Tree Genetics & Genomes (2026) 22:13
https://doi.org/10.1007/s11295-026-01735-y

RESEARCH

Design and validation of an SNP-based assay for detecting *Juglans ailantifolia* introgression into *Juglans cinerea* populations or living collections

Martin Williams[1] · Manuel Lamothe[2] · Aziz Ebrahimi[3] · Berni M. van der Meer[1] · Mojtaba Zamani Faradonbeh[3] · Anna O. Conrad[4] · James Warren[5] · Carolyn C. Pike[6] · Douglass F. Jacobs[3] · Nathalie Isabel[2]

Received: 3 October 2025 / Revised: 6 October 2025 / Accepted: 20 March 2026 / Published online: 19 May 2026

Abstract

Butternut (*Juglans cinerea* L.) is a rapidly declining tree species threatened by *Ophiognomonia clavigignenti-juglandacearum* (Oc-j), the causal agent of butternut canker. The introduction of Japanese walnut (*Juglans ailantifolia* Carrière), which hybridizes with butternut and exhibits greater resistance to the disease, presents both a conservation challenge and a potential avenue for genetic improvement. Identifying non-admixed butternuts is critical for conservation efforts and breeding programs aimed at preserving the species' genetic integrity while exploring resistance strategies. However, morphological identification of first and advanced generation hybrids is unreliable, necessitating the development of efficient detection tools. In this study, we developed and validated a cost-effective, high-throughput SNP genotyping tool using the MassARRAY platform to distinguish between *J. cinerea*, *J. ailantifolia*, and their hybrids. The panel consists of 28 nuclear SNPs and two chloroplast SNPs and demonstrates strong concordance with genotyping-by-sequencing (GBS) results using nearly 12,000 SNPs ($R^2 = 0.98$), indicating high predictive accuracy. Field testing this genotyping tool to 1,269 trees from natural stands and breeding programs in Canada and the U.S. revealed substantial hybridization, particularly in U.S. plantations, where hybrids comprised up to 18% of tested trees. Chloroplast analysis confirmed asymmetrical hybridization patterns, with *J. ailantifolia* acting predominantly as the maternal parent. This SNP panel provides a valuable tool for conservation, breeding, and regulatory applications by enabling rapid and accurate hybrid identification. Its implementation will help guide recovery strategies for butternut, ensuring that conservation efforts are informed by precise genetic data.

Keywords Species at risk · Resistance · Breeding · Genotyping · Butternut · Walnut · Single nucleotide polymorphism · Juglandaceae · MassARRAY

Martin Williams and Manuel Lamothe contributed equally to this work.

Communicated by E. Dirlewanger

✉ Martin Williams
Martin.Williams@nrcan-rncan.gc.ca

[1] Canadian Forest Service, Natural Resources Canada, Atlantic Forestry Centre, Fredericton, NB E3B 5P7, Canada

[2] Canadian Forest Service, Natural Resources Canada, Laurentian Forestry Centre, Quebec city, QC G1R 4K7, Canada

[3] Hardwood Tree Improvement and Regeneration Center (HTIRC), Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN, USA

[4] U.S. Forest Service, Northern Research Station, Delaware, OH, USA

[5] U.S. Forest Service, Northern Research Station, South Burlington, VT, USA

[6] U.S. Forest Service, State, Private and Tribal Forestry, South Burlington, VT, USA

Introduction

Juglans cinerea, L. (butternut; JC), also called white walnut, is one of approximately 23 species of walnut distributed around the world (Mu et al. 2020). This short-lived hardwood species is among the most cold-tolerant in the walnut genus and is prized for its edible nuts, highly workable wood, and ecological value to wildlife. It has been especially treasured by First Nations communities for its ethnobotanical and cultural significance (Lawrence et al. 2017).

Largely due to the introduction of the pathogen *Ophiognomonia clavigignenti-juglandacearum* (Oc-j), *J. cinerea* has undergone a rapid decline in the last decades and is currently listed as endangered in the IUCN Red List of Threatened Species (Stritch and Bartow 2019), by the Species at Risk Act in Canada (Committee on the Status of Endangered Wildlife in Canada (COSEWIC) 2017), and as vulnerable, imperiled and critically imperiled in 23 US states by NatureServe Explorer (NatureServe 2025).

J. cinerea and *Juglans nigra* L. (black walnut; JN) are the only two *Juglans* species native to Eastern North America. Despite their sympatric occurrence, there have been no substantiated reports of hybridization between the two species. Chlorotyping studies (Stanford et al. 2000; Aradhya et al. 2007; Dong et al. 2017; Ebrahimi et al. 2019) have all placed *J. cinerea* chloroplast genomes in the *Rhysocaryon* section with *J. nigra*. This was likely the result of an ancient chloroplast capture event (Ebrahimi 2022; Guzman-Torres et al. 2023) since both morphological studies (Dode 1909; Manning 1978; Lu et al. 1999) and nuclear marker-based phylogenetics (Stanford et al. 2000; Dong et al. 2017; Mu et al. 2020), including recent genome sequencing (Ebrahimi 2022; Guzman-Torres et al. 2023), have all placed *J. cinerea* in its own section, the *Trachycaryon*, with the *Cardiocaryon* as a sister section, which includes the *Juglans ailantifolia* Carrière (Japanese walnut; JA).

Juglans ailantifolia, native to Japan (Owhi 1965), was introduced to North America in the late 1800s (Bixby 1919) and widely planted as a horticultural tree in the 1900s. It can hybridize with both *J. cinerea* and *J. nigra* (Michler et al. 2006; Brennan et al. 2020) and due to this hybridization potential, it is listed as a potential genetic invasion threat (Environment Canada 2010). However, *Juglans cinerea X ailantifolia* hybrids could represent a potential opportunity for protecting butternut's genetic diversity as these hybrids show resistance to the canker caused by Oc-j (Boraks and Broders 2014; Brennan et al. 2020). Heritable resistance seems to be very limited (Ostry and Moore 2008; LaBonte et al. 2015) but not completely absent in wild-type *J. cinerea.* Recent evidence shows that some non-admixed butternut families do indeed have a degree of resistance, but most resistant families are hybrids (Brennan et al. 2020; Conrad et al. 2026). Whether that level of resistance is adequate to overcome the pathogen remains to be determined.

The rapid decline of *J. cinerea* has created an urgency to implement a tree improvement strategy to prevent extinction (Michler et al. 2006; Pike et al. 2020). Currently, few options are available, and one viable option is using backcross breeding with *J. ailantifolia* to impart or enhance the pathogen resistance that *J. cinerea* lacks. The American Chestnut Foundation (TACF) has already utilized this approach, namely introgressive hybridization, with Chinese chestnut to increase fungal blight resistance in American chestnut (Powell et al. 2019; Sandercock et al. 2024). This introgressive hybridization approach consists of repeated hybridization and backcrossing to transfer alleles of interest, in this case from Chinese chestnut, which is more resistant to the blight followed by backcrossing to the American chestnut to minimize the loss the native species genetic contribution. However, the hybrid program is still looking at whether sufficient resistance has been introgressed to protect them against the pathogen on the landscape.

Over the last 20 years, teams across the range have collected *J. cinerea* (from seed or by grafts) from the landscape to maintain this material in living collections. For example, the Forest Gene Conservation Association (FGCA) in Ontario, Canada selected putatively resistant trees from the landscape, grafted them to *J. nigra* rootstock and grew them within their breeding orchards with the goal of generating more resistant progeny through open pollination. Similar efforts have been done in the US (i.e. Bell (R9), Hardwood Tree Improvement and Regeneration Center (HTIRC) in West Lafayette, Indiana). However, lingering healthy trees are often hybrids, due to their increased resistance. If hybrids are erroneously selected for open pollinated *J. cinerea* seed orchards, they can obscure evidence of disease resistance within pure *J. cinerea*. Due to the difficulty of morphologically detecting hybrids between these two congeners, various non-genetic and genetic screening tools have been developed. Farlee et al. (2010) developed a visual screening method based on eleven traits to guide resource professionals, nut growers and landowners in conserving this declining species. Although it is a valuable resource for identifying pure *J. cinerea* trees, many individuals can still be misidentified or not fully assessed due, in part, to the lack of certain characters (flowering, seed) at the time of assessment. Misidentification and incomplete assessments are particularly problematic for selecting non-admixed disease-resistant trees. Different types of nuclear or organellar genetic markers such as restriction fragment length polymorphisms (Fjellstrom and Parfitt 1994), microsatellite markers (Hoban et al. 2008), chloroplastic Cleaved Amplified Polymorphic Sequences (CAPS; McCleary et al. 2009; Zhao and Woeste 2010), and sequence characterized

amplified regions markers from randomly amplified polymorphic DNA (Zhao and Woeste 2010) have also been developed and used to assess genetic composition of putatively pure *J. cinerea* but can be time-consuming, low throughput and can suffer from a lack of resolution.

With the advent of next generation sequencing technology, novel approaches have been developed for identifying interspecific markers, such as restriction site-associated DNA sequencing methods, which utilize restriction enzymes to reduce genome complexity before sequencing (e.g., Genotyping-by-Sequencing; GBS), or to produce Single Nucleotide Polymorphism (SNP) marker panels (e.g., MassARRAY; MA) to rapidly genotype samples at multiple loci, enabling the development of tools for either species or genotype detection and identification (e.g. Meirmans et al. 2010; Isabel et al. 2013; Godbout et al. 2017). The MassARRAY System combines PCR amplification with mass spectroscopy which provides a high-throughput genotyping approach with multiplexing capabilities (30-40 SNPs/panel). Compared to other approaches, the SNP marker panels are cost-efficient, high-throughput, and do not require complex downstream analysis, making them applicable to non-specialist end-users (for more details see Godbout et al. 2017) that simultaneously allows for both nuclear and organellar genetic marker assessment. Additionally, even in the case of low-quality DNA, species identification remains possible despite some missing data, as enough informative markers can often still be recovered.

