Assessment of genetic diversity of the invasive Callery pear, *Pyrus calleryana* Decne. using mitochondrial microsatellites and its molecular detection using LAMP

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Alina Pokhrel
August 2024

DEDICATION

Dedicated to my beloved family.

ABSTRACT

Pyrus calleryana Decne. (Callery pear; [PC]) is a popular ornamental tree in the urbanized areas of the United States of America (US), owing to its aesthetical value, with showy white blossoms in early spring and vibrant fall foliage. The tree native to Asia is now becoming one of the most problematic invasive tree species in the eastern US. From its introduction in the early 20th century, PC has been commercially used as rootstocks for propagating fruiting pears, alongside other closely related pear species. Several states are restricting the sale of rootstocks to limit the ongoing spread of PC. As such, there is a need to study the genetic diversity of PC and develop a field-based detection tool that can reliably detect PC rootstocks. For the first study, we identified candidate region in nuclear genome of PC and developed a rapid, sensitive and specific loop-mediated isothermal amplification (LAMP) assay to successfully detect invasive PC DNA and to differentiate it from the DNA of three closely related species, namely *P. communis, P. pyrifolia,* and *P. betulifolia.* The LAMP results were obtained within 30 minutes of incubation at 65 °C. The development of a LAMP assay for detecting invasive PC is a pivotal advancement in the management of invasive species. For the second study, we developed mitochondrial genome and five microsatellite markers localized therein, to analyze genetic diversity of Asian population, Southeastern escapees (SNesc), Tennessee escapees (TNesc), and US-released commercial cultivars (UScult). 72 specimens of Asian PC, 67 samples of the UScult, and 180 samples from SNesc/TNesc population were analyzed. Our data revealed high genetic diversity (H_e = 0.842) and presence of genetic structure in PC. The results indicated the possibility of divergence of TNesc from Asian population, and SNesc and UScult from TNesc. The observed high genetic diversity among PC cultivars illustrates the intricate genetic landscape shaped by multiple factors, including potential, highly likely unintended, mislabeling of the US-released cultivars. This study underscores the need for broader genomic studies to further elucidate the genetic architecture of cultivated PC composition and supports the development of robust method for cultivar truthing and invasion management.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Invasive plant species are those species that grow outside their naturally occurring habitats, proliferate, alter, and threaten the native species (Shezi and Poona, 2010; IUCN,2000). Invasive plant species are considered one of the major drivers of biodiversity loss (Gurevitch and Padilla, 2004; Maxwell et al., 2016) that can alter ecosystem functioning (Topin, 2018) including the changes in soil properties (Li et al., 2006; Pejchar and Mooney, 2009; Vilà et al., 2011). Invasive species have the potential to alter ecosystem structures and species compositions by suppressing or displacing native species. This can occur through direct competition for resources or indirect effects on nutrient cycling. Among the 680 animals listed in the 2003 IUCN Red List Database, causes of extinction were identified for 170 species. In approximately 20% of these cases, invasive species were identified as the sole factor driving animal extinction (Bellard et al., 2016). Invasive species have notable negative impacts on human health and well-being, agriculture, tourism, and ultimately threaten the economy, social and environmental stability, and sustainability of the planet (Pimental et al., 2005; McNeely et al., 2001; Kariyawasam et al., 2019). Invasive species interact with other drivers of global change to exert complex and combined effects (Pyšek and Richardson, 2010). The cost of damage by invading species was US$18.6 billion per year between 1970 and 2017 and expenditures for managing all invasive species averaged US$1.4 billion per year (Diagne et al., 2021). The majority of invasive plant species now established in the United States (US) were introduced for food, fiber, or ornamental purposes (Li et al., 2006; Richardson et al., 2000) and their spread is causing well-recognized issues. Thus, there is an urgent need to develop effective management strategies to limit their invasion.

Callery pear (*Pyrus calleryana* Decne.; *P. calleryana* [PC]), a member of the family Rosaceae, is a diploid plant with a chromosome number of 2n=34 (Phillips, et al., 2016) and an estimated genome size of 588 Mbp/1C (Dickson, et al., 1992). PC is a deciduous tree with a trunk less than 30 cm in diameter and a height ranging from 30 to 50 feet. It is an upright-branched ornamental tree native to China, Taiwan, Korea, Vietnam, and Japan with a broad ecological range (Cuizhi and Spongberg, 2003). *P. calleryana* has a pyramidal to rounded crown, alternately arranged oval, long, glossy dark green (above) leaves with crenate margins. White, five-petaled flowers bloom for 1-to-2 weeks, from mid-February to early April, depending upon location and climate, and produce a strong rancid odor, variously described as similar to semen, rotting flesh, or chlorine (Lapidos, 2013). The inedible fruits are hard, until softened by frost, and are green/brown in color, white to tan dotted, spherical to slightly oblong,

with 1 or 2 seeds (Vincent, 2005). The twigs of naturalized invasive PC often bear sharp thorn-like spurs. PC, which was brought to the US from Asia in the early 1900s to aid in management of fire-blight bacterium [*Erwinia amylovora*] in European pear (*Pyrus communis* L.), has been declared invasive (Vincent, 2005). Various experiments revealed that PC was resistant to pests and pathogens which made it a good rootstock for other related and outgroup species. 'Bradford', first officially released cultivar by USDA in 1963 (Whitehouse et al., 1963), is considered the most famous and widely planted ornamental cultivar of *P. calleryana* among approximately 25 cultivars that have been commercialized for landscape use (Sapkota et al., 2022; Whitehouse et al., 1963). PC has enjoyed great success in urbanized areas due to its aesthetics, with beautiful white blossoms in early spring and bright red leaf colors in the fall (Culley et al., 2008). Various traits such as its ability to reach sexual maturity at a young age, a self-incompatibility system that promotes outcrossing, resistance to pathogens and pests, and broad tolerance to various environmental conditions, contribute to the ability of PC to spread into a variety of environments (Culley and Hardiman, 2007). Though PC has gametophytic self-incompatibility, various cultivars and wild escapees may hybridize with compatible pear specimens yielding viable seeds that are then bird- and animal-dispersed (Culley, 2017; Culley and Hardiman, 2007). Even after cutting them down, the tree sprouts back again from the stump and roots (Bowen, 2023) and again may contribute to the invasion inoculum. It is an aggressive invader that has naturalized fallow fields, roadside edges, grassland habitats, and other early successional environments (Vincent, 2005; Woods et al., 2021). PC has the ability to alter habitats by decreasing soil pH and increasing soil C: N (Woods et al., 2021). It has been predicted that the species has the potential of becoming the most problematic invasive plant in the US (Coyle et al., 2021; Sapkota et al., 2021). PC is widely distributed throughout the eastern US and in several western states (EDDMapS, 2024). PC is considered an invasive plant in 29 states and the sale has been banned in some areas including the State of Ohio, Charlotte, NC, and Pittsburgh, PA (Blue Ridge PRISM, 2021). It is a hazard to people and property as well because of its narrow crotch angles that cause branches, especially in older trees to split when strained by wind, ice, or snow (Culley, 2017). The tree forms dense thickets with thorns, challenging native species by creating competition and rendering it difficult for individuals attempting removal due to its impenetrability. (Culley, 2017; Higgins, 2018). In optimal growing condition, PC trees produce hundreds to thousands of fruits in a single growing season and the dropping of the soft fruits on the ground especially during winter, can cause unsightly litter, as well as pose a risk to the people walking around (Fulcher, 2002). Land managers are facing never-ending issues to manage invasive species, which shifts resources from other

management needs (Renz et al., 2009) and the management implementation can be even more challenging than usual when the control measures are difficult to implement (Warrix and Marshal, 2018). There are very few documented effective management strategies for PC (Gawkins, 2019). The most effective control measure for wild trees is the complete mechanical removal along with the application of glyphosate and triclopyr-based herbicide on cut stumps to prevent sprouting (Swearingen et al., 2002), and the same applies for PC (Coyle et al., 2021; Culley and Hardiman, 2007; Culley et al., 2008; Vogt et al., 2020).

CHAPTER I

MOLECULAR DETECTION OF *PYRUS CALLERYANA* USING LOOP-MEDIATED ISOTHERMAL AMPLIFICATION (LAMP)

Abstract

Callery pear (*Pyrus calleryana* Decne.) was introduced to the United States from Asia to help manage fire blight in edible pear species. Since then, *P. calleryana* has become invasive and is growing as a noxious weed throughout much of the eastern US. From its introduction in the early 20th century, *P. calleryana* has been commercially used as rootstocks for propagating fruiting pears, alongside other pear species such as *P. communis, P. pyrifolia*, and *P. betulifolia*. Several states are restricting the sale of *P. calleryana* rootstocks to limit the ongoing spread of fruit-producing *P. calleryana*. As such, there is a need to develop a field-based detection tool that can reliably detect *P. calleryana* rootstocks. We hypothesized that *P. calleryana* specific primers could be designed and utilized to reliably detect *P. calleryana* DNA in the field within minutes using loop-mediated isothermal amplification (LAMP). We used OrthoFinder to identify the candidate regions in nuclear genome of *P. calleryana* and developed primers for a rapid, sensitive and specific loop-mediated isothermal amplification (LAMP) assay to successfully detect invasive *P. calleryana* DNA and to differentiate it from the DNA of other three closely related species. *P. calleryana* can be detected as early as 20 minutes by incubating Warmstart colorimetric 2× LAMP master mix, LAMP primers and 1 µl of 5ng /µl DNA at 65 °C. The detection limit for 30 min reaction is 0.5 pg ≈ 10 copies of the target region. This tool will help implement regulatory measures and assist land managers in restricting the spread of *P. calleryana.* The tool may also be potentially adjusted to detect other invasive plant species.

Introduction

Invasive species

Invasive plant species are those species that grow outside their naturally occurring habitats, proliferate, alter, and threaten the native species (Shezi and Poona, 2010; IUCN,2000). Invasive plant species are considered one of the major drivers of biodiversity loss (Gurevitch and Padilla, 2004; Maxwell et al., 2016) that can alter ecosystem functioning (Topin, 2018) including the changes in soil properties (Li et al., 2006; Pejchar and Mooney, 2009; Vilà et al., 2011). Invasive species have the potential to alter ecosystem structure and species composition by suppressing or displacing native species. This can occur through direct competition for resources or indirect effects on nutrient cycling. Among the 680 animals listed in the 2003 IUCN Red List Database, causes of extinction were identified for 170 species. In

approximately 20% of these cases, interactions with invasive species were identified as the sole factor driving animal extinction (Bellard et al., 2016). Invasive species have notable negative impacts on human health and well-being, agriculture, tourism, and ultimately threaten the economy, social and environmental stability, and sustainability of the planet (Pimental et al., 2005; McNeely et al., 2001; Kariyawasam et al., 2019). Invasive species interact with other drivers of global change to exert complex and combined effects (Pyšek and Richardson, 2010). The cost of damage by invading species was estimated at US$18.6 billion per year between 1970 and 2017 and expenditures for managing all invasive species averaged US$1.4 billion per year (Diagne et al., 2021). Most invasive plant species now established in the United States (US) were introduced for food, fiber, or ornamental purposes (Li et al., 2006; Richardson et al., 2000) and their spread is causing well-recognized issues. Thus, there is an urgent need to develop effective management strategies to limit their invasion.

Callery pear (*Pyrus calleryana* Decne.; PC), a member of the family Rosaceae, is a diploid plant with a chromosome number of 2n=2C=34 (Phillips, et al., 2016) and an estimated genome size of 588 Mbp/1C (Dickson, et al., 1992). PC is a deciduous tree with a height ranging from nine to 15 m. It is an upright-branched ornamental tree native to China, Taiwan, Korea, Vietnam, and Japan with an environmentally broad native range (Cuizhi and Spongberg, 2003). PC has a pyramidal to rounded crown, alternately arranged oval, long, glossy dark green (above) leaves with crenate margins. White, five-petaled flowers bloom for 1-to-2 weeks, from mid-February to early April, depending upon location and climate, and produce a strong rancid odor, variously described as similar to semen, rotting flesh, or chlorine (Lapidos, 2013). The fruits, which are inedible to people, are hard, until softened by frost, and are green/brown in color, white to tan dotted, spherical to slightly oblong, with 1 or 2 seeds (Vincent, 2005). The twigs and branches of naturalized invasive PC often bear sharp thorn-like spurs.

