon, S.; De Kochko, A. Genome size variations in diploid African *Coffea* species. *Ann. Bot.* **2003**, *92*, 709–714. [\[CrossRef\]](#)
36. Bilinski, P.; Albert, P.S.; Berg, J.J.; Birchler, J.A.; Grote, M.N.; Lorant, A.; Quezada, J.; Swarts, K.; Yang, J.; Ross-Ibarra, J. Parallel altitudinal clines reveal trends in adaptive evolution of genome size in *Zea mays*. *PLoS Genet.* **2018**, *14*, e1007162. [\[CrossRef\]](#)
37. Díez, C.M.; Gaut, B.S.; Meca, E.; Scheinvar, E.; Montez-Hernandez, S.; Eguiarte, L.E.; Tenaillon, M.I. Genome size variation in wild and cultivated maize along altitudinal gradients. *New Phytol.* **2013**, *199*, 264–276. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.