In this paper, we had three main objectives: (1) develop and select a high-quality SNPs panel to distinguish *J. cinerea*, *J. ailantifolia*, and their hybrids, (2) Measure the assignment efficiency of the interspecific panel (limited number of SNPs) by comparing it to GBS genotyping approach (thousands of SNPs), and (3) field test the panel to assess the extent of admixture of trees collected from plantations and wild stands in Canada and the US.

Materials and methods

Sample collection

Leaflet samples were collected fresh and stored at -80°C or at room temperature after drying the leaves with silica packs. A total of 1,497 samples from unique trees were used in this study. Canadian samples came from various locations from 3 Canadian provinces (NB, QC, ON) where the species is native. Forest Gene Conservation Association (FGCA) samples were all selected as putatively resistant trees in the province of ON. They were grafted to *J. nigra* root stock and are currently planted in various living orchards in the province of ON. Other Canadian samples were collected from various wild stand in all 3 provinces. US samples, came from diverse sources, including a planting of open-pollinated seedlings from seed collections across the US and from research plantings, which presumably could have included surviving or more resistant material (HTIRC), surviving trees from a decades-long study with commercial nursery-sourced material (MTE), and from a planting established to examine resistance in *J. cinerea* with material sourced from seed orchards (Bell, R9). Matrogrny 'Bell' trees were believed to be *J. cinerea*, with some hybrid trees intentionally included in the study design.

DNA extraction

Since samples were collected and extracted from many collaborators, different DNA extraction procedures were used. Details for samples prepared for GBS can be found in the SNP assay development section. DNA from Canadian samples (FGCA, Canada) were extracted using the Plant DNeasy kit (Qiagen, Germantown, MD, USA) according to the manufacturers protocol except that the lysis step was done for 1 h at 65° C. DNA from US samples were extracted using a modified cetyltrimethylammonium bromide (CTAB) method outlined by Doyle and Doyle (1987).

SNP assay development

The genotyping assay was built from the results of two distinct sources of GBS to maximize the presence of interspecific fixed SNPs on as many chromosomes as possible. The first GBS analysis was conducted on 63 individuals from the *Juglans* genus, including 53 *J. cinerea*, four *J. ailantifolia*, three F1 hybrids between *J. ailantifolia* and *J. cinerea*, three *J. nigra*, all identified as putative members of their respective species (material obtained from Danny Rioux, Canadian Forest Service). DNA was extracted from dried leaf material with the Nucleospin 96 Plant II kit (Macherey-Nagel, Bethlehem, PA, USA) following the manufacturer's protocol for vacuum processing with the following modifications: (a) cell lysis using buffers PL2 and PL3 (lysis step heated for 1 h at 65° C); and (b) elution with Tris-HCl 10 mM pH 8.0 buffer. DNA samples were quantified and assessed for quality with a NanoDrop spectrophotometer and a Qubit fluorometer (ThermoFisher Scientific, Waltham, MA, USA). The GBS libraries were prepared at the Plateforme d'Analyses Génomiques of the Institut de Biologie Intégrative et des Systèmes (IBIS; Laval University, Quebec City, Canada) according to the procedure described in Poland et al. (2012) with *Pst*I and *Msp*I restriction enzymes. Barcoded adapters were ligated to individual DNA. The libraries were size selected on a blue pippin 2% agarose cassette and then amplified by PCR. Each amplified library was

loaded on three different Ion Torrent™ System P1 v3 chips (ThermoFisher Scientific) for sequencing. The resulting sequences from three chips were merged at the end of the process. Adapters were removed and raw reads trimmed to 120 bp using Cutadapt v1.9dev1 (Martin 2011) before being aligned to the *Juglans regia* L. reference genome, Chandler v1.0 (Martinez-Garcia et al. 2016) using CLC Genomic Workbench v5.5.2 (Qiagen, Germantown, MD, USA.). Eleven SNPs (JUGv-001 to JUGv-035) were selected based on their position in the genome and/or the function of the gene they belong to.

The second GBS dataset also used dried leaf segments (approximately 1.5 cm long and 1 cm wide) that were placed in a 96-well plate (1.2 mL) and ground using the GenoGrinder 2000 (Spex CertiPrep, Metuchen, NJ, USA) at 650 strokes/min for 15 to 30 s. This step was crucial for the DNA extractions, conducted using silica columns in 96-well plates and a modified cetyltrimethylammonium bromide (CTAB) method outlined by Doyle and Doyle (1987). The DNA concentration was measured with the Quant-iT Pico-Green dsDNA Kit (Invitrogen, Carlsbad, CA, USA) on a Synergy™ HT plate reader (Bio-Tek Instruments, Winooski, VT, USA) and adjusted to 50 ng/µL. Illumina library were prepared following the method by Poland et al. (2012) using 7.5 µL of DNA in a 30 µL restriction-ligation reaction with *Hind*III-HF, *Mse*I, and T4 DNA Ligase. Individual barcoded libraries were pooled by 96-well plate, cleaned using Ampure XP beads (Beckman-Coulter, Indianapolis, IN, USA), PCR amplified and cleaned again with Ampure XP beads before being loaded onto a Bioanalyzer 2100 in a DNA7500 chip (Agilent, Santa Clara, CA, USA). Libraries were adjusted to 10 nmol before being submitted to the UC Davis Genome Center (Davis, CA) for SR100 sequencing on a HiSeq4000 instrument (Illumina, San Diego, CA, USA). For this GBS, the TASSEL GBS pipeline (Glaubitz et al. 2014) and BWA (Li and Durbin 2009) were used to align 64 bp tags to the Chandler v2.0 genome (Marrano et al. 2020), keeping only tags that appeared at least 10 times in the dataset with a MAPQ score >= 20. Concordance between replicate samples was used to set filtering thresholds. SNP calls with a depth < 5 were set to missing, and samples with more than 90% missing data were discarded. Indels, SNPs with more than two alleles, and SNPs with over 50% missing data were also discarded. Beagle 5.0 (Browning et al. 2018) was used for imputation with window = 12 and step = 4. These filters resulted in a median concordance between replicate samples of 95.9% (mean = 95.5%), a median heterozygous mismatch rate of 3.9% (mean = 4.2%), and a median homozygous mismatch rate of 0.2% (mean = 0.3%). A total of 11,952 SNPs were kept from this dataset and 17 unlinked SNPs (JUGv-102 to JUGv-122) were selected for the assay based on 100% frequency difference between 18 putative *J. ailantifolia* and 30 putative *J. cinerea* samples.

Two species-specific chloroplast CAPS markers (CPS2 and CPS10) developed by McCleary et al. (2009) to distinguish *J. ailantifolia* and *J. cinerea* were also included in the assay. New primers were redesigned based on the alignment of the original CAPS primers with the chloroplast genomes of *J. cinerea*, *J. ailantifolia*, *J. nigra*, *J. regia*, and *Juglans mandshurica* L. (Dong et al. 2017).

Genotyping

Genotyping was performed using the Agena Bioscience iPLEX Gold platform at the Genome Quebec service facility (Montreal, Canada). In short, a multiplex PCR was performed on 20 ng of template genomic DNA in a 5uL reaction mixture containing: 0.1 µL (0.5 U) HotStarTaq DNA polymerase (Qiagen, Germantown, MD, USA), 0.63 µL of 10X HotStarTaq PCR Buffer, 0.33 µL of 25 mM $MgCl_2$, 0.25 µL of 10 mM dNTP mix, 0.55 µL of forward and reverse primer pool (1 µM; see Supplementary Table S1 for all primer sequences) and 1.15 µL of water, with cycling parameters of and initial hold of 95 °C for 15 min, followed by 45 cycles (95 °C for 20s, 56 °C for 30s, 72 °C for 60s), then a final 72 °C hold for 3 min. Shrimp-Alkaline-Phosphatase (SAP) treatment was then performed to remove any unused nucleotides; reactions were added to 0.2 µL of SAP Buffer, 0.3 µL of SAP, and 1.5 µL of water and incubated at 37 °C for 40 min, then 85 °C for 10 min. A primer extension reaction was then performed with 0.94 µL of extension primer mix, 0.2 µL of iPlex Terminator, 0.2 µL of iPlex Buffer, 0.041 µL of iPlex Thermo Sequenase and 0.62 µL of water. The extension cycle parameters were an initial hold of 95 °C for 30s, followed by 45 cycles of (95 °C for 5s, then 5 subsequent rounds of 52 °C for 5s, and 80 °C for 5s, a final hold of 72 °C for 3 min. The products are then desalted using 6 mg of resin and spotted on a 384-point SpectroCHIP (Agena Bioscience, San Diego, CA) using the Chip Prep Module. The distinct masses were determined by MALDI-TOF mass-spectrometry and results were analyzed using MassARRAY Typer Analyser software. The resulting mass heights were clustered with the Mclust v6.1 (Scrucca et al. 2016) in R v4.3.3 (R Core Team 2023).