PC, which was brought to the US from Asia in the early 1900s to aid in management of fire-blight bacterium (*Erwinia amylovora*; (Burril) Winslow et al., 1920) in European pear (*Pyrus communis* L.), has been documented invasive by herbarium records (Vincent, 2005). Various experiments revealed that PC was resistant to pests and pathogens which made it a good rootstock for other related and outgroup species (Culley and Hardiman, 2007; Culley, 2017). ‘Bradford’, the first ornamental PC cultivar officially released by USDA in 1963 (Whitehouse et al., 1963), is considered the most famous and widely planted among approximately 25 PC cultivars that have been commercialized for landscape use (Sapkota et al., 2022; Whitehouse et al., 1963). PC has enjoyed great success in urbanized areas due to its

hardiness and aesthetics, with beautiful white blossoms in early spring and bright red leaf colors in the fall (Culley et al., 2008). Various traits such as its ability to reach sexual maturity at a young age, a self-incompatibility system that promotes outcrossing, resistance to pathogens and pests, and broad tolerance to various environmental conditions, contribute to the ability of PC to spread into a variety of environments (Culley and Hardiman, 2007). Though PC has gametophytic self-incompatibility, various cultivars, rootstock sprouts, and wild escapees may hybridize with compatible pear specimens yielding viable seeds that are then bird- and animal-dispersed (Culley, 2017; Culley and Hardiman, 2007). Even after cutting down PC trees and burning or freezing the cut stems, the trees sprout back from the stump and roots (Bowen, 2023; Maloney et al., 2023) and may again contribute to the invasion inoculum. It is an aggressive invader that has naturalized fallow fields, roadside edges, grassland habitats, and other early successional environments (Vincent, 2005; Woods et al., 2021). PC can alter habitats by decreasing soil pH and increasing soil C: N (Woods et al., 2021). Also, the leaf and flower leachate of PC has shown allelopathic effect on grassland species reducing their germination rate (Woods et al., 2023). The species has the potential of becoming one of the most problematic invasive plants in the US (Coyle et al., 2021; Sapkota et al., 2021). PC is widely distributed throughout the eastern US and in several western states (EDDMapS, 2024). PC is considered an invasive plant in 29 states and the sale has been banned in some areas including the State of South Carolina, Ohio, Charlotte, NC, and Pittsburgh, PA (Blue Ridge PRISM, 2024). The cultivated trees pose a hazard to people and property as well because of the narrow crotch angles that cause branches, especially in older trees, to split when strained by wind, ice, or snow (Culley, 2017). The naturalized tree in the wild forms dense thickets with thorns, challenging native species by creating competition and rendering it difficult for individuals attempting removal due to its impenetrability. (Culley, 2017; Higgins, 2018). Under optimal growing condition, PC trees produce hundreds to thousands of fruits in a single growing season; the dropping of the frost-softened fruits on the ground especially during winter can cause unsightly litter as well as pose a risk to the people walking (Fulcher, 2002).

Land managers are facing never-ending issues to manage invasive species, which shifts resources from other management needs (Renz et al., 2009). The management implementation can be even more challenging than usual when the control measures are difficult to implement (Warrix and Marshal, 2018). There are very few documented effective management strategies for PC (Gawkins, 2019). The most effective control measure for escaped trees is the complete mechanical removal along with the application of glyphosate and triclopyr-based herbicides on the cut stumps to prevent sprouting (Swearingen et al., 2002), and the same applies for PC

(Coyle et al., 2021; Culley and Hardiman, 2007; Culley et al., 2008; Maloney et al., 2023; Vogt et al., 2020).

Comparative genomics

Comparative genomics seeks to understand the similarities and variations in genomic features, tracing their origin, changes, and losses across various evolutionary branches (Xia, 2013). It offers a means to uncover the connection between genomes by detailing conserved (or homologous) chromosomes or chromosomal regions among closely related species. Identification and characterization of the blocks of genes conserved across different species, i.e., syntenic relationships among the genomes of plants within the same family serves as a valuable tool for pinpointing markers in specific regions (Fleury et al., 2012). Conventional plant identification methods that are based on morphological characteristics face limitations, as morphology can be influenced by environmental and/or phenological factors (Hebert et al., 2003). Comparisons of whole genomes, plastidial DNA (ptDNA) and mitochondrial DNA (mtDNA) have become increasingly relied upon for identifying, classifying, and determining the evolutionary relationships of plant species (Jansen et al., 2005; Leigh et al., 2013; Nowicki et al., 2018; Wang et al., 2019; Wang et al., 2024). The ptDNA, primarily inherited maternally in angiosperms, proves useful in identifying distinct species and even different populations of the same species. It offers the advantages of convenience, accuracy, and cost-effectiveness in exploring the systemic evolution, classification, and identification of plant species (Li et al., 2015; Semerikova and Semerikov, 2014). Similar to plastidial genomes, mitochondrial genomes in plants also have predominantly maternal inheritance (Morley and Neilsen, 2017), but have generally lower mutation rates compared to ptDNA (Smith et al., 2015; Wolfe et al., 1987). The ptDNA have conserved architecture, i.e., remain relatively unchanged throughout evolution. Mitochondrial genomes have variable sizes, both among and within lineages, complex structures, multiple RNA-editing processes, frequent reorganizations, and gene loss during evolution (Bi et al., 2016; Smith and Keeling, 2015). Regions of nuclear DNA have also been used for phylogenetic and evolutionary studies (Zuntini et al., 2024). The ptDNA resources of various *Pyrus* species including PC can be found in the genetic databases. PC has a plastidial genome size of 159,965 bp (Nowicki, et al., 2022). Among *Pyrus* species, *P. phaeocarpa* R. has the largest documented plastidial genome size of 160,203 bp (Xiang et al., 2019). There has been little progress in the development of mtDNA and whole genomic resources for *Pyrus* species. *P. pyrifolia* has been reported to have a circular mitochondrial genome that is 458,873 bp long (Chung et al., 2017). Comparative genomics of nuclear,

plastidial, and mitochondrial genomes of PC will help in identifying unique molecular markers specific to at least the species level.

LAMP

Loop-mediated isothermal amplification (LAMP) is a single-tube amplification technique in which the reaction can be processed at a constant temperature (Notomi et al., 2000; Tomita et al., 2008). LAMP can be performed within a relatively short time, as short as 15 to 30 min using a simple device and has been widely used for nucleic acid detection throughout the tree of life (Kurosaki et al., 2007; Ravan et al., 2016; Kitamura et al., 2018). Amplification and detection of genes are completed in a single step by incubating the DNA template, DNA polymerase, and the reaction substrates at constant temperatures (Sahoo et al., 2016). LAMP has a sensitivity of 10- to 1000-fold higher than that of conventional PCR (Notomi et al., 2000). The high specificity and sensitivity as well as the rapid and simple features of LAMP make it clearly different from the existing conventional molecular tests (Ranjbar and Afshar, 2015; Tomita et al., 2012). LAMP normally uses a set of six primers that are complementary to eight regions of the target DNA. The use of four core primers, namely forward (F3) and backward (B3) outer primers and forward (FIP) and backward (BIP) inner primers result in high specificity and efficiency, and the addition of even a single or both forward and backward loop primers (FLP and BLP) enhances the performance and rapidity of LAMP assay (Duan et al., 2014; Parida et al., 2008; Notomi et al., 2000; Nagamine et al., 2001). The amplification duration can as well be reduced to less than 30 minutes by using loop primers (Nowicki et al., 2015; Ranjbar and Afshar, 2015). The inner primers, each are formed by the concatenation of two distinct sequences corresponding to the sense and antisense sequences of the target DNA. An isothermal *Bst* DNA polymerase (acting at 55 to 65 °C) with high strand displacement activity in addition to replacement activity is the prime requirement in this technique (Oscorbin et al., 2017).

The LAMP reaction leads to the formation of amplification products as well as insoluble magnesium pyrophosphates, which can both be used for direct detection (Focke et al., 2013; Mori et al., 2001). Detection of turbidity by the naked eye is the simplest and the most cost-effective method for the assessment of positive or negative LAMP reactions (Goto et al., 2009). Various DNA intercalating or pH-sensitive dyes such as SYBR green, Picogreen, or propidium iodide can be added to the solution pre- or post-reaction to observe the reaction result through color change or fluorescence under UV light (Goto et al., 2009). In addition to DNA, LAMP can detect RNA by the reverse transcription LAMP reaction (RT-LAMP).

Though LAMP has many advantages, the proper designing of primers is a major constraint in successfully deploying this type of assay (Torres, et al., 2011). Despite the availability of free online tools for primer design, there are situations where the preferred target sites may not be automatically chosen. In such cases, manual primer design remains necessary (Wong et al., 2018). LAMP products being highly stable are less prone to degradation and thus may sometimes pose a risk of unintended carryover contamination when results are visualized through post-amplification detection steps that require opening of the reaction tubes, which may lead to subsequent false-positive results (Bi et al., 2012; Dhama et al., 2014; Fischbach et al., 2015; Hsieh et al., 2014). Optimization of reliable, rapid, and sensitive on-site diagnostic protocol based on the LAMP reaction could be an effective tool for the early detection of invasive PC in scouting efforts and implementation of bans by several US states. The use of battery-powered portable miniPCR machine as light as 500 g, that can be directly connected to the cellphone through Bluetooth allows the adaptability of the protocol in a variety of settings, real-time data monitoring, and rapid diagnostics.

The rationale of the study

P. calleryana has been used commercially since early 1900s as rootstocks for the grafting and propagation of the fruiting pears alongside other pear species such as *P. communis*, *P. pyrifolia* Burm.f. Nakai, and *P. betulifolia* Bunge. Viable fruits can be produced from the shoots originating from PC rootstocks even when the grafted tree dies. Also, PC species and naturalized escapees may hybridize with compatible specimens and produce viable seeds. The seeds may then get dispersed by birds or animals. The most effective control measure for PC is complete removal and replacement. To limit the ongoing spread of invasive PC, several states are restricting the sale of rootstocks. As such, there is a need to develop a field-based detection tool that can reliably detect PC rootstocks. Such a tool will help land managers restrict the spread of PC. As different states are planning or implementing to ban the sale and use of PC, the application of this tool will certainly support the evoked regulatory measures. The tool may also be potentially adjusted to detect other invasive plant species.

Objectives

Broad: To develop a rapid field detection tool for *Pyrus calleryana* (Callery pear) using plant tissue samples

Specific:

- **a.** To identify and characterize candidate regions of *P. calleryana* through comparative genomics of ptDNA, mtDNA, and nrDNA across *Pyrus* species
- **b.** To evaluate candidate regions and test their suitability in the development of the LAMP protocol
- **c.** To develop and optimize a LAMP assay for the field detection of *P. calleryana*

Hypothesis

1. *P. calleryana* DNA can be extracted in the field using a plant tissue sample and detected within minutes using a field detection tool and the LAMP method

2a. There exist candidate regions in plastidial, mitochondrial, and nuclear genomes that ensure the species-specificity of *P. calleryana*

2b. Candidate regions can be screened and used to develop LAMP assay

2c. LAMP can be used as a tool to detect *P. calleryana* DNA in the field

Materials and methods

Plant materials

The PC leaf samples requested from J. Hartshorn, Clemson University and Sapkota et al., (2021) were utilized for this study. The four vouched cultivars of PC were also requested from Arnold Arboretum of Harvard University. Similarly, the available accessions of other three species of interest were requested from GRIN, USDA, which included 96 leaf samples of *P. pyrifolia*, 48 samples of *P. betulifolia*, and 87 samples of *P. communis* (Supplementary Table 1).

Identification of plastidial (ptDNA) and mitochondrial (mtDNA) candidate regions of Pyrus calleryana

The bioinformatics comparisons of plastidial (ptDNA) and mitochondrial (mtDNA) genomes were performed across *Pyrus* species whose rootstocks are widely sold by the commercial nurseries and fruit tree producers, along with *P. calleryana* rootstocks (Table 1.1). mVISTA, which is a computational tool for comparative genomics (Frazer, et al., 2004) was used for comparative genomics of both ptDNA and mtDNA. The ptDNA of *Pyrus communis* (NCBI, MN577870), *Pyrus pyrifolia* (NCBI, MZ673796), and *Pyrus betulifolia* (NCBI, ON478188) were aligned against the chloroplast genome of *P. calleryana* (NCBI, OM541581) in

mVISTA and then used to identify the regions that presented sequence similarities and differences. Shuffle-LAGAN (Brudno et al., 2003) was used as an alignment algorithm behind mVISTA because this tool is able to find rearrangements that occur within a global alignment framework. The mVISTA output view for a 100 bp sliding window that employs PC as the comparative basis is shown in Figure 1.1. The heights of the peaks indicate the percent of conservation between the organisms at any given coordinate.

For comparative genomics of the mtDNA genomes, firstly a draft mtDNA genome of PC was assembled using NOVOWrap v1.20 (Wu et al., 2021), because there was no mtDNA genome available for PC. The details of genome assembly are presented in Chapter 2. Keeping our draft mtDNA genome as the base, the mitochondrial genomes of *P. communis* (NCBI, NC_065229.1), *P. pyrifolia* (NCBI, KY563267), and *P. betulifolia* (NCBI, MW080658) were aligned in mVISTA in a similar fashion as was done for the ptDNA.

Design of LAMP primers through NEB LAMP Primer Design Tool

The six LAMP primers required for the LAMP reaction include the core set of two outer primers (forward, F3; backward, B3) and two inner primers (forward inner, FIP; backward inner, BIP), and two optional loop primers (forward loop, FLP; backward loop, BLP) (Notomi et al., 2000). **NEB LAMP primer Design Tool** (version 1.4.1) was used to design LAMP primers from the identified, potential candidate regions. This tool allows to input target DNA sequences to obtain sets of core primers optimized for LAMP. In the subsequent step, it gives an option to generate loop primers based on a given set of core primers. All designed primers were synthesized by Integrated DNA Technologies (IDT Inc.; Morrisville, NC).

Whole genome-based LAMP primer designer (GLAPD)

GLAPD is a software used to design target organism(s)-specific LAMP primer sets using whole genome sequences as input (Jia et al., 2019). This tool first identifies all possible single primer regions from the reference genome pool, aligns them with the target genome(s) and background (off-target) genome(s) to ensure that the primers are specific to the targets and then combines them into LAMP primer sets (Figure 1.2). GLAPD has an optional flag in the command line for whether to generate loop primers or not. In this study, the available whole genomes, ptDNA, and mtDNA genomes of different *Pyrus* species of focus were retrieved from databases such as NCBI, the Genome Database for Rosaceae GDR and Comparative Genomics CoGe. Then, the GLAPD was used separately on ptDNA, mtDNA,

and whole genome assemblies and inputted to design LAMP primers. The PC genome assembly was done under default parameters of MaSuRCA (Zimin et al., 2013) using the four Illumina paired-end sequence read archives (SRAs) of PC available in NCBI (SRR7135497; SRR7135498; SRR7135500; SRR16505594). GLAPD required the single concatenated file as input, so the conversion of multifasta draft nuclear assemblies into a single fasta file was done using a custom script in Python v3.12.4. ptDNA and mtDNA required no such formatting as all were already available as a single contig each. The target group included the genomes of PC whereas the background group included the genomes of other three *Pyrus* species named earlier.

DNA extraction

The documented accessions of *P. calleryana* and other relevant species were requested from Germplasm Resources Information Network (GRIN USDA), herbaria, and arboreta for DNA extraction and subsequent testing (Supplementary Table 1). The genomic DNA of different *Pyrus* species was extracted using approximately 100 mg of leaf per sample. The samples were homogenized using a Bead Mill 24 (Fisher Scientific, Pittsburgh, Pennsylvania, USA) and subsequently used for DNA extraction using EZNA DNA DS Mini Kit (Omega Bio-Tek, Norcross, Georgia, USA), following the manufacturer's protocol. The concentration and purity of extracted DNA was assessed using Nanodrop Spectrophotometer (Thermo Fisher Scientific, Wilmington, Delaware, USA).

Evaluation of F3B3 LAMP primers

First, LAMP outer primers (F3/B3) were tested and optimized using polymerase chain reaction (PCR). PCR was done in a 10 µl reaction mixture consisting of 5 µl of AccuStart II PCR supermix (2×), 1 µl each of 10 µM F3 and B3, 2 µl of sterile water, and 1 µl of 1ng/µl DNA. The PC DNA sample 'Keowee mother-1' (Source: J. Hartshorn, Clemson University; Collected on Aug 2022) was used as a positive control. Sterile deionized water was used as a negative control for each primer pair. The gradient PCR followed by the touch-down PCR was done to ensure the specificity of the amplified products. The PCR program used for temperature optimization was: initial denaturation at 94 ℃ for 3 min, followed by 10 cycles of denaturation at 94 ℃ for 15 s, gradient annealing at 53 - 60 ℃ and a touch-down of 0.5 ℃/cycle for 30 s and an extension at 72 ℃ for 30 s, followed again by 30 cycles of denaturation at 94 ℃ for 15 s, gradient annealing at 48 -55 ℃ for 30 s, and an extension at 72 ℃ for 30 s, with final extension at 72 ℃ for 4 min. The amplified PCR products were

subjected to Agarose Gel Electrophoresis with 2% w/v agarose gel stained with ethidium bromide and a 100 bp Invitrogen ladder (Thermo Fisher Scientific) as a reference. The visualization was done under UV light using UVP GelStudio PLUS (Analytikjena, Upland, California, U.S.) and documented using VisionWorks 8.22.18309.10577. The F3B3 primer pairs that showed positive amplification with PC DNA were then tested with other *Pyrus* species, including the outgroup species. The samples that showed positive amplification were gel-extracted using the QIAquick Gel Extraction Kit (QIAGEN, Maryland, USA), and Sanger sequenced (Molecular Cloning Laboratories; MCLAB, California, USA) using both F3 and B3 primers, respectively, to further ensure species specificity. The sequencing results were assembled using Sequencher v5.4.1 (Gene Codes Corporation, Michigan, USA). The assembled data were then aligned in MAFFT v7 (Katoh et al., 2002) under default parameters to visualize polymorphisms among the analyzed species. BioEdit sequence alignment editor v7.2 (Hall, 1999) was used to visualize the alignment results.

Optimization of the LAMP assay

For optimizing LAMP assay, the protocol for *Bsm* DNA Polymerase, Large Fragment (ThermoFisher Scientific) and SYBR Green nucleic acid stain (ThermoFisher Scientific) was followed initially. The recommended protocol contained a 25 µl of reaction volume including 2.5 µl of 10× *Bsm* buffer, 4 µl of $MgCl_2$, 3.5 µl of dNTP mix, 1 µl of *Bsm* DNA polymerase, 5 µl of of LAMP primers, 8 µl of of nuclease free water and 1 µl of 5 ng/µl DNA. The reaction mixture was then incubated in a thermocycler at 60 °C for 30 to 60 minutes. After incubation, 1 µl of SYBR Green I stain 10,000× stock diluted to 1:10 was added to each tube. Green color indicated the successful amplification and orange color indicated no amplification. After several tests, this protocol was abandoned because it gave false positive results for all tested samples, including the water negative control.

After failure of that initial kit, WarmStart Colorimetric LAMP 2× Master Mix (NEB, MA, USA) was used to perform LAMP in a 10 µl reaction volume, following the manufacturer's protocol. For this, a stock solution of 10× primer mix containing 2 µM of F3;B3 each, 16 µM of FIP;BIP each, and 4 µM of FLP;BLP each (whenever applicable) was made for each of the tested LAMP primer sets. The final 10 µl reaction volume contained 5 µl of master mix, 1 µl of primer mix, 3 µl of nuclease free water and 1 µl of 5 ng/µl DNA. To identify the minimum detection time for PC DNA, the LAMP program was run in a thermocycler at 65 °C for different time periods, starting from 30 minutes and going down to 15 minutes with five minutes decrement in each set of serial dilution of PC DNA. To

determine the lowest limit of detection (LOD) of PC DNA, ten-fold serial dilution was done, starting from 5 ng/µl and going down to 5 fg/µl (≈ 10 copies of target DNA sequence).

OrthoFinder for comparative genomics

OrthoFinder: phylogenetic orthology inference for comparative genomics (Emms and Kelly, 2019) tool was leveraged towards identifying nuclear genomic sequences specific to PC. The available datasets of PC (SRR7135497; SRR7135498; SRR7135500, SRR16505594), *P. communis* (ID: GCA_037177615.1; GCA_963583255.1), *P. pyrifolia* (ID: GCA_016587475.1), and *P. betulifolia* (ID: GCA_007844245.1) were utilized for our bioinformatics approach. The orthologs of assemblies from all *Pyrus* species including PC were inferred with OrthoFinder under default parameters. The genes orthologous in 90% of the taxa were extracted with a custom python script into interleaved sequence files. Nucleotide similarity scores were calculated for each position of interleaved sequences with MStatX (Collet, 2012). These similarity scores were averaged for each gene with a custom python script and the genes with the lowest average similarity scores were manually searched for nucleotide polymorphism(s)as compared to PC draft genomes using SeaView (Guoy et al., 2010). The potential candidate sequence with at least three nucleotide polymorphisms specific to PC was extracted and NCBI BLAST (blastn) search was done against the genomes of other closely related species to ensure the specificity of that candidate region. Based on that candidate sequence, LAMP primers were designed using NEB LAMP Primer Design tool and synthesized by IDT (Morissville, NC, USA). Those primers were tested using WarmStart Colorimetric LAMP 2× Master Mix (NEB, MA, USA), as described above. For ensuring the specificity of primers, LAMP was run for 85 DNA samples of PC comprising of Asian, US commercial cultivars, and US escaped populations (Sapkota et al., 2022). For cross-checking the specificity, samples of *P. communis, P. pyrifolia* and, *P. betulifolia* from GRIN USDA were also tested.

Direct DNA extraction method for fresh P. calleryana tissue samples

To determine whether viable DNA can be extracted in real-time field conditions from PC tissue samples for the molecular detection, several experiments were carried out to extract DNA from fresh tissue samples taken from PC, including leaf, flower, bark, and root tissues, and were tested using LAMP. For direct DNA extraction, samples from the PC cultivars maintained at the University of Tennessee's (UTK: Knoxville, TN) Nursery Compound were used. The plants of four different cultivars of PC; Bradford, Autumn Blaze, Aristocrat, and

Cleveland Select were purchased from Ty Ty Nursery, GA in October of 2022, and are maintained during the study at the UTK research nursery. During the time of tissue sampling (July 2023), the trees were around 3 m tall. To find the combination of optimal amount of plant tissue and buffer solution required for downstream LAMP detection, different amounts of fresh tissue (leaf, flower, bark, and roots) were collected in a 2 ml microcentrifuge tube and variable amount of DNA elution buffer (EZNA DNA DS Mini Kit) was added to it. Then, a disposable plastic pestle was used to gently grind the tissue and suspend DNA in the buffer solution (Figure 1.3B). The DNA concentration of each sample was assessed using nanodrop. The DNA was diluted to 10 ng/µl and 5 ng/µl and tested with LAMP primers at 65 °C (Table 1.3).

Results

Plastidial (ptDNA) and mitochondrial (mtDNA) candidate regions of Pyrus calleryana and LAMP primers

From the mVISTA comparison, 19 potential candidate regions from ptDNA (Table 1.2) and 16 from mtDNA (Table 1.3) were retrieved that could potentially be specific to *P. calleryana*. Out of the total candidate regions, nine regions from ptDNA and 13 regions from mtDNA yielded primers under the default parameters of the NEB LAMP primer design tool.

Evaluation of F3B3 primers

The amplified F3B3 PCR products were visualized using UV light after electrophoresis in ethidium bromide-stained agarose gels, with a 100 bp Invitrogen ladder (Thermo Fisher Scientific) as a reference. The expected size of PCR product was ~300 bp. After optimization, all of the nine ptDNA F3B3 primers and 13 mtDNA F3B3 primers showed positive amplification with PC DNA (ID: Keowee mother-1). The primers were then tested with DNA samples of other *Pyrus* species and the outgroup species *Malus rockii* Rehder (Figure 1.4; Figure 1.5). The samples that showed positive amplification were gel-extracted and Sanger sequenced to further ensure species specificity. The sequencing results were assembled and aligned to visualize polymorphism among the species. Based on PCR and Sanger sequencing data, two promising candidate loci from ptDNA and three loci from mtDNA that were polymorphic to other tested *Pyrus* species were selected for downstream analysis (Figure 1.6; Figure 1.7).

LAMP specificity assessment using NEB and GLAPD primers

Initially, all NEB LAMP primers containing at least one loop primer for the promising loci from both ptDNA and mtDNA were tested under LAMP at 65 ℃ for 30 min. Out of all tested primers, the mitochondrial primer set mt169 kbp containing all six LAMP primers (Table 1.3; Figure 1.7) showed negative amplification of *P. pyrifolia* DNA samples. However, all the primers that resulted in positive amplification of PC DNA, also resulted in positive amplification of *P. betulifolia* and *P. communis* DNA. The primer set mt169 kbp was then tested with several samples of *P. pyrifiolia* obtained from USDA GRIN for further validation (Figure 1.8A). Though mt169k could differentiate PC DNA from *P. pyrifolia* DNA, we were not able to develop a primer set that could differentiate PC from all three non-target species of interest from our initial screening.

GLAPD did not yield any primers for ptDNA or whole genome inputs, however, this analysis yielded five primer sets for mtDNA. The GLAPD primer set mt26k differentiated PC DNA from both *P. betulifolia* and *P. pyrifolia*, however, the primer set mt67k yielded positive amplification for *P. communis* DNA (Figure 1.8B). As such, from our second attempt, we were able to generate the primer set that could differentiate PC from at least two other non-target species of interest.

The LAMP primer sets without the loop primers were also tested in an attempt to develop PC-specific LAMP protocol. Out of 15 mtDNA primer sets without the loop primers, only four primer sets showed positive amplification with PC DNA and those four primer sets were tested with other three species as well (Figure 1.8C). Except for the generally unresponsive mt87k, all other primer sets showed negative amplification for *P. betulifolia* and *P. pyrifolia*, whereas the result was positive for *P. communis.*

OrthoFinder towards LAMP primer identification

The candidate sequence specific to PC identified by utilizing the OrthoFinder approach coded for an MYB gene located on chromosome 12 (Li et al., 2016). Two sets of LAMP primers resulted from that candidate sequence (Table 1.4). The core set of inner and outer primers for both the sets of primers was the same, the difference being the two alternative loop backward primers, LB1 and LB2. No convenient loop forward primer LF was generated for this region. The two primer sets, hereafter referred to as MYB1 and MYB2, contain the LB1 and LB2 respectively. In a preliminary test, the primer sets proved to be specific to PC, showing positive amplification within 30 min for only the PC DNA samples (Figure 1.9) within 30

min. When tested without the loop primer, there was no positive amplification after 1 hr long assays. For specificity confirmation, the primers were tested using a collection of DNA from non-target *Pyrus* species (Supplementary Table 1). Final confirmation was performed using the primer set having LB1 (Table 1.5). Out of 95 PC samples, 75 samples showed clear change in color to yellow representing positive amplification, whereas there were five ambiguous results, i.e., there was no distinct change in color when tested for 30 min. However, when the reaction time was increased to 40 min, there were 80 positive amplifications (Table 1.5). For *P. communis*, none of the 87 tested samples resulted in a positive amplification: all were negative. Out of 90 *P. pyrifolia* samples tested, 78 samples showed negative amplification and the remaining 12 samples had ambiguous results. Finally, out of 48 tested samples of *P. betulifolia*, 38 samples gave a negative result, 6 samples showed ambiguous results, and four samples yielded a positive amplification result. All these results show the high specificity of the developed primers. To ensure that the negative LAMP results of *Pyrus* species other than PC were not due to degraded DNA, the same DNA samples were tested using mitochondrial SSR 417k developed in our previous study (see Chapter 2) which had proved to have cross-amplification capacity. Agarose Gel Electrophoresis with 2% w/v agarose gel stained with ethidium bromide and a 100 bp Invitrogen ladder showed successful amplification for all samples to underscore the viability of the tested DNA samples (Supplementary Figure 1).