Validation

A validation plate containing 90 putative samples where 86 were *J. cinerea*, *J. ailantifolia* and hybrids, as well as 4 samples of *J. nigra* were used to validate the SNPs assay results. Species assignments were previously determined either by observed allelic proportions from GBS results or

from microsatellite marker analysis. We classified them as pure *J. cinerea* (JC), pure *J. ailantifolia* (JA), backcross toward *J. cinerea* (BC JC), backcross toward *J. ailantifolia* (BC JA), complex hybrids between these two species (Complex hybrids), or *J. nigra*. In addition to the validation plate samples, a subset of 55 samples (some used in the initial SNP Validation steps and some from the field testing) previously assessed by GBS ($n=11,952$ SNPs) were compared to MassARRAY ($n=26$ SNPs) to estimate how often MassARRAY successfully assigns genotypes in comparison to GBS.

Hybrid analysis

Both Buerkle's hybrid index (Buerkle 2005) using function esth in R package gghybrid v1.0 (Bailey 2024) and STRUCTURE v2.3.4 (Pritchard et al. 2000) were used to estimate the hybridization level of 1,269 samples from various natural origins in Eastern Canada (Ontario, Quebec, and New Brunswick), as well as samples from plantation trees from the USA. Both methods use Markov Chain Monte Carlo (MCMC) for estimating results. For Hybrid Index (HI) we used a burn-in of 2,000 and 3,000 more iterations (nitt = 5000) while for STRUCTURE, we used a burn-in of 20,000 and 100,000 iterations with default parameters; 95% credible intervals were computed for both methods. The two methods mentioned above provide admixture proportion in the form of a Qscore (STRUCTURE) and posterior probability (Buerkle's HI) but do not provide an easy solution to differentiate the genealogical F1 and F2 + classes since they both result in admixture proportions around 50%. To characterize and differentiate these different genetic classes, we used the hybrid classification tool NewHybrids (Anderson and Thompson 2002). NewHybrids uses a Bayesian statistical method to compute posterior probabilities, enabling the assignment of individuals to specific hybrid categories with an associated confidence level. We evaluated the data for six genotype frequency classes from two generations of potential interbreeding between the 2 species of interest. A total of 115,501 sweeps were done before genotypes were assigned to their specific classes.

Results

MassARRAY assay efficiency and how assignments compare to GBS

The definitive version of the assay holds 26 nuclear and two chloroplastic SNPs (CAPS02 and CAPS10) for distinguishing between *J. cinerea* and *J. ailantifolia*, as well as 2 SNPs (JUGv001 and JUGv010) that can help identify the presence of *J. nigra* and *J. regia* (Supplementary Table S1) with the main objective of removing samples that contain these species from hybrid analysis. Originally, the assay was supposed to contain six more SNPs but three did not amplify well and three produced clustering results of poor quality.

The genotyping success rate for the 26 nuclear SNPs of the assay was 96.2%, while the two chloroplast CAPS-transformed SNPs had a success rate of 99.5%, for the 1,375 samples retained (Supplementary Table S2); this excludes 122 low quality DNA samples with 42% and higher missing data (Supplementary Table S3).

Given the nearly fixed nature of the genetic markers used in our assay, the Bayesian approach employed by STRUCTURE (Pritchard et al. 2000) to calculate a probabilistic degree of admixture can be a valuable tool for estimating the level of hybridization between the two species. However, the simpler and more straightforward Buerkle's hybrid index (Buerkle 2005) is equally effective for estimating hybridization levels. The correlation between the admixture proportion calculated by the two methods (HI vs. Qscore) shows an R^2 of 0.999, indicating that both methods can confidently be used without compromising accuracy (Supplementary Table S4).

As part of the initial SNP panel validation, 86 samples were selected based on morphological and genotyping assessment. Based on their morphology at time of collection and the profile of 2 SNPs (JUGv_001 and JUGv_010) from the panel, we removed four non-*J. cinerea* and non-*J. ailantifolia* samples. The remaining samples were then classified as either admixed or pure, based on previous GBS and microsatellite marker data assessment (Hoban 2008). Among these, 58 were pure, including 40 *J. cinerea* and 18 *J. ailantifolia* from across the range (Fig. 1; Supplementary Table S5). The remaining 28 samples were identified as hybrids.

Using a hybrid index threshold of 0.95, all but one sample were correctly identified. One sample (juglansval12) identified as an advanced generation *J. cinerea* backcross with GBS was instead classified as pure *J. cinerea* using MassARRAY genotyping in which, all SNPs were *J. cinerea* specific. The hybrid index of the 55 samples genotyped using both GBS (11,952 SNPs) and MA (26 SNPs) shows a very high correlation ($R^2=0.98$), indicating that the 26 SNPs selected allow for confident assignment of hybrid status (Fig. 2).

Field testing the panel on wild and plantation trees in North America

After panel validation, we field tested the panel on 1,269 individuals selected from US and Canadian sources. The FGCA individuals (162) consists of individuals that were

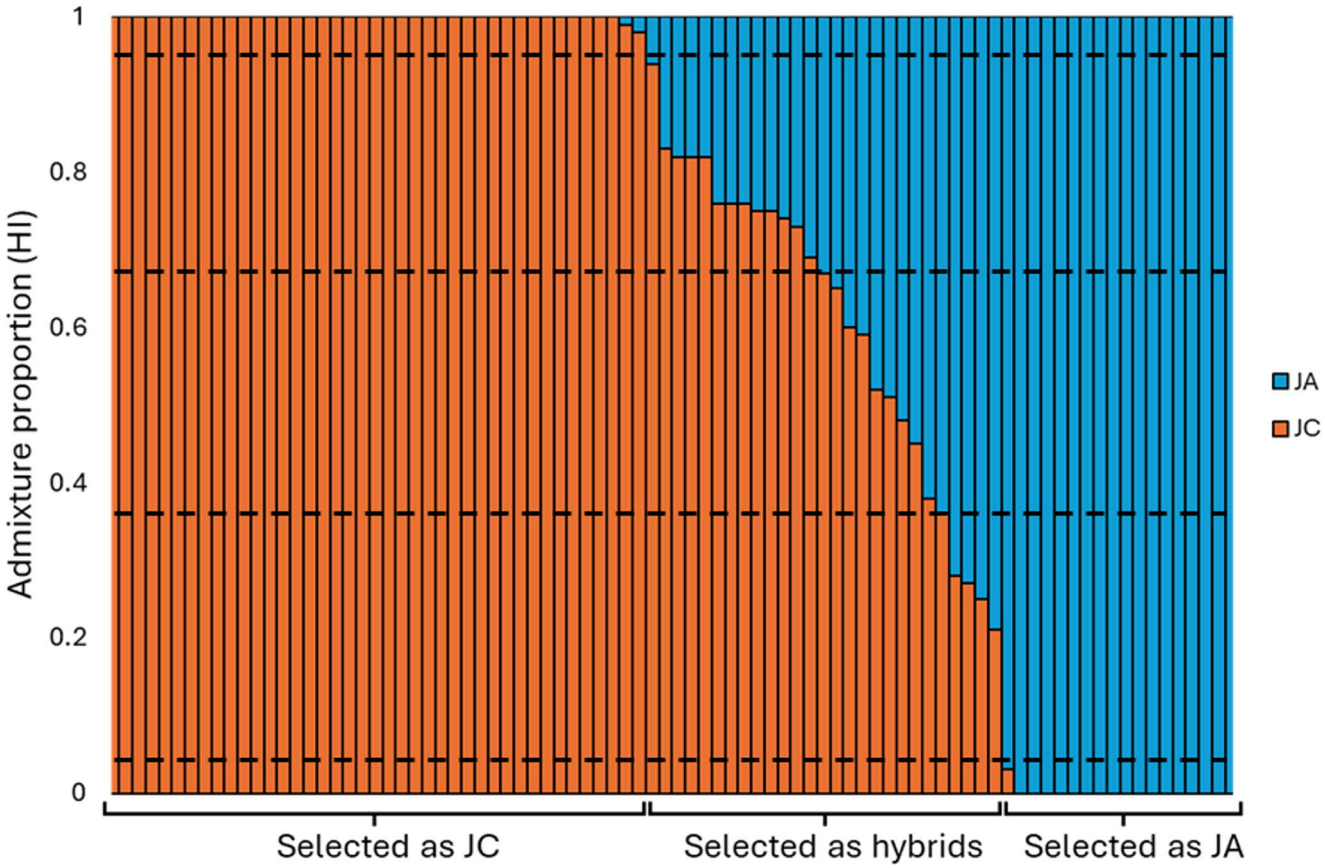


Fig. 1 Hybrid index graph depicting the admixture proportion of all validation samples pre-selected based on previous GBS or microsatellite genotyping assignment. Orange and blue indicate the admixture proportion for *J. cinerea* (JC) and *J. ailantifolia* (JA), respectively. The dashed lines refer to the admixture proportion of 0.95, 0.65, 0.35 and 0.05

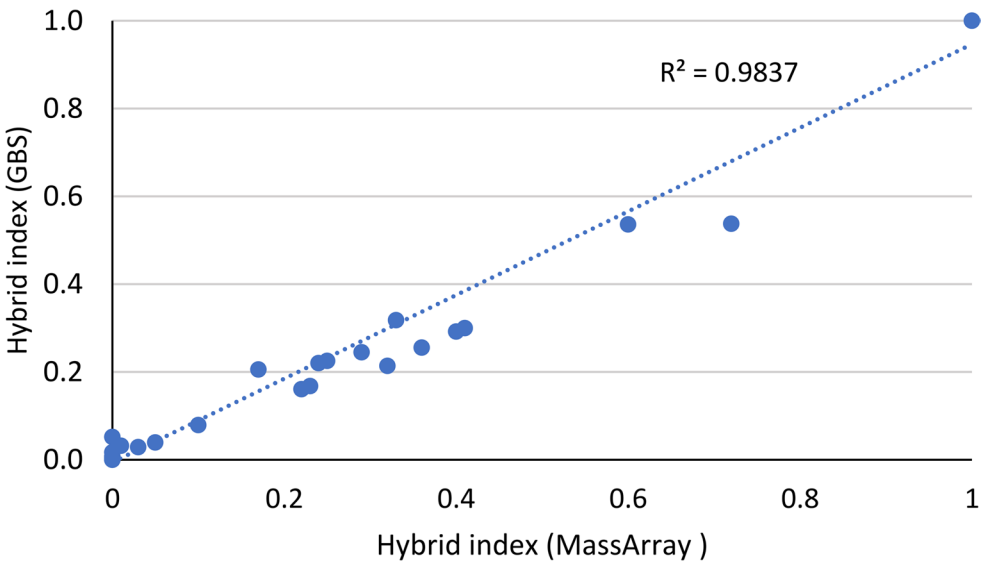


Fig. 2 Scatterplot showing the relationship between the hybrid index (admixture proportion) as determined by GBS (11,952SNPs) and MassARRAY (26 nuclear SNPs) genotyping from 55 individuals

selected from the wild for their putative resistance to canker. The other samples (Canada; 98, HTIRC; 445, Bell (R9); 365, MTE; 49 and US; 150) were sourced from the wild, from seed orchards or germplasm repositories, or from a commercial nursery, with unknown source. The results show that although many trees appear as non-admixed *J.*

cinerea (81.7%), there are still many hybrid trees (18.2%) within these plantations (Fig. 3A). We can see some important differences between the samples collected from Canada and the US. First, all but two (JUNL2024_743 and juglansprod7, Supplementary Table S6) Canadian accessions collected from the wild and in the FGCA archived gardens were assigned as pure *J. cinerea* while in the US, the proportion of hybrids observed were 25%, 23% and 17% for HTIRC, Bell (R9) and other US (MTE + US) datasets, respectively. Using the admixture proportion, we observed that over 60% of the hybrids exhibit introgression predominantly directed towards *J. cinerea* (HI between 0.05 and 0.35), only 12% towards *J. ailantifolia* (HI between 0.65 and 0.95) and 28% as complex hybrids with intermediate proportions (HI between >0.35 and <0.65) (Supplementary Table S6). Using NewHybrids (NH), we were able to distinguish individuals identified as complex hybrids, separating those classified as first-generation hybrids (F1) which are expected to be heterozygous at all the nuclear SNP positions, from those resulting from more intricate hybridization patterns between congeners across multiple generations. Out of 62 trees with HI values >0.35 and <0.65, only 11 were identified as F1 with a NH posterior probability (PP) ≥ 0.95 (Supplementary Table S7). Another 28 were assigned as complex hybrids (PP ≥ 0.95) and 1 as BC_A (PP ≥ 0.95). Another 6 were assigned as F1, 10 as Complex hybrids, 2 as BC JC and 3 as BC JA with posterior probabilities between 0.5 and 0.95. The last one had a probability below 0.5. Out of the 11 F1 hybrids, 9 individuals belonged to the US sample group while the other 2 F1 individuals came from the HTIRC plantation.

Geographic origin of the hybrids

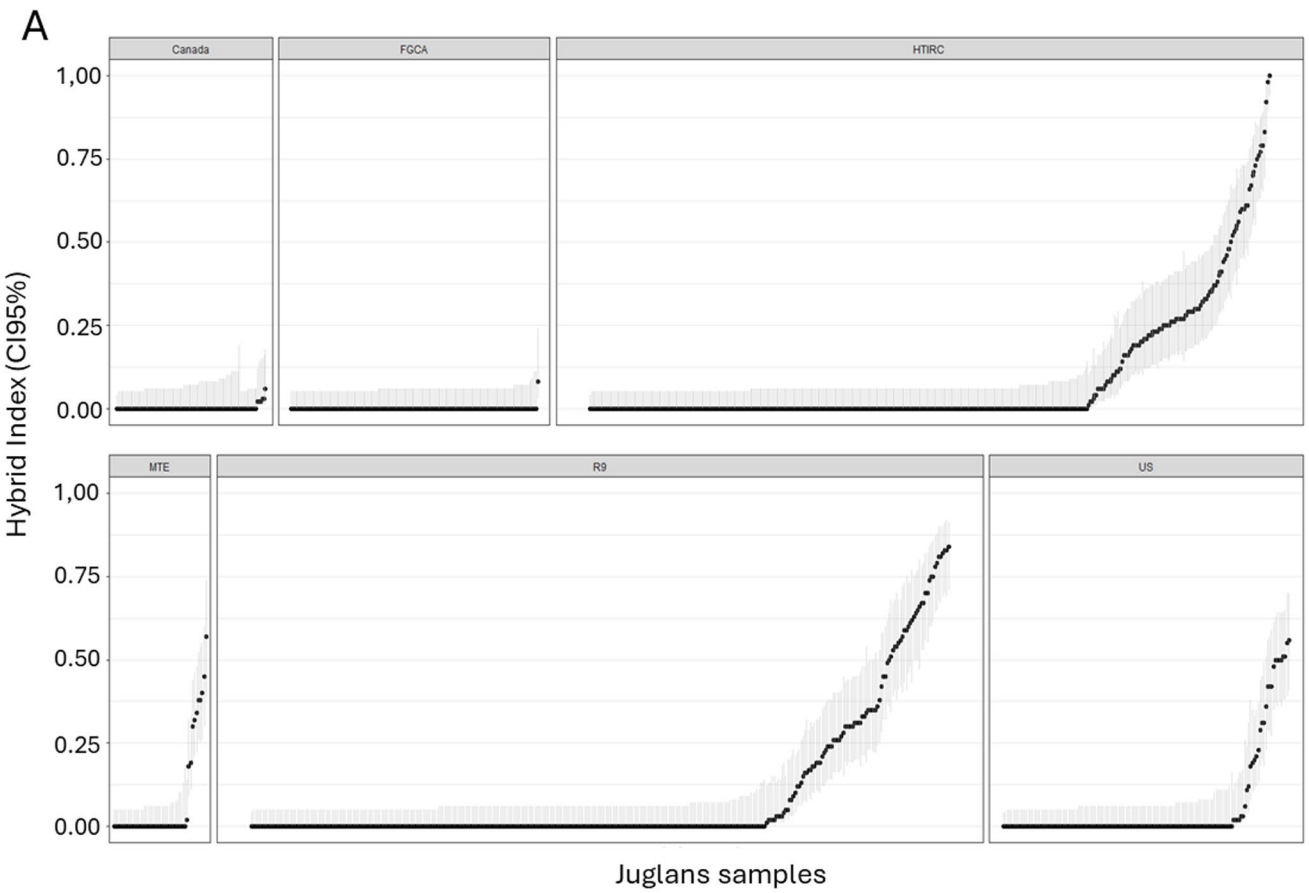
As part of the initial testing of the panel, various non-admixed *J. cinerea* samples from across its natural range were selected to assess whether the SNPs selected were fixed across the landscape. This was important since it has been demonstrated that *J. cinerea* populations are genetically differentiated at the landscape level (Hoban et al. 2010; Schumacher et al. 2022). Although the sampling was not specifically designed to look at the level of hybridization across the landscape, individuals from various regions in both Canada and the US allowed us to determine the origin of the hybrids (Supplementary Table S8). As can be seen from the map (Fig. 3B), hybrids are interspersed across the landscape, but the greatest proportion are found South of the Great Lakes in Northern Indiana. There are also proportionally less hybrids in the southern states below the 40th parallel North and very few hybrids from Canada. The two individuals from Canada identified were very near the 0.95 HI admixture cut-off (both 0.94) and both individuals (juglansprod7, JUNL2024_743) had posterior probabilities ≥ 0.95 classifying them as JC according to NH (see Supplementary Table S7).

Direction of hybridization

As part of panel development, two previously developed CAPS (McCleary et al. 2009) were used and redesigned for use as markers in the SNP-based assay. These markers are useful to identify the direction of the crosses since in *Juglans*, chloroplasts are maternally inherited (Zhang and Sodmergen 2003). As expected, most non-admixed individuals exhibit their respective species-specific chloroplast (Table 1). For the initial validation samples (86; 90 minus 4 *J. nigra* samples), all pure *J. cinerea* and *J. ailantifolia* samples had their respective chlorotypes except for one *J. cinerea* (Table 1). For the validation hybrids, most (24) had *J. ailantifolia* specific chlorotype 2 and three had chlorotype 1. For the field-tested samples, out of 1036 pure *J. cinerea* (HI ≥ 0.95), all but 6 individuals had chlorotype 1 (AC) specific to *J. cinerea* (Table 1). For the other 6 samples, 3 were missing and the other 3 had chlorotype 2 (GT). For non-admixed *J. ailantifolia*, the 2 individuals had their respective chlorotype 2 (GT). Among the 231 hybrids (without Validation group) observed (18.2% of individuals), 11 missed (null alleles) either one or both chloroplast markers. The majority (71.4%) harbored the main *J. ailantifolia* chlorotype 2 while 19% had the *J. cinerea* chlorotype. We also found a rarer third chlorotype (AT) that was observed in only 11 of the field-tested samples. They were all hybrids suggesting it is also specific to *J. ailantifolia*. At the country level, all non-validation samples from Canadian sources had the *J. cinerea* specific chlorotype 1. As expected, all four *J. nigra* samples used for initial validation exhibited the same chlorotype as JC (Supplementary Table S2).