Sensitivity of LAMP assay

The sensitivity of the LAMP assay was tested using a ten-fold serial dilution of PC DNA (Keowee mother-1). Sensitivity was tested separately using two different primer sets mt169k and MYB1 (Figure 1.10 and Figure 1.11). The primer set mt169k contained all six LAMP primers whereas MYB1 contained the five LAMP primers. For mt169k, positive amplification of PC DNA was detected within 20 minutes using 5 ng DNA under LAMP. The detection limit for 30 min reaction using the mt169k was 0.5 pg, which is a DNA mass equivalent to about 7 copies of the target region. This indicates the high sensitivity and rapidity of LAMP with all six primers, making LAMP a reliable tool for even low abundance targets. For MYB1, positive amplification of PC DNA was detected within 30 minutes using 5 ng DNA under LAMP. The detection limit for 30 min reaction was 5 ng of the target region, whereas extending the reaction time up to 50 min yielded 100-fold increase in detection – 50 pg.

Applicability of direct DNA extraction for LAMP

To evaluate the feasibility of isolating DNA in the field and performing a LAMP assay within minutes, the DNA extraction method was rigorously tested. PC DNA was extracted successfully from fresh tissue samples, i.e., from each of leaf, bark, flower, and root samples by grinding each plant tissue in 100 to 300 µl elution buffer in a tube. The extracted DNA was then used immediately in LAMP assays. The pipetted 1 µl of freshly extracted DNA was mixed with the LAMP primer set 'mt169k' and LAMP mix and incubated at 65 ℃. In many cases, there was no positive amplification without dilution. However, after approximately 200-fold dilution to obtain 10 ng/ul, positive detection of PC DNA could be seen as early as 20 minutes (Table 1.6).

Discussion

Invasive species pose complex ecological and economic challenges, making their rapid and accurate detection crucial for effective management. Numerous LAMP-based methods have been described in the literature, specifically to detect foodborne pathogens, pathogenic microorganisms, or invasive pests in plants (review by Wong et al., 2018). LAMP has also been used to detect genetically modified crops (Li et al., 2015; Zhang et al., 2013), herbal medicines (Li et al., 2016), and animal species (Lee et al., 2016). But, LAMP methods for the detection of invasive plant species are not described in the literature. In this study, a rapid, sensitive, and specific LAMP assay was developed to successfully detect invasive *Pyrus calleryana* DNA and to differentiate it from the DNA of other three closely related species, namely *P. communis, P. pyrifolia,* and *P. betulifolia.* LAMP assay was developed and optimized using a set of four core primers and one loop primer at 65 ℃. The color of reaction mix changed from pink to yellow in positive tubes and remained unchanged in the negative samples. Fast and easy interpretation of the assay results, a very simple method for direct DNA extraction in the field, and a requirement of stable water bath or heat block makes this assay an easy-to-use method. One of the things to consider for direct DNA extraction in field is the dilution because in many cases, there was no positive amplification without dilution possibly because of high concentration of DNA, high viscosity of the solution, or the presence of inhibitory substances that might interfere with LAMP.

Only the loop backward (LB) primer was generated based on the core primers by NEB LAMP primer designing tool for the MYB target locus. The absence of successful amplification even after 1 hr when tested with only four core primers emphasizes the importance of loop primers in LAMP (Nagamine et al., 2002; Soroka et al., 2021). The

addition of one loop primer increased the specificity and speed of LAMP reaction in our study. The lower detection time and limit when using the primer set containing both the loop primers further underscores the role of loop primers. The LAMP assay was rigorously tested on a comprehensive set of DNA samples to validate its specificity and sensitivity. There were only ~15% false negatives in the PC samples when tested for 30 min and only ~11% when tested for 40 min, and no false positive in *P. communis* and *P. pyrifolia* samples, indicating the desirable high specificity of the designed primers. The false negative results in PC could be due to the suboptimal quality of PC DNA used in this study as major parts of the PC samples were obtained from historical herbarium specimens, or the reaction time shorter than that required for successful target amplification. There were about 4% false positives in case of *P. betulifolia.* The high risk of carryover contamination, which often leads to false positive results is one of the major challenges of LAMP (Hsieh et al., 2014), which is less likely to have happened in our study as post-LAMP opening of tubes to verify the success of amplification was not required. Another possible cause of false positive results is the use of multiple primers in LAMP which will increase the chances of primer-primer hybridization (Kim et al., 2023). Furthermore, as LAMP is highly sensitive – as additionally documented in our assays, cross-contamination between samples can produce false positive results (Suleman et al., 2016).

Though LAMP has many advantages, the proper design of primers is a major constraint to success with this type of diagnostic assay (Torres, et al., 2011). A robust assay requires selection of unique and specific genomic regions that are well conserved within the target genomes (Domingo et al., 2021). Even in this study, designing specific primers was the most challenging task for the LAMP assay. Initially, attempts were made to design primers through comparative genomics of plastidial and mitochondrial genomes of different *Pyrus* species. These approaches, however, proved inadequate due to high level of sequence conservation in those genomes across *Pyrus* species, leading to a lack of specificity in distinguishing PC from at least one other related species. Numerous independent bioinformatics approaches and algorithms including GLAPD, MAFFT alignment, ProgressiveMauve alignment (results not shown) were employed towards identifying the unique putative target markers, but none of these approaches yielded primer sets to specific PC. Ultimately, the application of OrthoFinder (Emms and Kelly, 2015), a comparative genomics pipeline, using the draft whole genomes of PC and other *Pyrus* species, led to the identification of unique regions specific to PC, thus facilitating the design of specific markers.

The development of a LAMP assay for detecting invasive PC is a pivotal advancement in the management of invasive species. This assay is especially relevant as numerous US states have begun restricting or banning the sales of PC rootstocks to curb its spread. This LAMP tool offers a desirable detection method for field use due to the portability of miniPCR machine, simple color-change interpretation of results, and the fact that it requires no detailed prior knowledge of molecular detection. With just a simple training, it can be deployed easily in field which makes it highly accessible and practical for on-site detection.

APPENDIX: Tables and figures

Table 1.1. *Pyrus* species used in comparative genomics of ptDNA and mtDNA.

***Pyrus* species**	**Genome**	**Genome length (bp)**	**NCBI number**	**Source**
Pyrus calleryana	Plastidial	169,965	OM541581	Nowicki, et al., 2021
Pyrus communis	Plastidial	160,171	MN577870	Liu, et al., 2019
Pyrus pyrifolia	Plastidial	160,214	MZ673796	Wang, Y., 2021
Pyrus betulifolia	Plastidial	160,119	ON478188	Sun, M., 2022
Pyrus calleryana	Mitochondrial	458,892		This study (Chapter 2)
Pyrus communis	Mitochondrial	443,525	NC_065229	Sun, M. and Chen, X., 2022
Pyrus pyrifolia	Mitochondrial	458,873	KY563267	Kim, et al., 2017
Pyrus betulifolia	Mitochondrial	469,928	MW080658	Yuan, et al., 2020

Table 1.2. Potential candidate regions from plastidial (ptDNA) genome of *Pyrus calleryana*, and the LAMP primers generated by NEB LAMP primer design tool.

Approx. position of candidate regions in ptDNA	LAMP core primers	LAMP loop primers
4.2 kbp	Not Available	Not Available
6.8 kbp	Not Available	Not Available
7.5 kbp	Not Available	Not Available
9.8 kbp	F3-GTTCTGAATACATTCAGACAGA B3-AAGTAAAAGCGGGTAGCG FIP-CGGCTCCTTTATGGAAGATGTATAAGCTGACATAGATGTTATGGGT BIP-CACGTTCAGTTTTGAATTAGAGACGGTTAGCTTGGAAGGCTAGG	LF- Not available LB-TGATGAATCAACGTCGACTATAACC
15 kbp	Not Available	Not Available
32 kbp	Not Available	Not Available
35 kbp	Not Available	Not Available
51 kbp	F3-CTTGTACTGTACTGAAACTTTGA B3-TACTCATGTGCCAGGAAC FIP-GGACTTGGGTCTATGTCAATTAGAAGATCCAAGAAATTCTATTAGATCCT BIP-GGATGATGCGTCATGAATGGTAGATTTGAACTGGTGACACG	LF- Not Available LB-CTCAGCTGGTAGAGCAGAGGAC
59 kbp	F3-TTATTTTGTGCCGAAGCAA B3-TGAGCCAAGGTAGTATTTGC FIP-GTACCTGCAGTAGCGTTTAAGTAATTTATAAAGCACAGGCTGAAAC BIP-ATGCGAAGATATGATGAAAAGAGCTCCCCTGTTAAGTAATCATGC	LF- Not Available LB-CCAGAGAATTGGGGGTTCCAATC

Table 1.2. Continued

66 kbp	Not Available	Not Available
72 kbp	F3-TTTTTTCTAGGTGAAATGCCTA B3-GGATAATTATGGGTAGAGCCAA FIP- CATAGAAACGATGGAACCCACTATAGGTATCAAAACCACATAAAAGTTG BIP-TCTTAAACGGTGAGGTCCTCTCTATAGAGTGTGAACTGTACAAGTT	LF- Not Available LB-CACCGGAGCCCTTCTTTCATTTAAT
83 kbp	Not Available	Not Available
92 kbp	F3-ACAAAAAGAAAGGAAGTGACTT B3-GAGATCCATGAATCGGGAT FIP-CAAGTAGTGTTCGATCGATTACGTATTTTGTCGATAACCTCAGACC BIP-ACGGCTCTTCTGCTCAGAAATGTGCATATAGATACAAATGGTCC	LF- Not Available LB-ATGTTCCTGTAAATTCTTGCTCCCA
103 kbp	F3-CCTATGTTGTGTTGAAGGGATA B3-CCTCAGTCCACAGAGACAA FIP-CTGTATACTTTCCCCGGTTCCGCTATATGATCCGATCGATTGC BIP-TTAGTTAGTGATCTCGGCTCAGTATGTAGGACTGGTGCCAA	LF-TGCTACCGCGGGCTTTA LB-GAGTCCTTTCTTCCGTGATGAA
116 kbp	Not Available	Not Available
118 kbp	Not Available	Not Available
135 kbp	F3-AGGCCATTCCATTTCGACAA B3-ACAGATCCAGTGGAGACGG FIP-GACCTTTCCCTTTGGGGGTAGGGACCCACGCCCAAATTCC BIP-GTTGTGGGCGAGGAGGGATTCGCTCAGAGGATTAGAGCACG	LF-AAAATCATGATCGGGATAGCGGAC LB- Not Available

Table 1.2. Continued

144 kbp	F3-CTTTCTCCTCAGTCCACAG	LF-TCCGTGATGAACTGTTGGC
	B3-CCTATGTTGTGTTGAAGGGATA	LB-TTCCGTTGCTACCGCG
	FIP-TCGGCTCAGTGAGTCCTTTCAGACAAAATGTAGGACTGGT	
	BIP-ACTGTCTTTTCTGTATACTTTCCCCATCCGATCGATTGCGTAA	
154 kbp	F3-ACAAAAAGAAAGGAAGTGACTT	LF- Not Available
	B3-GAGATCCATGAATCGGGAT	LB-ATGTTCCTGTAAATTCTTGCTCCCA
	FIP-CAAGTAGTGTTCGATCGATTACGTATTTTGTCGATAACCTCAGACC	
	BIP-ACGGCTCTTCTGCTCAGAAATGTGCATATAGATACAAATGGTCC	

Not available: NEB LAMP Primer Design Tool did not yield primers

Table 1.3. Potential candidate regions from mitochondrial (mtDNA) genome of *Pyrus calleryana*, and the LAMP primers generated by NEB LAMP primer design tool.