Discussion

Characterizing the level of disease resistance within species is challenging. Having introduced congeners with known disease resistance that can hybridize with the native species can make this characterization even more difficult, especially if the introduced species has been present over a long period of time and when hybrids are difficult to identify morphologically. This is the situation with *J. cinerea* that has co-existed with *J. ailantifolia* since being introduced in the late 1800s. Although listed as a possible threat due to genetic swamping (Committee on the Status of Endangered Wildlife in Canada (COSEWIC) 2017), introgression with the *J. ailantifolia* may be a viable option to preserve *J. cinerea* in the future if insufficient resistance is found in the native species. However, this introduction complicates resistance phenotyping since on-going research is still


A
Canada
FGCA
HTIRC
MTE
R9
US
Hybrid Index (CI95%)
1,00
0.75
0.50
0.25
0.00
Juglans samples

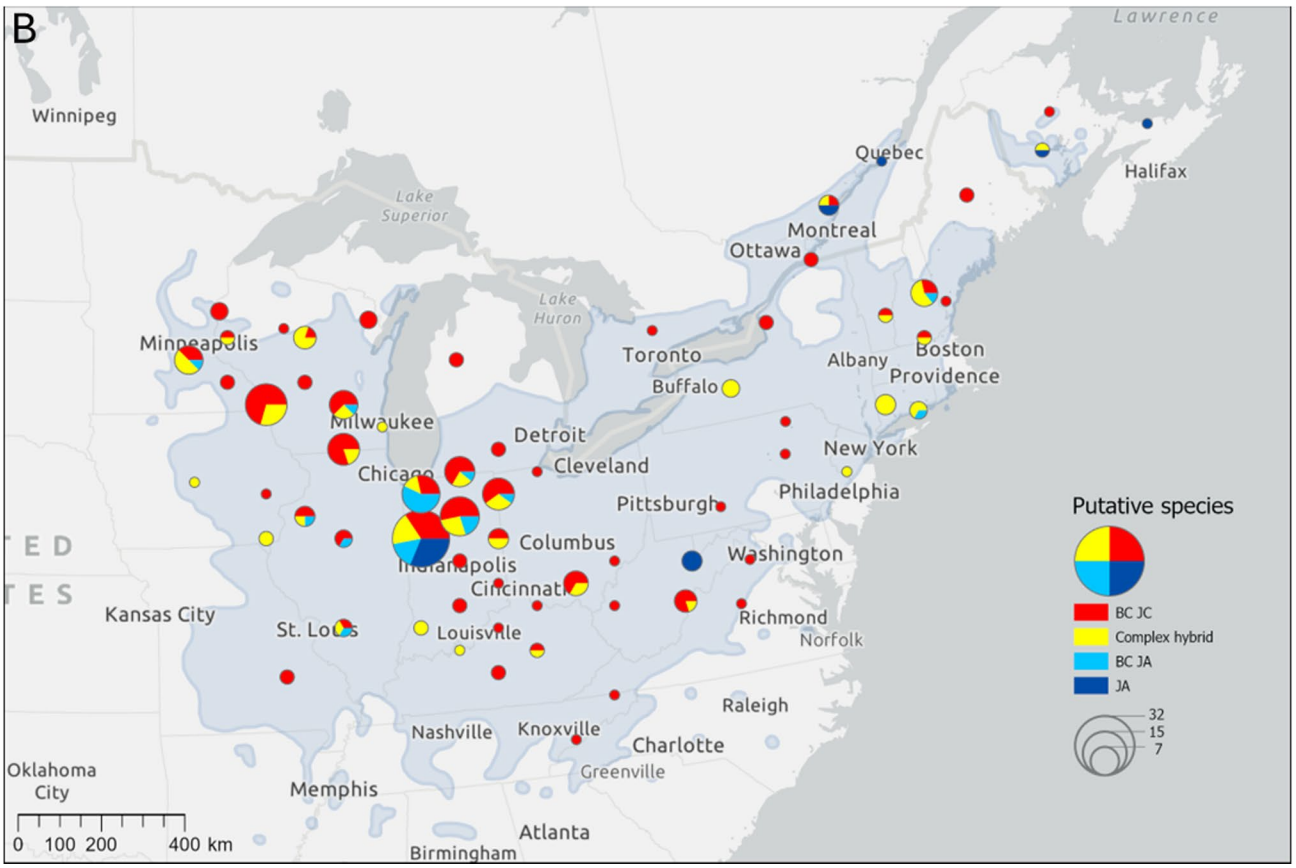

B
Winnipeg
Lake Superior
Lake Huron
Minneapolis
Milwaukee
Chicago
Detroit
Cleveland
Toronto
Buffalo
Ottawa
Montreal
Quebec
Halifax
Albany
Boston
Providence
New York
Philadelphia
Pittsburgh
Columbus
Indianapolis
Cincinnati
Washington
Richmond
Norfolk
Raleigh
Charlotte
Greenville
Nashville
Knoxville
Louisville
St. Louis
Kansas City
Oklahoma City
Memphis
Atlanta
Birmingham
0 100 200 400 km
Putative species
BC JC
Complex hybrid
BC JA
JA
32
15
7

Fig. 3 (**A**) Bar graphs showing the hybrid index values and 95% credible intervals for the 1269 wild and plantation individuals assessed from the US and Canada. (**B**) Map showing the origin of the hybrids (260) in the study as determined by their Hybrid Index. Samples with identical coordinates have been organized in pie chart with size determining the number of samples. The HI values are based on the proportion of *Juglans cinerea* admixture with *J. ailantifolia*. BC JC > 0.05 to ≤ 0.35; Complex hybrid > 0.35 to ≤ 0.65; BC JA > 0.65 to < 0.95; JA 0.95 to 1.0. Pure JC individuals (which are in the majority) were not included in the map to highlight the origin of the hybrids within our study

trying to characterize and find a source of disease resistance within non-admixed *J. cinerea*.

The genomic panel developed herein is functional and can provide end-users with the capabilities to characterize the hybrid status of their trees and support the next steps for improving the efforts of identifying disease resistance within the species or introducing it via hybrid introgression. We aimed to develop a SNP-based assay that could separate the pure *J. cinerea* from the ones admixed with *J. ailantifolia*. Moreover, we wanted a tool that was cost effective, precise, and easily scalable, that could provide various end users (i.e., arborists, researchers, Species at Risk regulators, etc.) with a quick genotyping option regarding their trees of interest (e.g. see Godbout et al. 2017). To do this, we used GBS sequence data from the two congeners to identify species specific SNPs and tested the success rate of the assignment of the MA panel to identify hybrids and compared those assignments to the same samples that were tested using GBS. Finally, we use the panel to field test and characterize the hybrid status of both natural and plantation individuals from the US and Canada.

Table 1 Table representing the putative species and chlorotype for all individuals (Validation and testing) within the study. The final TOTAL represents all samples in the field testing (not including the Validation sample source)

Sample source	Putative species	*n* samples	Null chlorotype *n* samples	Chlorotype 1 J. cinerea (AC) *n* samples	Chlorotype 2 J. ailantifolia (GT) *n* samples	Chlorotype 3 JC/JA (AT) *n* samples
Validation	*J. cinerea*	41	0	40	1	0
Validation	*J. ailantifolia*	18	0	0	18	0
Validation	Hybrid: BC JA	4	0	1	3	0
Validation	Hybrid: BC JC	14	1	0	13	0
Validation	Complex hybrids	9	0	1	8	0
Total		86	1	42	43	0
Canada (FGCA + Canada)	*J. cinerea*	258	0	258	0	0
Canada (FGCA + Canada)	*J. ailantifolia*	0	0	0	0	0
Canada (FGCA + Canada)	Hybrid: BC JA	0	0	0	0	0
Canada (FGCA + Canada)	Hybrid: BC JC	2	0	2	0	0
Canada (FGCA + Canada)	Complex hybrids	0	0	0	0	0
Total		260	0	260	0	0
HTIRC	*J. cinerea*	332	2	328	2	0
HTIRC	*J. ailantifolia*	2	0	0	2	0
HTIRC	Hybrid: BC JA	12	1	0	9	2
HTIRC	Hybrid: BC JC	73	4	14	54	1
HTIRC	Complex hybrids	26	2	2	22	0
Total		445	9	344	89	3
Bell (R9)	*J. cinerea*	281	1	279	1	0
Bell (R9)	*J. ailantifolia*	0	0	0	0	0
Bell (R9)	Hybrid: BC JA	17	0	1	14	2
Bell (R9)	Hybrid: BC JC	41	0	9	29	3
Bell (R9)	Complex hybrids	26	1	7	16	2
Total		365	2	296	60	7
US other (MTE + US)	*J. cinerea*	165	0	165	0	0
US other (MTE + US)	*J. ailantifolia*	0	0	0	0	0
US other (MTE + US)	Hybrid: BC JA	0	0	0	0	0
US other (MTE + US)	Hybrid: BC JC	16	0	7	8	1
US other (MTE + US)	Complex hybrids	18	3	2	13	0
Total		199	3	174	21	1
Total	*J. cinerea*	1036	3	1030	3	0
Total	*J. ailantifolia*	2	0	0	2	0
Total	Hybrid: BC JA	29	1	1	23	4
Total	Hybrid: BC JC	132	4	32	91	5
Total	Complex hybrids	70	6	11	51	2
TOTAL (not including Validation)	All putative species	1269	14	1074	170	11

Interspecific SNP panel accurately identifies non-admixed J. cinerea

The generated panel consists of 30 SNPs integrating both nuclear and chloroplastic SNPs. The addition of two chloroplastic SNPs allow for the identification of chloroplast inheritance. To maximize the utility of each SNP across the species' natural range, we selected *J. cinerea* samples from diverse regions across the landscape for SNP validation. Two more SNPs were added to confirm the presence of *J. nigra*. Although *J. nigra* is morphologically easily differentiated from the other two species and does not hybridize with *J. cinerea*, those two SNPs provide some increased confidence in the case of sampling from grafts, where *J. nigra* rootstocks are used.

We then compared the success rate of our selected SNPs compared to previously used methods of analysis. This included GBS, which provides a powerful approach to generate a high number of SNP markers for hybrid analysis. Although useful for hybrid determination, the genotyping method is mainly used for deeper genetic analysis like genome-wide association (GWAS) and genome-environment association (GEA) studies. The method requires better DNA quality and requires more complex bioinformatics pipelines to acquire and interpret marker data, which makes it less attractive for end users looking for quick taxonomic identification. In our case, the method was used to evaluate the predictive power of our 26 nuclear SNPs (MA) against 11,952 SNPs (GBS). The very high correlation between the two approaches ($R^2 = 0.98$) clearly demonstrates the efficiency of the panel to separate non-admixed individuals from hybrids, which is key for many end user applications, particularly when it only requires the knowledge of whether individuals are hybrid or not for regulatory purposes (Jacobs et al. 2023). However, the lower number of SNPs limits our ability to effectively determine advanced generations of hybrids, which would require many more SNPs (Padilha da Silva and Caetano 2020). The application of MA genotyping on previously confirmed samples demonstrates its ability to differentiate the non-admixed congeners with the panel able to confirm 100% of the non-admixed *J. cinerea* and *J. ailantifolia* pre-selected as such. All the samples pre-selected as hybrids were also correctly identified.

Plantation genotyping shows substantial numbers of hybrids

After validation on known samples, we field tested the genotyping approach to various individual trees located in natural stands and in plantations from the US and Canada were selected as putatively resistant stock (FGCA) and others to some extent (HTIRC, Bell_R9). The others were collected in natural stands (Canada), from germplasm repositories (US) or from a commercial nursery in the US (MTE). Although not surprising, our results show a substantial number of hybrids in the plantations which highlights the difficulty of selecting non-admixed *J. cinerea* from the landscape using morphological traits alone. Although useful as a guide since a good proportion are accurately identified (>80%), morphological identification can become problematic especially when this selection is used to remove hybrid trees from the landscape for regulatory purposes, for example, to minimize the threat of genetic swamping from *J. ailantifolia*. Though a reassessment of the risk of genetic swamping versus the risk of extinction is another topic that could be examined, since this is currently listed as a threat but could be a potential solution for species conservation by limiting functional extirpation in the event of insufficient resistance within the native species. Misidentification can also be troublesome when the goal is the phenotyping for resistance traits in *J. cinerea* since it could contribute to a false conclusion that there is resistance within the species when in reality, the genetic contribution for the resistance is based on *J. ailantifolia* introgression.

Results herein also show that a greater proportion of hybrids have the *J. ailantifolia* chloroplast, indicative of *J. ailantifolia* as the mother in at least one previous generation since the chloroplast is maternally inherited in *Juglans* (Zhang and Sodmergen 2003). In our study, 75% of the identified hybrids had the *J. ailantifolia* specific chloroplast. In Hoban et al. (2009), the proportion was even higher where over 90% of the hybrids within their dataset had the *J. ailantifolia* chloroplast demonstrating a high level of directional hybridization towards the exotic congener. Another interesting observation in our dataset is regarding the origin of these hybrids. It is important to point out that the focus of our study was not to identify the landscape level distribution of the hybrids nor to delimit any hybrid zones since only one species is native, but to focus on characterizing the hybrid status of trees found mostly within managed plantations where many are being phenotyped for resistance to the pathogen. What we found was that all but a few hybrids originate from the US. The reason for this discrepancy is unknown but likely involves many factors such as differential flowering phenology in *Juglans*, both in type and in timing of overlap (Gleeson 1982; Kimura et al. 2003; Bai et al. 2007); excess native congener pollen density due to higher numbers of adjacent *J. cinerea* trees and asymmetrical crossing barriers in angiosperms (Tiffin et al. 2001). For example, bud flush phenology is different between non-admixed *J. cinerea*, *J. ailantifolia* and their hybrids with *J. ailantifolia* (the introduced species) being the earliest to flush followed by the hybrids and *J. cinerea* (Aziz Ebrahimi, unpublished data) as observed in poplar species and

their hybrids (Roe et al. 2014a, b). Phenological differences in flowering times across the latitudinal gradients may have influenced regional patterns of hybridization, potentially limiting hybrid formation in some areas while favoring it in others (LeBoldus et al. 2013). However, without extensive studies looking at flowering phenology for non- admixed *J. cinerea*, *J. ailantifolia* and various hybrids growing at different latitudes, we cannot conclude that this is the main driver for these observations. The earlier introduction (reported in 1967) of the pathogen in the US (Renlund 1971), and possibly much earlier (Broders et al. 2015) in comparison to its introduction in Canada in the 1990s is another factor that could also have contributed to higher hybrid survival and presence in different regions. *J. cinerea* decline of 77% had already been reported between 1966 and 1986 (Anderson and LaMadeleine 1978; Ostry et al. 1994). More recent work by Morin et al. (2018) had shown a steep decline in both *J. cinerea* number (58%) and volume (44%) across the US since the 1980s. It is possible that increased pathogen susceptibility of *J.* cinerea could have removed the more susceptible congener, leaving the hybrids on the landscape to be selected and planted in living collections. However, tree hybrid status at the landscape level was not characterized within these surveys and hence it cannot be empirically confirmed whether this was a contributing factor.

Future application in conservation and tree improvement

Protecting species against existential threats like invasive pests and pathogens often requires multiprong conservation strategies to protect both the habitat where the species is found while integrating ex situ strategies such as seed banking to minimize genetic diversity loss. Genomic tools are increasingly being developed to support these efforts for real world applications in forest health (Sniezko and Koch 2017; LeBoldus et al. 2024) in early detection of pests and pathogens, to characterize biodiversity but are also very useful to support tree improvement activities which are often one of the only options available to protect the trees from decline. Out of all the tools, genotyping is one of the most used since it can support many of these efforts to improve trees by predicting the traits of progeny through marker assisted selection methods such as genomic selection (Meuwissen et al. 2001) which can save years off the breeding cycle for tree improvement. It is also very important in species identification and in this case for interspecies hybrid determination.

The hybrid SNP panel presented in this study was developed for specific applications for the conservation of *J. cinerea*. It was initially developed to help in the characterization of trees that were already present in living collections across North America. Some of these living collections are currently being assessed for canker susceptibility and mischaracterizing these trees as non-admixed would bias the conclusion that disease resistance within non-admixed trees can be achieved. The tool will be able to support the characterization of new selections from wild populations and minimize resource loss by grafting or selecting mischaracterized trees.

Ex situ conservation is another avenue that will benefit from this genomic tool by its ability to characterize the hybrid status of seed-bearing trees for conservation. Currently, the only ex situ strategy for the long-term conservation of remaining *J. cinerea* genetic diversity is through cryo-conservation of embryogenic axes (Williams et al. 2019, 2023). Both the trees used as mother trees for these collections as well as the trees regenerated from this backup will be able to be assessed.

Another benefit is that the tool will help assess the extent of introgression within natural populations and provide insights into the level of introgression in the next generation where there is regeneration. Although landscape level introgression analysis can provide useful data into the genetic impact on endangered species because of genetic swamping and loss of genetic diversity. It can also complicate the legal status of species, especially if the hybrids do not have the same protection status as their non-admixed congeners.

To be successful, multiple sources of resistance will likely be required before a viable reintroduction program can be established for pure *J. cinerea.* This effort has become increasingly important to the *J. cinerea* research community to determine whether tree improvement in *J. cinerea* can be achieved and be long lasting without congener introgression. Given the endangered status of the species and its rapid decline, we suggest these research efforts be undertaken sooner rather than later to prevent the further loss of the adaptive potential and its ecological function for benefits at a larger landscape scale. As previously mentioned, some regional populations have already been lost because of the pathogen (Committee on the Status of Endangered Wildlife in Canada (COSEWIC) 2017; Morin et al. 2018) and the pace has not slowed. This is particularly evident in New Brunswick, which was the last region to report on the presence of the pathogen and where, in the last few years, it has become widespread, leading to increasing mortality (22% loss in 7 years) and declining genetic diversity (van der Meer et al. 2025). From a resistance breeding standpoint, these results clearly show the importance of characterizing trees genetically before phenotyping them for inclusion in a tree improvement program, especially one focused on selecting disease resistance within non-admixed *J. cinerea*. Deploying this genomic tool will be key to supporting the

various research activities currently being undertaken and help drive the discussion on the requirement and the value of using hybrids to support the conservation and recovery efforts of the species, where time is of the essence.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11295-026-01735-y.