Approx. position of candidate regions in mtDNA	**LAMP core primers**	**LAMP loop primers**
21 kbp	F3-ATGGTGCTCGTTTTTCCA	LF-ACTAGGTCCAATGTGAGTAGGA
	B3-TTATTCCAGATGCTCGCC	LB-CGGGGGTTTCCGAGGAA
	FIP-TATTGGCCAAACCACCTGGGATTATGAAGCATGGCTAAGTGA	
	BIP-GTGGGTCAAGAAATATTGAATGGTGAAGCTGAAAAAAACCAGAGG	
22 kbp	F3-AGCTGGAAAATGCAGTCAGA	LF- Not available
	B3-TGGCTCACACTCTGTCCA	LB-CCAGTTCCAGAAGGAAAGTAAC
	FIP-CTCTGCGCAAAGTTTCCTCCCGTGGGGTAGTATCAGCGATCA	
	BIP-CCTTGCGCGCTTTAGGAGGAGGAACCCGTATTTTCGTCGT	

Table 1.3. Continued

27 kbp	F3-GAAGTCTCTTGCGGGTAT	LF-GCGTGGGAGAAAGAAGGA
	B3-TGACCAGAAGAATTGTGTGA	LB-GGAGTTAAGCTGTCTGGAGG
	FIP-GGTTGGTTGTTGACCAACGGTTTGATCTCTGGCTGTGAC	
	BIP-CTTCACTAATCCTACAGGCTTGACAGTTGAGGCATTCTTGATTG	
66 kbp	F3-GTGATCTTTCGTTTTTTCCATT	LF- Not available
	B3-CTGAAACTGCGCTGAATG	LB-GTCGTAAATCGCCGGTTGG
	FIP- AGCAGTAAAAGAATGAAGGAAAGGACATAATAGATTGCCATACATTGAGA	
	BIP-TTGCTCCCCCAAAAAAAATGAAGTAAAGCCTTATTTCCATGACTT	
87 kbp	F3-TACTGCCCAGAGGAGAAG	LF-AAAATGTCACGGGTCCCA
	B3-TGGAGGAAATAACGATGTGG	LB-CAATTGGTCTACATGGTCATTGTAG
	FIP-AGCGCCTTTTTTTGGGGTACTCATAGGTAGGAATGACAGGA	
	BIP-CAAGGAGCGCTACATCCCTTAAAACAGGGCAGGGATTC	
90 kbp	Not Available	Not Available
102 kbp	F3-GTCCTTTGGACCCTTCGC	LF-GATAAAGGGGAAAGGGGCAAC
	B3-CCCAAGCCTACGTTTGCT	LB- Not available
	FIP-TGCTGGGCTTGAATGCTTACCTCAAGCTAGCCCCCTATGTTG	
	BIP-AGAAGGGGCCCCCCCTAATGAGTAAGCTTGAGCTGCGC	
169 kbp	F3-CATGTAATTAATGTTCCAAAGGAAG	LF-GTCATCCTCGCTAGCTTTTTTAAC
	B3-TCCCAGCGAACTACCATT	LB-GCTTTTCTTGTGCATGACGA
	FIP-TCCCCACCCTCCTTACTAAATTTAAAATGATGAAACGTATTTCGGA	
	BIP-GTTTTAGCGGAAACGCGACCGTGACTTTGTTCACCCGG	

Table 1.3. Continued

231 kbp	F3-CACATATCTAACCCAGTTTGG	LF-CGTCCCCCTGCTTGAAT
	B3-CGGATTGCTTTTTCTTTCTCTC	LB- Not available
	FIP-GATCCTCACTGCCTGTTCGGTGGTGGGTCATATTGAGGTA	
	BIP-GAAACCCCGAGAAGCCACAACTGGGTAAAGTAGTATCCCTT	
243 kbp	F3-CCTTATTAGATGACTCGGCC	LF- Not available
	B3-GAAGAATGTGACTCAGGAAAT	LB- Not available
	FIP-CAGGAGAACAAGTGCAGGAATTTTAATTAGCCTGAGCATAGTAAAGTG	
	BIP-TGAAAACTGAACCCTGGTCAGCTCATTAAGGTCCTGAAGGAG	
304 kbp	F3-GTAAGGTGCTGCTCTTCGC	LF-GGGCCTAAGTGAAAGTGAAGT
	B3-TCTCGGGGCATCTGAGAA	LB- Not available
	FIP-AGGGTCCCCCCACTGACTTGATACCCCTGGTTGGGATGG	
	BIP-CTCCCCCCCCCCCCAAAAAAATGATTCGAGCCGTATGAAGG	
338 kbp	Not Available	Not Available
385 kbp	F3-GGTTGTTGTACTGAAAGTGG	LF-TCCTTTCTACTGCTACAGGTTC
	B3-TCACTATTTCACCCAAGAGAG	LB-GTATAAGGTAGCGCCGGAGT
	FIP-ATCTGTCCTTAGGAACATGGATTTTGAAGCAAATGAATCGAGTAGC	
	BIP-GAAGCGAATGGATAGCTCGTGGCCGACCGAGTAATGTAGGA	
395 kbp	Not Available	Not Available
402 kbp	F3-TTGCAGGAGAGGGGTTCA	LF-TCTTTTTGTGGGGGGGTAAT
	B3-AGCAACTCTACTACCGCCTT	LB- Not available
	FIP-GCCCTCTCGAAACCTAGCTGTTAGTACCCGTTCGGGGATT	
	BIP-TATCTCCGGAACAGGGGAGGAAGCTAACCCAACCGAAGGAA	

Table 1.3. Continued

457 kbp	F3-TTTGCCTAGACGATCTCAA	LF-CGCGACAAGAAGGACGTAAC
	B3-GGGGAAAGGAAACAGAGG	LB-TGCTTTTTAGTGTGGGTGGGG
	FIP-GGTTGCCAATTAATCAACTGCTATTTTTCCCTTCTTTTTTAACACCG	
	BIP-AAGAATAAGGAAGCTAAACATGGGTTGCAGACTGAAATTGACGAA	

Not available: NEB LAMP Primer Design Tool did not yield primers

Table 1.4. List of LAMP primers specific to *Pyrus calleryana.*

Primer ID	Primers	Tm
MYB_F3	Forward outer: 5'- CTGTAAGACCGCCTTGGA AG -3'	55.5 °C
MYB_B3	Backward outer: 5'- CCTCATGGTIGGAACACAT -3'	55.8 °C
MYB_FIP	Forward inner: 5'- GCCACCAGTGTACCAAGCAGCGCCAGACCATCCCCAAAG -3’	71.6 °C
MYB_BIP	Backward inner: 5'- GGCCCCCGGTTCAAAGGATTGCAATGGTGGATTCAGACT -3’	70.1 °C
MYB_LB1	Loop backward 1: 5'- CCCAGAAGAGGAAAACAACACAAAC -3'	56.6 °C
MYB_LB2	Loop backward 2: 5’- TCCCAGAAGAGGAAAACAACACAAA -3'	56.9 °C

Table 1.5. LAMP specificity test using MYB1 primer set. LAMP was run for 30 min at 65 °C. The color of the right four columns correspond to the LAMP result. Yellow color represents positive amplification, pink represents the negative amplification, and the orange color represents the ambiguous result.

Pyrus species	Sample size	LAMP positive '+' result (30 min)	LAMP positive '+' result (40 min)	LAMP negative '-' result (30 min)	Ambiguous result '+/-' (30 min)
Pyrus calleryana	95	75	80	15	5
Pyrus communis	87	0		87	0
Pyrus pyrifolia	90	0		78	12
Pyrus betulifolia	48	4		38	6

Table 1.6. *Pyrus calleryana* detection using LAMP on directly extracted DNA. '+' in yellow indicates positive amplification whereas '-' in pink indicates negative amplification; '+/-' indicates an ambiguous result. The DNA was extracted using tissue samples from a single tree of the *P. calleryana* cultivar 'Cleveland Select' maintained in containers with soilless substrate at the University of Tennessee research nursery.

Plant tissue	Tissue (mg) + buffer (µl)	LAMP result (30 minutes)			LAMP result (25 minutes)			LAMP result (20 minutes)		
		Without dilution	10ng/µl	5ng/µl	Without dilution	10ng/µl	5ng/µl	Without dilution	10ng/µl	5ng/µl
Leaves	20mg+100µl	-			-			-		
	20mg+300µl	-	+	+	-	+	+	-	+	+
	10mg+100µl	-			-			-		
	10mg+200µl	-	+	+	-	+	+	-	+	+
	10mg+300µl	-	+	+	-	+	+	-	+	+
Flowers	20mg+100µl	-			-					
	20mg+300µl	-	+	+	-	+	+	-	+	+
	10mg+100µl	-			-					
	10mg+200µl	-			-					
	10mg+300µl	-	+	+	-	+	+	-	+	+

Table 1.6. Continued

Bark	20mg+100µl	-			-					
	20mg+300µl	-	+	+	-	+	+	-	+	+
	10mg+100µl	-	+	+	-	+	+	-	+	+
	10mg+200µl	-	+	+	-	+	-	-	+	+
	10mg+300µl	+	-	+	-	+	+	-	+	+
Roots	20mg+100µl	+	+	+	-	+	+	-	+	+
	20mg+300µl	+	+	+	+	+	+	+	+	+
	10mg+100µl	+	+	+	+	+	+	+	+	+
	10mg+200µl	+	-	+	+	+	-	-	+	+
	10mg+300µl	+	+	+	+	+	+	+/-	+	+

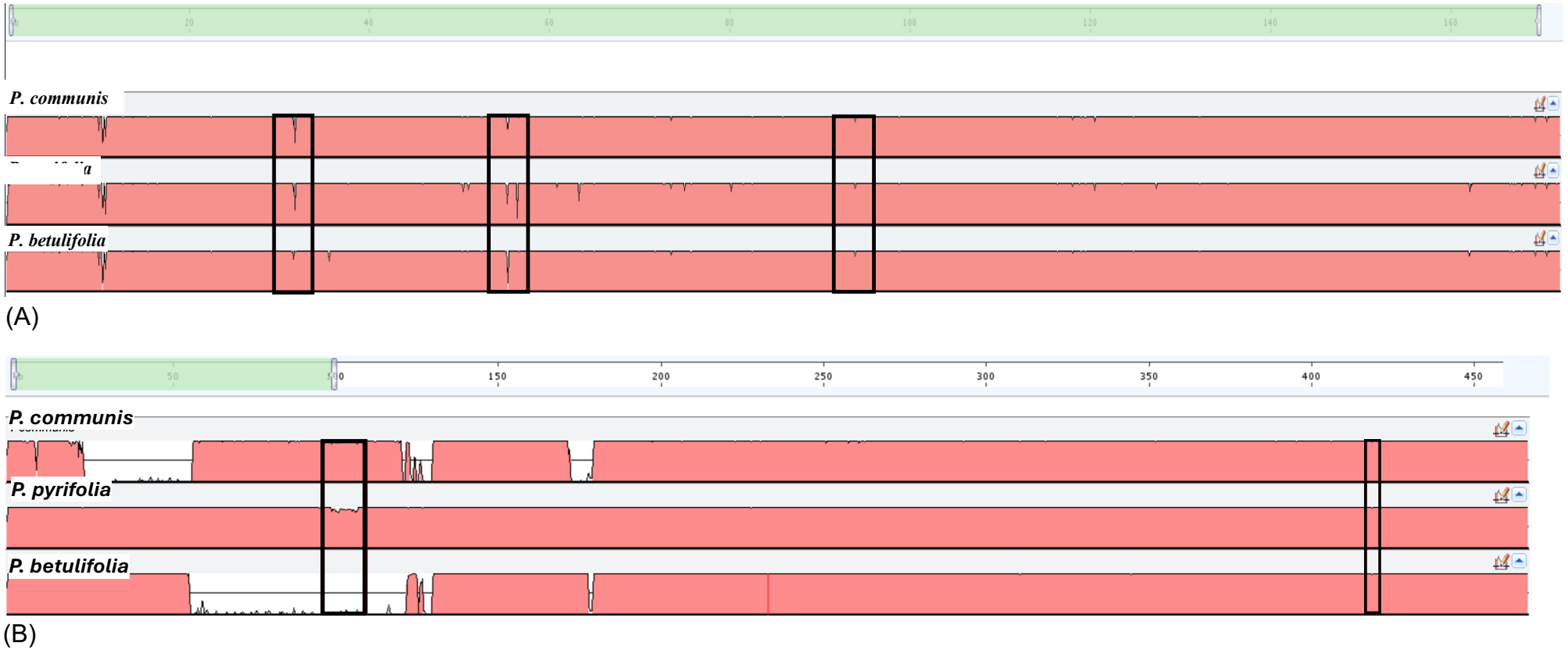


Figure 1.1. mVISTA view with 100 bp sliding window for the comparative genomics of plastidial genomes (A) and mitochondrial genomes (B) of *Pyrus* species with *P. calleryana* as the baseline. Some of the candidate regions of lower conservation or higher dissimilarity with *P. calleryana* are marked with black outlines.