Acknowledgements We thank Martine Blais, Marie Deslauriers, Marie-Claude Gros-Louis, Ricardo Morin and Danny Rioux (NRCan) for providing samples from the Province of Quebec, Glenna Halicki Hayden (Ontario Forest Research Institute), Heather Zurbrigg (Forest Gene Conservation Association) and Melissa Spearing for Ontario samples and Aaron Dowding (Nature Conservancy of Canada) for providing NB samples. We thank Menominee Tribal Enterprises, Tommy Stueck (MTE), David Mausel and Nick LaBonte (USDA Forest Service) for providing butternut samples from the US. We thank Chantelle Kostanowicz and Taylor Swanburg (NRCan) for DNA extractions of Canadian samples. Genome Quebec for their contribution in the design and development of the MassARRAY SNP panel as well as for the genotyping of all samples. Finally, we would like to thank Ian DeMerchant for producing the map used in figure 3B.

Author contributions M. Lamothe, M. Williams, A. Ebrahimi, M. Zamani, J. Warren, A. Conrad and B.M. van der Meer conducted material collection and preparation for genotyping the samples. M. Williams, M. Lamothe and A. Ebrahimi conducted analyses on the samples. M. Williams and M. Lamothe wrote the first draft. Conceptualization and supervision were carried out by M. Williams, N. Isabel; funding acquisition by N. Isabel, E. Ebrahimi, D. Jacobs and M. Williams; resources were provided by A. Conrad, E. Ebrahimi, J. Warren, C. Pike, D. Jacobs, N. Isabel and M. Williams. All authors read and approved of the final manuscript.

Funding Open access funding provided by Natural Resources Canada library. This work was supported by The Genomics Research Development Iniative (NRCan) awarded to NI and MW, by the HTIRC and USDA Forest Service Gene Conservation, Resistance and Restoration program (Award # 22-CS-11242314-077 awarded to AE, DFJ and AOC as well as a USDA National Institute of Food and Agriculture (Award Number #: 2023-67012-39608) awarded to AE.

Data availability Data is provided within the manuscript or supplementary information files.

Declarations

Competing interests The authors declare no competing interests.

References

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229. https://doi.org/10.1093/genetics/160.3.1217

Anderson RL, LaMadeleine LA (1978) The distribution of butternut decline in the eastern United States. USDA Forest Service Survey Rep. S-3-78. Northeastern Area State and Private Forestry, Broomall, PA

Aradhya MK, Potter D, Gao F, Simon CJ (2007) Molecular phylogeny of Juglans (*Juglandaceae*): a biogeographic perspective. Tree Genet Genomes 3:363–378. https://doi.org/10.1007/s11295-006-0078-5

Bailey RI (2024) Bayesian hybrid index and genomic cline estimation with the R package gghybrid. Mol Ecol Resour 24:e13910. https://doi.org/10.1111/1755-0998.13910

Bai WN, Zeng YF, Zhang DY (2007) Mating patterns and pollen dispersal in a heterodichogamous tree, *Juglans mandshurica* (Juglandaceae). New Phytol 176:699–707. https://doi.org/10.1111/j.1469-8137.2007.02202.x

Bixby W (1919) The butternut and the Japan walnut. Am Nut J 10(6):76–83

Boraks A, Broders KD (2014) Butternut (Juglans cinerea) health, hybridization, and recruitment in the northeastern United States. Can J Res 44:1244–1252. https://doi.org/10.1139/cjfr-2014-0166

Brennan AN, McKenna JR, Hoban SM, Jacobs DF (2020) Hybrid breeding for restoration of threatened forest trees: evidence for incorporating disease resistance in *Juglans cinerea*. Front Plant Sci 11:580693. https://doi.org/10.3389/fpls.2020.580693

Broders K, Boraks A, Barbison L, Brown J, Boland GJ (2015) Recent insights into the pandemic disease butternut canker caused by the invasive pathogen *Ophiognomonia clavigignenti-juglandacearum*. Pathol 45:1–8. https://doi.org/10.1111/efp.12161

Browning SR, Browning BL, Zhou Y, Tucci S, Akey JM (2018) Analysis of human sequence data reveals two pulses of archaic Denisovan admixture. Cell 173:53-61.e9. https://doi.org/10.1016/j.cell.2018.02.031

Buerkle CA (2005) Maximum-likelihood estimation of a hybrid index based on molecular markers. Mol Ecol Notes 5:684–687. https://doi.org/10.1111/j.1471-8286.2005.01011.x

Committee on the Status of Endangered Wildlife in Canada (COSEWIC) (2017) COSEWIC assessment and status report on the butternut *Juglans cinerea* in Canada. COSEWIC, Ottawa, ON, Canada

Conrad AO, Ebrahimi A, Warren J, Fetter K, Pike C, Kell C, McKenna J, Williams M, Faradonbeh MZ, Jacobs DF (2026) Susceptibility to butternut canker disease in field trials is influenced by disease level and hybridization. Can J of For Res 56:1–13. https://doi.org/10.1139/cjfr-2025-0234

Dode LA (1909) Contribution to the study of the genus *Juglans*. Bull de la Société Dendrologique de France 11:22–90

Dong W, Xu C, Li W, Xie X, Lu Y, Liu Y, Jin X, Suo Z (2017) Phylogenetic resolution in *Juglans* based on complete chloroplast genomes and nuclear DNA sequences. Front Plant Sci 8:1148. https://doi.org/10.3389/fpls.2017.01148

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Ebrahimi A (2022) Morpho-physiological and genomics analyses reveal adaptations of hardwood trees to abiotic stressors. Purdue University Graduate School. Thesis. https://doi.org/10.25394/PGS.21669269.v1

Ebrahimi A, Mathur S, Lawson SS et al (2019) Microsatellite borders and micro-sequence conservation in *Juglans*. Sci Rep 9:3748. https://doi.org/10.1038/s41598-019-39793-z

Environment Canada (2010) Recovery strategy for the butternut (*Juglans cinerea*) in Canada. Species at Risk Recovery Strategy Series. Environment Canada, Ottawa, ON, Canada

Farlee LD, Woeste K, Ostry ME, McKenna JR, Weeks S (2010) Identification of butternuts and butternut hybrids. Purdue Extension, FNR-420-W Purdue University, West Lafayette, IN

Fjellstrom RG, Parfitt DE (1994) Walnut (*Juglans* spp.) genetic diversity determined by restriction fragment length polymorphisms. Genome 37:690–700. https://doi.org/10.1139/g94-097

Glaubitz JC, Casstevens TM, Lu F, Harriman J, Elshire RJ, Sun Q, Buckler ES (2014) TASSEL-GBS: a high-capacity genotyping by sequencing analysis pipeline. PLoS ONE 9:e90346. https://doi.org/10.1371/journal.pone.0090346

Gleeson SK (1982) Heterodichogamy in walnuts: inheritance and stable ratios. Evolution 36(5):892–902. https://doi.org/10.1111/j.1558-5646.1982.tb05461.x

Godbout J, Tremblay L, Levasseur C, Lavigne P, Rainville A, Mackay J, Bousquet J, Isabel N (2017) Development of a traceability system based on a SNP array for large-scale production of high-value white spruce (*Picea glauca*). Front Plant Sci. https://doi.org/10.3389/fpls.2017.01264

Guzman-Torres CR, Trybulec E, LeVasseur H, Akella H, Amee M, Strickland E, Pauloski N, Williams M, Romero-Severson J, Hoban S, Woeste K, Pike CC, Fetter KC, Webster CN, Neitzey ML, O'Neill RJ, Wegrzyn JL (2023) Conserving a threatened North American walnut: a chromosome-scale reference genome for butternut (*Juglans cinerea*). G3: Genes, Genomes, Genetics 14:jkad189. https://doi.org/10.1093/g3journal/jkad189

Hoban S, Anderson R, McCleary T, Schlarbaum S, Romero-Severson J (2008) Thirteen nuclear microsatellite loci for butternut (*Juglans cinerea* L.). Mol Ecol Resour 8:643–646. https://doi.org/10.1111/j.1471-8286.2007.02030.x

Hoban SM, Borkowski DS, Brosi SL, McCleary TS, Thompson LM, Mclachlan JS, Pereira MA, Schlarbaum SE, Romero-Severson J (2010) Range-wide distribution of genetic diversity in the north American tree *Juglans cinerea*: a product of range shifts, not ecological marginality or recent population decline. Mol Ecol 19:4876–4891

Hoban SM, McCleary TS, Schlarbaum SE, Romero-Severson J (2009) Geographically extensive hybridization between the forest trees American butternut and Japanese walnut. Biol Lett 5(3):324–327. https://doi.org/10.1098/rsbl.2009.0031

Isabel N, Lamothe M, Thompson SL (2013) A second-generation diagnostic single nucleotide polymorphism (SNP)-based assay, optimized to distinguish among eight poplar (*Populus* L.) species and their early hybrids. Tree Genet & Genomes 9(2):1101–1110. https://doi.org/10.1007/s11295-012-0569-5

Jacobs DF, Dumroese RK, Brennan AN, Campbell FT, Conrad AO, Delborne JA, Fitzsimmons S, Flores D, Giardina CP, Greenwood L, Martín JA, Merkle SA, Nelson CD, Newhouse AE, Powell WA, Romero-Severson J, Showalter DN, Sniezko RA, Strauss SH, Westbrook J, Woodcock P (2023) Reintroduction of at-risk forest tree species using biotechnology depends on regulatory policy, informed by science and with public support. New For 54:587–604. https://doi.org/10.1007/s11056-023-09980-y

Kimura M, Seiwa K, Suyama Y, Ueno N (2003) Flowering system of heterodichogamous Juglans ailanthifolia. Plant Species Biol 18:75–84. https://doi.org/10.1111/j.1442-1984.2003.00088.x

LaBonte NR, Jacobs DF, Woeste KE (2015) Estimating heritability of disease resistance and factors that contribute to long-term survival in butternut (*Juglans cinerea* L). For Ecol Manag 336:110–117. https://doi.org/10.1016/j.foreco.2014.10.014

Lawrence M (2017) Increasing the sustainability of a threatened tree species, butternut (Juglans cinerea), for use in traditional dyes for Cherokee basketry. https://mds.marshall.edu/asa_conference/2017/accepted_proposals/82/ (accessed 30 October 2023)

LeBoldus JM, Isabel N, Floate KD, Blenis P, Thomas BR (2013) Testing the "hybrid susceptibility" and "phenological sink" hypotheses using the P. balsamifera - P. deltoides hybrid zone and septoria leaf spot [Septoria musiva]. PLoS One 27(12):e84437. https://doi.org/10.1371/journal.pone.0084437.