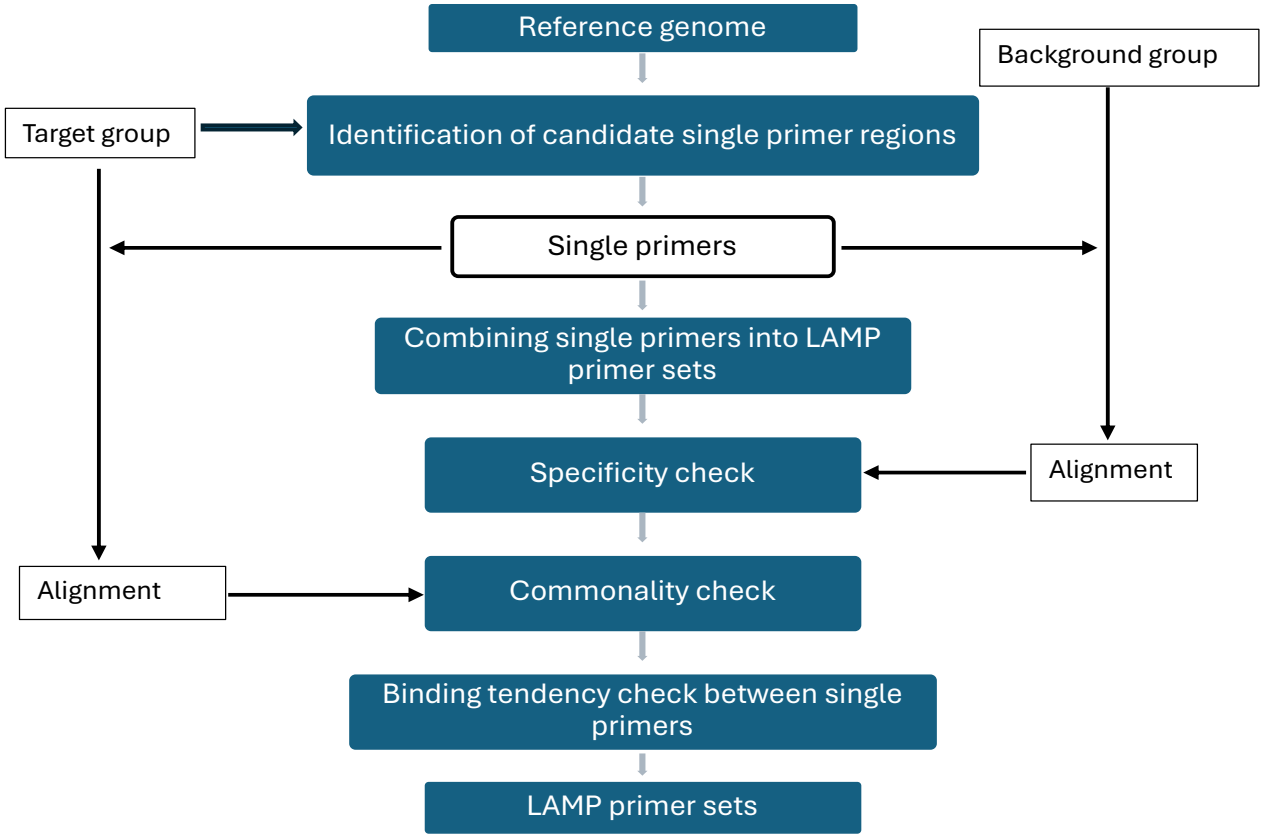


Figure 1.2. Workflow for GLAPD

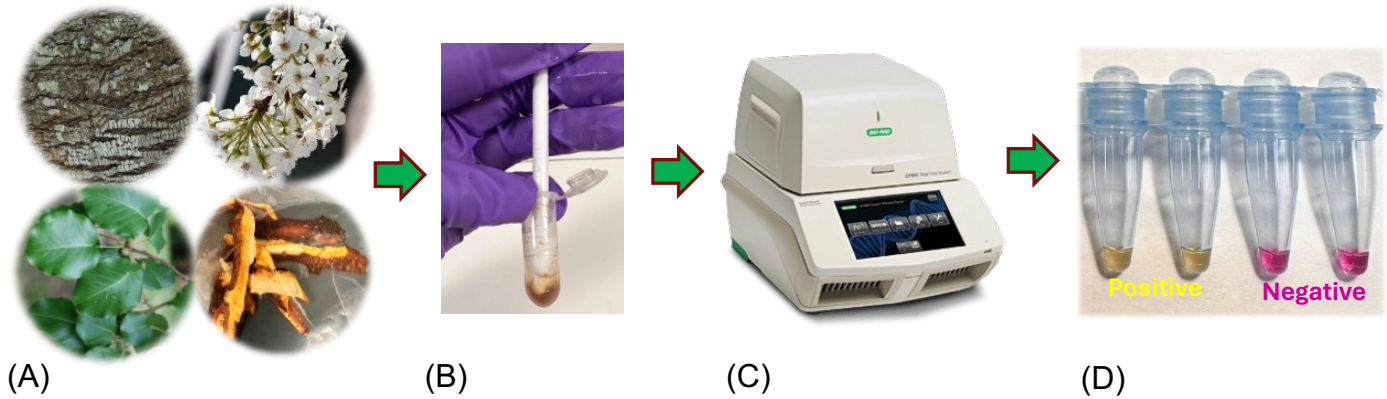


Figure 1.3. Method of direct DNA extraction from fresh tissue samples of *Pyrus calleryana* and the LAMP detection. (A) Different tissue samples used; (B) Tissue samples (leaf, flower, root, bark) were collected into a 2 ml tube containing elution buffer. Samples were then ground in the tube to obtain a solution containing DNA. (C) The directly extracted DNA was then tested using LAMP program in PCR machine. (D) The pink color indicated negative amplification, whereas the yellow color indicated positive amplification of *P. calleryana.*

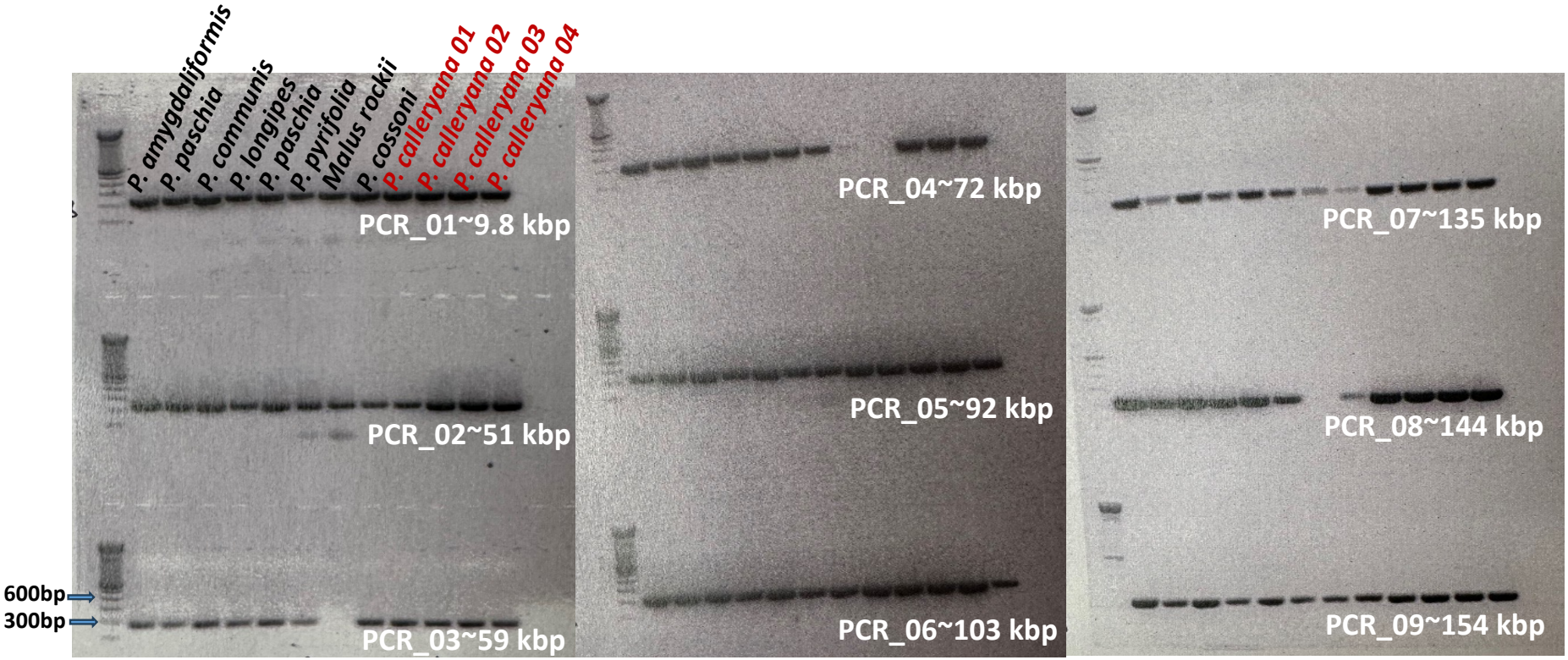


Figure 1.4. Evaluation of plastidial F3B3 primers mentioned in Table 1.1. The expected PCR product size is about 300 bp. The numeric figures followed by PCR number (E.g. PCR_01~9.8 kbp) indicate the approximate position of each primer sets in the plastidial genome of *Pyrus calleryana*. The samples used for testing primers included: *Pyrus amygdaliformis, Pyrus paschia, Pyrus communis, Pyrus longipes, Pyrus paschia, Pyrus pyrifolia, Malus rockii, Pyrus cossoni*, and four samples of *Pyrus calleryana* (left to right). The missing reactions in this figure were successfully amplified in subsequent PCRs.

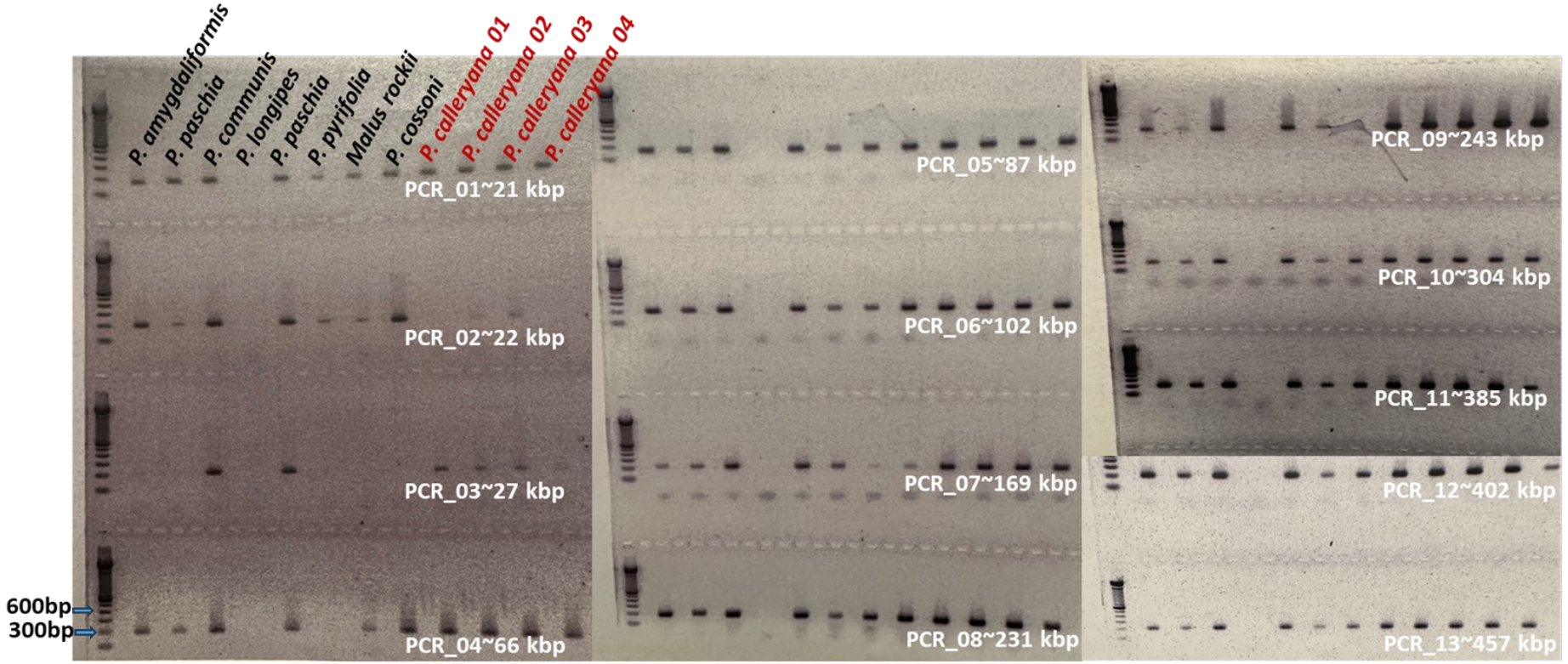


Figure 1.5. Evaluation of mitochondrial F3B3 primers mentioned in Table 1.2. The expected PCR product size is about 300 bp. The numeric figures followed by PCR number (E.g. PCR_01~21 kbp) indicate the approximate position of each primer sets in the mitochondrial genome of *Pyrus calleryana*. The samples used for testing primers included: *Pyrus amygdaliformis, Pyrus paschia, Pyrus communis, Pyrus longipes, Pyrus paschia, Pyrus pyrifolia, Malus rockii, Pyrus cossoni*, and four samples of *Pyrus calleryana* (left to right). The missing reactions in this figure were successfully amplified in subsequent PCRs.