LeBoldus JM, Lynch SC, Newhouse AE, Søndreli KL, Newcombe G, Bennett PI, Muchero W, Chen JG, Busby PE, Gordon M, Liang H (2024) Biotechnology and genomic approaches to mitigating disease impacts on forest health. Annu Rev Phytopathol 62:309–335. https://doi.org/10.1146/annurev-phyto-021622-114434

Li H, Durbin R (2009) Fast and accurate short read alignment with burrows–wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lu AM, Stone DE, Grauke LJ (1999) *Juglandaceae*. In: Wu ZY, Raven PH (eds) Flora of China: *Cycadaceae* through *Fagaceae*, vol 4. Science Press/Missouri Botanical Gardens Press, Beijing/St Louis, pp 277–285

Manning WE (1978) The classification within the *Juglandaceae*. Ann Mo Bot Gard 65(4):1058–1087

Marrano A, Britton M, Zaini PA, Zimin AV, Workman RE, Puiu D, Bianco L, Pierro EAD, Allen BJ, Chakraborty S, Troggio M, Leslie CA, Timp W, Dandekar A, Salzberg SL, Neale DB (2020) High-quality chromosome-scale assembly of the walnut *(Juglans regia* L.) reference genome. Gigascience 9:giaa050. https://doi.org/10.1093/gigascience/giaa050

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17(1):10–12. https://doi.org/10.14806/ej.17.1.200

Martínez-García PJ, Crepeau MW, Puiu D et al (2016) The walnut (*Juglans regia*) genome sequence reveals diversity in genes coding for the biosynthesis of non-structural polyphenols. Plant J 87(5):507–532. https://doi.org/10.1111/tpj.13207

McCleary TS, Robichaud RL, Nuanes S, Anagnostakis SL, Schlarbaum SE, Romero-Severson J (2009) Four cleaved amplified polymorphic sequence (CAPS) markers for the detection of the *Juglans ailantifolia* chloroplast in putatively native *J. cinerea* populations. Mol Ecol Resour 9:525–527. https://doi.org/10.1111/j.1755-0998.2008.02465.x

Meirmans PG, Lamothe M, Gros-Louis MC, Khasa D, Périnet P, Bousquet J, Isabel N (2010) Complex patterns of hybridization between exotic and native North American poplar species. Am J Bot 97:1688–1697. https://doi.org/10.3732/ajb.0900271

Meuwissen THE, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157(4):1819–1829. https://doi.org/10.1093/genetics/157.4.1819

Michler CH, Pijut PM, Jacobs DF, Meilan R, Woeste KE, Ostry ME (2006) Improving disease resistance of butternut (*Juglans cinerea*), a threatened fine hardwood: a case for single-tree selection through genetic improvement. Tree Physiol 26:121–128. https://doi.org/10.1093/treephys/26.1.121

Morin RS, Pugh SA, Steinman JR, Crocker SJ (2018) Regional patterns of declining butternut (*Juglans cinerea* L.) suggest site characteristics for restoration. For Ecol Manage 424:11–20. https://doi.org/10.1016/j.foreco.2018.04.046

Mu XY, Tong L, Sun M, Zhu YX, Wen J, Lin QW, Liu B (2020) Phylogeny and divergence time estimation of the walnut family (*Juglandaceae*) based on nuclear RAD-Seq and chloroplast genome data. Mol Phylogenet Evol 147:106802. https://doi.org/10.1016/j.ympev.2020.106802

NatureServe (2025) Explorer: Juglans cinerea. Available at: https://explorer.natureserve.org/Taxon/ELEMENT_GLOBAL.2.154365/Juglans_cinerea. Accessed 25th February 2025

Ohwi J (1965) Flora of Japan. Smithsonian Institution, Washington, DC

Ostry ME, Mielke ME, Skilling DD (1994) Butternut—strategies for managing a threatened tree. USDA Forest Service Gen Tech Rep. NC-165. North Central Forest Experiment Station, St. Paul, MN

Ostry ME, Moore M (2008) Response of butternut selections to inoculation with *Sirococcus clavigignenti-juglandacearum*. Plant Dis 92:1336–1338. https://doi.org/10.1094/PDIS-92-9-1336

Pike CC, Williams M, Brennan A, Woeste K, Jacobs J, Hoban S, Moore M, Romero-Severson J (2020) Save our species: a blueprint for restoring butternut (*Juglans cinerea*) across Eastern North America. J For 119:fvaa053. https://doi.org/10.1093/jofore/fvaa053

Poland JA, Brown PJ, Sorrells ME, Jannink JL (2012) Development of high-density genetic maps for barley and wheat using genotyping-by-sequencing. Plant Genome 5:103–113. https://doi.org/10.3835/plantgenome2012.05.0005

Powell WA, Newhouse AE, Coffey V (2019) Developing blight-tolerant American chestnut trees. Cold Spring Harb Perspect Biol 11:a034587. https://doi.org/10.1101/cshperspect.a034587

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. https://doi.org/10.1093/genetics/155.2.945

R Core Team (2023) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Renlund D (1971) Forest pest conditions in Wisconsin, annual report. Wisconsin Department of Natural Resources, Madison, WI

Roe AD, MacQuarrie CJ, Gros-Louis MC, Simpson JD, Lamarche J, Beardmore T et al (2014a) Fitness dynamics within a poplar hybrid zone: I. Prezygotic and postzygotic barriers impacting a native poplar hybrid stand. Ecol Evol 4:1629–1647. https://doi.org/10.1002/ece3.1029

Roe AD, MacQuarrie CJ, Gros-Louis MC, Simpson JD, Lamarche J, Beardmore T et al (2014b) Fitness dynamics within a poplar hybrid zone: II. impact of exotic sex on native poplars in an urban jungle. Ecol Evol 4:1876–1889. https://doi.org/10.1002/ece3.1028

Sandercock AM, Holliday JA, Westbrook JW (2024) A genome-guided strategy for climate resilience in American chestnut restoration populations. Proc Natl Acad Sci U S A 121:e2403505121. https://doi.org/10.1073/pnas.2403505121

Schumacher E, Brown A, Williams M, Romero-Severson J, Beardmore T, Hoban S (2022) Range shifts in butternut, a rare, endangered tree, in response to past climate and modern conditions. J Biogeogr 49:866–878. https://doi.org/10.1111/jbi.14350

Scrucca L, Fop M, Murphy TB, Raftery AE (2016) mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R J 8(1):289–317

Silva JPD, Caetano AR (2020) Determining minimum numbers of di-allelic diagnostic markers required to identify introgressions in diploid cross-species hybrid individuals from different types of inter- and backcross populations. Genet Mol Biol 21;43(3). https://doi.org/10.1590/1678-4685-gmb-2019-0324

Sniezko RA, Koch J (2017) Breeding trees resistant to insects and diseases: putting theory into application. Biol Invasions 19:3377–3400. https://doi.org/10.1007/s10530-017-1482-5

Stanford AM, Harden R, Parks CR (2000) Phylogeny and biogeography of *Juglans* (*Juglandaceae*) based on *mat*K and ITS sequence data. Am J Bot 87(6):872–882. https://doi.org/10.2307/2656895

Stritch L, Barstow M (2019) Juglans cinerea. IUCN Red List Threatened Species eT62019689A62019696. https://doi.org/10.2305/IUCN.UK.2019-1.RLTS.T62019689A62019696.en. Accessed on 23 September 2025

Tiffin P, Olson MS, Moyle LC (2001) Asymmetrical crossing barriers in angiosperms. Proc R Soc Lond B Biol Sci 268(1469):861–86. https://doi.org/10.1098/rspb.2000.1578

van der Meer BM, Solomon JP, Kostanowicz CA, Williams M (2025) The urgency of genetic decline in New Brunswick butternut, *Juglans cinerea* L. PPP:1–11. https://doi.org/10.1002/ppp3.70079

Williams M, Brophy M, van der Meer BM (2023) Cryogenic storage increases the longevity of butternut (*Juglans cinerea*, L.) seed embryogenic axes. Cryobiology 110:103–106. https://doi.org/10.1016/j.cryobiol.2023.01.002

Williams M, Forbes K, Williams C, Beardmore T (2019) In vitro germination protocol for the propagation of the endangered butternut tree (*Juglans cinerea* L). NPJ 20:117–122. https://doi.org/10.3368/npj.20.2.117

Zhang Q, Sodmergen YL (2003) Examination of the cytoplasmic DNA in male reproductive cells to determine the potential for cytoplasmic inheritance in 295 angiosperm species. Plant Cell Physiol 44:941–951. https://doi.org/10.1093/pcp/pcg121

Zhao P, Woeste KE (2010) DNA markers identify hybrids between butternut (*Juglans cinerea* L.) and Japanese walnut (*Juglans ailantifolia* Carr). Tree Genet Genomes 7:511–533. https://doi.org/10.1007/s11295-010-0352-4