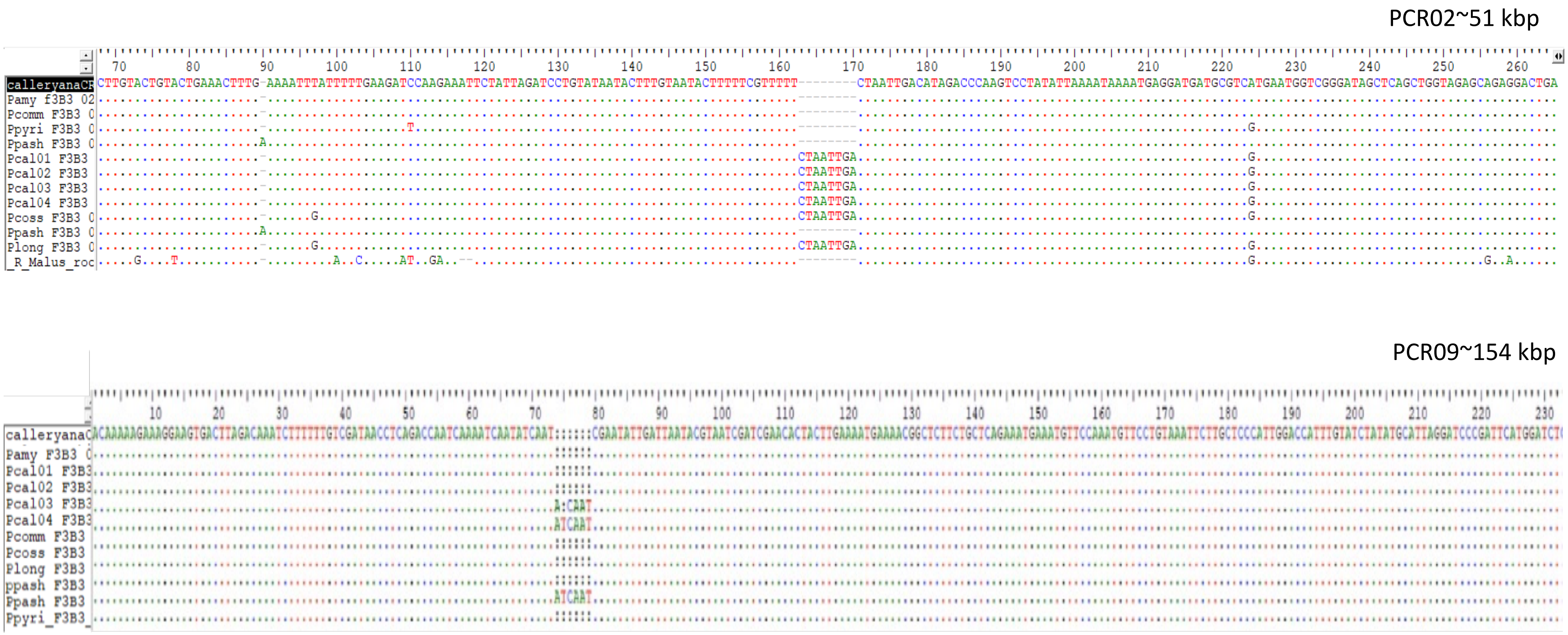
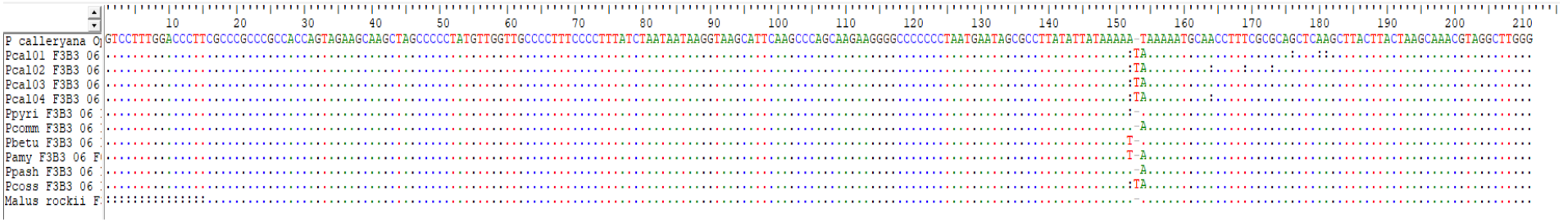


Figure 1.6. Visualization of polymorphism between *Pyrus calleryana* and other *Pyrus* species (PCR02 and PCR09 corresponding to Figure 1.4) by aligning the assembled sequences in BioEdit. Two promising candidate loci from ptDNA selected for downstream analysis.

PCR07~169 kbp



PCR08~231 kbp

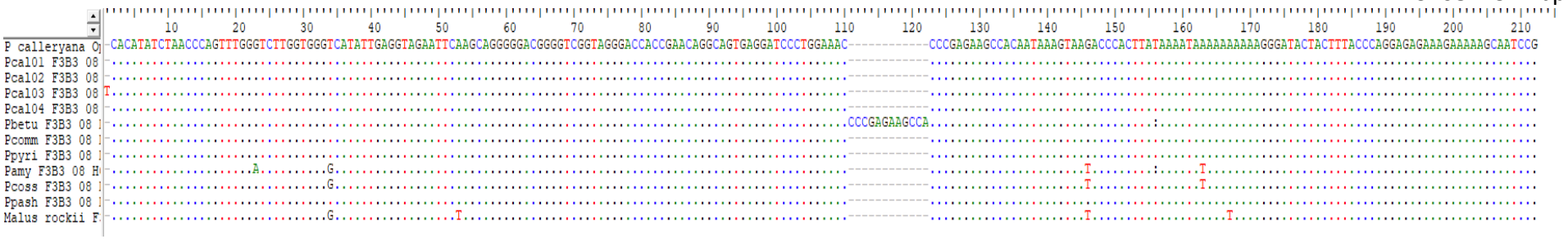
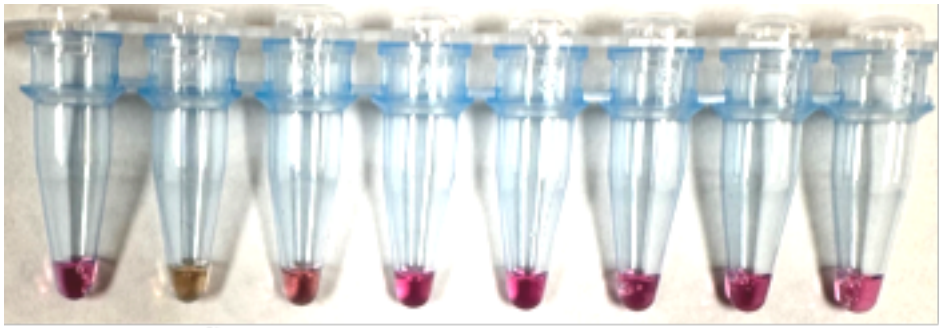


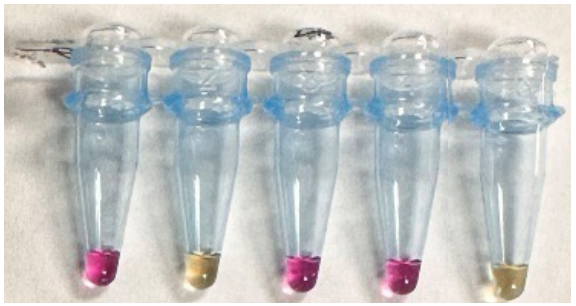
Figure 1.7. Visualization of polymorphism between *Pyrus calleryana* and other *Pyrus* species (PCR07 and PCR08 corresponding to Table 1.3; Figure 1.5) by aligning the assembled sequences in BioEdit. Three promising candidate loci from mtDNA selected for downstream analysis.

(A)



NTC PC

(B)



NTC PC PB PP Pcom

mt21k

mt87k

mt170k

mt385k

(C)

NTC PC1 PC2 PB PP Pcom

Figure 1.8. LAMP results for various primer sets; **yellow color** represents the positive amplification and **pink color** represents the negative amplification. (A) Promising locus (mt169k) under test for differentiating *Pyrus calleryana* (PC) from *P. pyrifolia*. PC positive samples turned yellow at 18 minutes of LAMP reaction. PC: *Pyrus calleryana*; NTC: Non-template control. Non-labeled strips contain *P. pyrifolia* DNA (USDA-GRIN samples). (B) The GLAPD primer set mt26k could differentiate PC DNA from both *P. betulifolia* and *P. pyrifolia*, however, it yielded positive amplification for *P. communis* DNA. (C) Evaluation of NEB LAMP primer sets without the loop primers. The labels mt21k, mt87k, mt170k, and mt385k (Table 1.3; Figure 1.5) indicate the approximate location of the target locus in the mitochondrial genome of PC. PC1: *Pyrus calleryana* 1 (20_PC_054); PC2: *Pyrus calleryana* 2 (Keowee mother-1); PB: *Pyrus betulifolia* (20_PC_084); PP: *Pyrus pyrifolia* (20_PC_086); Pcom: *Pyrus communis* (20_PC_087).

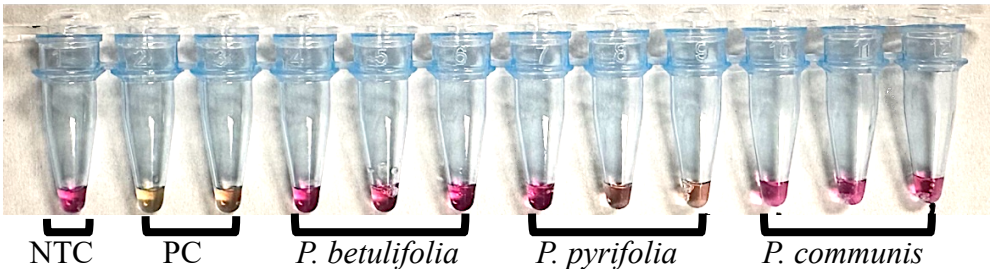


Figure 1.9. LAMP results using MYB LAMP primer sets after incubating at 65 °C for 30 min using 1 µl of 5 ng/µl DNA. Yellow color represents the positive amplification and pink represents the negative amplification. NTC: Non-template control; PC: *Pyrus calleryana.*

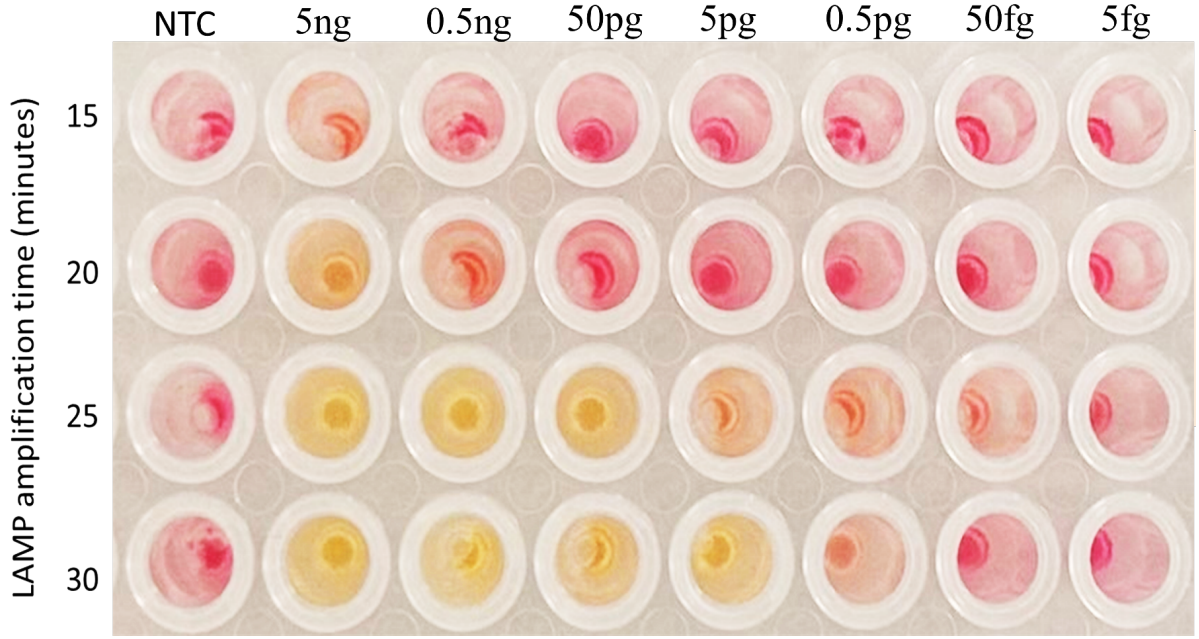


Figure 1.10. PCR test results via LAMP for time vs. serial dilution of the *Pyrus calleryana* DNA using the locus mt169k (Table 1.3; Figure 1.5) having all six LAMP primers. The x-axis represents the different amount of template DNA starting from 5 ng and going down to 5 fg, and y-axis represents the LAMP amplification time in 5 min increments. The pink color indicates no amplification, whereas the yellow color indicates positive amplification. NTC: Non-template control.

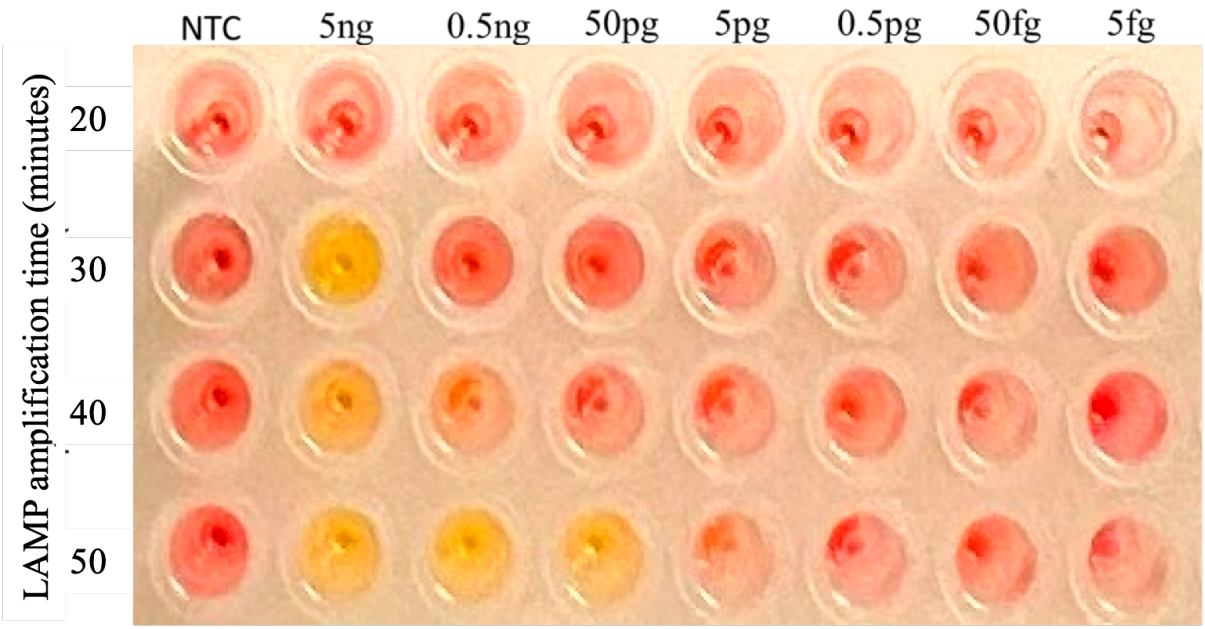


Figure 1.11. PCR test results via LAMP for time vs. serial dilution of the *Pyrus calleryana* DNA using the MYB1 primer set having four core LAMP primers and one loop backward (LB1) primer. The x-axis represents the different amount of template DNA starting from 5 ng and going down to 5 fg, and y-axis represents the LAMP amplification time in 10 min increments. The pink color indicates no amplification, whereas the yellow color indicates positive amplification. NTC: Non-template control.

CHAPTER II
ASSESSMENT OF THE GENETIC DIVERSITY IN INVASIVE CALLERY PEAR POPULATIONS USING MITOCHONDRIAL MICROSATELLITES

Abstract

Callery pear (*Pyrus calleryana* Decne.; PC) is a popular ornamental tree in the urbanized areas of the United States of America (US). The hardiness of PC led to the development and release of several hybrid cultivars making it popular among local gardeners and landscapers. The extensive planting of cultivars in managed areas contributed to the now widespread appearance of invasive PC in almost all habitats in several eastern and southern states of the US. Now, PC is declared an invasive species in most of the eastern US states. Self-incompatibility, tolerance to various environmental conditions, pest resistance, intraspecific hybridization among the cultivars, possibly interspecific hybridization with other escaped *Pyrus* species, and seeds dispersed by various vertebrates could have contributed to the spread and persistence of *P. calleryana* to various environments. As effective management strategies are currently lacking, study of the genetic characteristics of this species may prove useful in generating informed management strategies and decisions. Nuclear genomic short sequence repeats (gSSRs) have been used in previous studies to assess the genetic diversity of PC, however, neither the mitochondrial genome of PC nor the mitochondrial short sequence repeats (mtSSRs) have been developed previously for the genetic diversity assessment. In this study, we developed a mitochondrial genome and five microsatellite markers localized therein, to analyze PC samples collected from the species native range in Asia, samples from Southeastern US escapees (SNesc), Tennessee escapees (TNesc), and US-released commercial cultivars (UScult). All 5 mtSSR markers were used to assess the genetic diversity was assessed using 72 specimens of Asian PC, 69 samples of the UScult, and 180 samples from the SNesc/TNesc collections. Our data revealed high genetic diversity (H_e = 0.842) and presence of genetic structure in PC. The results indicated the possibility of UScult being the intermediary between the Asian populations and the US escaped populations. The observed high genetic diversity among PC cultivars illustrates the intricate genetic landscape shaped by multiple factors, including potential, likely unintended, mislabeling of the US-released cultivars. This study underscores the need for broader genomic studies to further elucidate the genetic architecture of cultivated PC composition and supports the development of robust method for cultivar truthing and invasion management.

Introduction

Callery pear introduction, spread, and impact

Callery pear (*Pyrus calleryana* Decne. [PC]), a flowering pear native to Asia, was introduced to the United States (US) in early 1900s (between 1908 to 1919) primarily as a breeding resource to save the edible European pear *Pyrus communis* which was severely affected by the fire blight disease caused by the bacterium *Erwinia amylovora* (Burril) Winslow et al. (Meyer, 1918; Culley, 2017). PC germplasms collected from Asia were brought to the United States Plant Introduction Stations in Corvalis, Oregon and Glen Dale, Maryland and screened for *E. amylovora* resistance. The screening involved extensive planting of PC seeds followed by inoculating the grown seedlings to assess their susceptibility to the bacterium. The resistant rootstocks of PC were then used widely to propagate other pear species. In 1952, John Creech, a USDA breeder identified the ornamental potential of the PC after he saw a 33-year-old tree with globular growth form, glossy leaves, attractive flowers, and lack of sharp spurs (Creech, 1973). He used this tree as a scion to graft another PC rootstock which resulted in the development of the first ornamental cultivar of PC, the 'Bradford' (Culley, 2017; Whitehouse et al., 1963). The PC cultivar Bradford was sold commercially as an ornamental tree starting in 1962 and soon became one of the most widely planted street trees in the eastern US. With the widespread popularity of Bradford cultivar, at least 25 additional ornamental cultivars of PC have been developed and commercially released since then, from various sources of PC accessions including Asian seed collections and several unknown parents (Culley and Hardiman, 2009; Vincent 2005). Considering that PC is self-incompatible (Culley, 2017; Culley and Hardiman, 2010; Culley and Hardiman, 2007) and rarely produces viable fruits in the native range (Culley and Hardiman, 2007; but see Whitehouse et al., 1963), the invasive potential of its cultivars was initially thought to be limited (Swerington et al., 2002); however, abundant fruit set was documented when different cultivars were planted together. The cross-pollination of PC with either other *Pyrus* species or its own genetically diverse germplasm resulted in the production of viable seeds which are readily spread by birds and wildlife (Griffiths and Huxley, 1992; Vincent, 2005). The escaped populations were not identified until 1964, but soon after that, the escaped populations of PC started appearing and are now found in at least 37 US states (Georgia-Center for Invasive Species and Ecosystem Health, 2023), encroaching the forested areas, roadsides,

riversides, and grasslands, or disturbed areas as early successional species (Culley, 2017; Sapkota et al., 2022). Collectively, the release of new cultivars and the practice of grafting provided the PC trees with sufficient genetic variation to facilitate cross-pollination and form self-sustaining populations (Coyle et al., 2021; Nebhut and Dukes, 2023). Along with the intra- and possible inter-specific hybridization, various traits such as high seed germination rates, a seedling's ability to reach sexual maturity at just a few years old, abundant flowering and fruiting, resistance to pathogens and pests and herbivory, and broad tolerance to various environmental conditions, contribute to the ability of PC to spread into a variety of environments (Culley and Hardiman, 2007; Hardiman and Culley, 2010; Morewood et al., 2004; Warrix et al., 2017). Furthermore, the ability of PC to undergo seed and bud dormancy enables it to produce seeds resilient to harsh temperatures, leading to the formation of a seed reservoir. This reservoir poses an invasion threat for as long as 11 years, with seed germination rates ranging from 45% to 87% (Serota and Culley, 2019; Zhang et al., 2022). Also, the extended leaf phenology in both spring and fall, and frost tolerance that exceeds that of native species, contributes to PC's capacity to outcompete native trees (Maloney et al., 2022). Likewise, the high rates of gene flow, genetic diversity, and high mutation rate enhances the ability of PC to adapt to new environments rapidly (Sapkota et al., 2022).

Pyrus calleryana is an aggressive invader that has naturalized various early successional environments (Vincent, 2005; Woods et al., 2021). This tree species can alter habitats by causing reduction in soil pH and increasing soil C:N (Woods et al., 2021; Woods, Dietsh and McEwan, 2022). Once established, this tree can also reduce native plant species success through allelopathy (Woods, Dietsh and McEwan, 2022; Woods et al., 2023). Not only is PC causing negative impacts on the native ecosystem, but also causing nuisance to people. It has become a hazard to people and property because of its narrow crotch angles that cause branches, specially of old trees, to split when strained by wind, ice, or snow (Culley, 2017; Stewart, 1999). The escaped tree forms dense thickets with thorns, which challenge the native species by creating competition and render it difficult for individuals who attempt removal, due to thickets' impenetrability (Culley, 2017; Higgins, 2018). The woody fruits become soft in winter due to frost and the dropping of soft fruits on the ground causes unsightly litter, as well as poses a risk to the people walking around (Fulcher, 2002). The spring flowers produce a strong rancid odor, variously described as semen, rotting flesh, or chlorine (Lapidos, 2013). Therefore, the invasive

characteristics and ecological impacts of PC underscore the urgent need for management strategies to mitigate its spread and negative effects on both native ecosystems and human environments.

Genetic diversity study of Callery pear

Pyrus calleryana is considered a noxious weed in the eastern US. There are very few documented effective management strategies for PC (Coyle et al., 2021; Gawkins, 2019; Vogt et al., 2020). The most effective control measure for escaped invasive trees is the complete mechanical removal along with the application of glyphosate and triclopyr-based herbicide on cut stumps to prevent re-sprouting (Swearingen et al., 2002), and the same applies for PC (Culley and Hardiman, 2007; Culley et al., 2008; Warrix and Marshall, 2018; Coyle et al., 2021). Once the tree is established in an area, however, its long-lived seeds, frequent thorny phenotype, fire resistance, and capacity to resprout even after top-killing make the control of the established populations cost- and labor-intensive (Coyle et al., 2021; Hay, 2021; Serota and Culley, 2019; Warrix and Marshall, 2018).

The deeper understanding of the genetics of invasive species is crucial to formulate effective management strategies. The genetic and evolutionary mechanisms play a crucial role in determining the successful establishment and spread of invasive species (Barret, 2015). Normally, it is assumed that the invasive populations experience a demographic bottleneck during the time of introduction, which can negatively impact their genetic variation and evolutionary potential (Dlugosch et al., 2008). There are several invasive plant species that are reported to be successful with lower genetic diversity compared to the native populations and there are many invasive plants that show comparably greater genetic diversity than the non-invasive natives (Bhattacharya et al., 2022). By comparing the genetic diversity of populations from native and introduced/invasive ranges, one can evaluate how the stochastic and the deterministic factors influence the extent and distribution of genetic diversity in invasive species (Zhu et al., 2017). The study of population genetic structure and diversity of invasive species helps in elucidating the invasiveness potential and future invasion dynamics, which will subsequently help in their effective control or eradication (2022; Pyšek et al., 2020).

The genomic basis of adaptation and invasiveness has remained largely unexplored, mainly due to the not uncommon poor availability of genomic resources of invasive plant species

(Matheson and McGaughran, 2022). So is the case with PC. Very few studies have been accomplished to evaluate the genetic diversity in PC's native range (Kato et al., 2013; Liu et al., 2012; Sapkota et al., 2021) and there is a limited study of escaped populations in the US (Dunn, 2018; Nebhut et al., 2022; Sapkota et al., 2022). The fine-scale study of PC carried out by Sapkota et al. (2022) indicated high genetic diversity and gene flow within Tennessee and Georgia/South Carolina escaped populations, which underscored the invasion potential of the species.

The microsatellites or short sequence repeats (SSRs) have been used in different population genetics and phylogenetics studies (Hadziabdic et al., 2012; Kamvar et al., 2014; Nowicki et al., 2019). SSRs are characterized by high repetitiveness and mutation rates, extensive polymorphism and reproducibility, and provide useful information for analyzing genetic diversity, genetic relationships and population structures of plants (Shrivastaba et al., 2019; Vieira et al., 2016; Wang et al., 2024). The maternal inheritance, lack of recombination, relatively higher mutation rates, and copy number of mtDNA as compared to nrDNA make mitochondrial short sequence repeats (mtSSRs) suitable for genetic diversity and phylogenetics study (Wang et al., 2014).

Rationale of the study

Previous studies have utilized nuclear genomic short sequence repeats (gSSRs) to assess the genetic diversity of PC, however, neither the mitochondrial genome nor the mtSSRs of PC have been developed previously for use in genetic diversity assessment. Developing the mitochondrial genome of PC provides a valuable addition to the existing genomic resources for this species. Even though PC has been cultivated in the US for more than a century, few studies have assessed the genetic diversity and invasion potential of US commercial cultivars (Hardiman and Culley, 2010; Sapkota et al., 2021; Santamour and McArdle, 1983). By comparing with the published results from nuclear SSRs, our data will provide insights into the genetic variability between native species and the US-released commercial cultivars and wild type/open pollinated escapees. Also, the presence or absence of agreements between nuclear and mitochondrial signals will provide information about the intensity of gene flow or recombination (Sapkota et al., 2022). This study of mtSSRs is also important because the plastidial genome (ptDNA) of PC had no SSRs (Nowicki et al., 2022), so utilizing the highly polymorphic mtSSRs is an easier and faster

way to directly assess and analyze geneflow of PC in comparison to the single nucleotide polymorphisms (SNPs) harboured by nuclear and cytoplasmic genomes. The results will be valuable to compare and contrast the genetic diversity within and between different PC populations using different molecular markers such as gSSRs and mtSSRs. Another valuable application deriving from our results can be a method for cultivar truthing using samples of the confirmed original cultivars.

Hypothesis

We hypothesized that genetic differentiation exists between native and escaped populations and that a high genetic diversity can be observed within and between PC populations across the native, introduced cultivated, and escaped naturalizing populations of PC.

Objectives

Broad: To assess genetic diversity of Asian populations, naturalized US-escapes, and US commercial cultivars of PC using mtSSRs-based genotyping data

Specific:

- **a.** To develop mitochondrial genomic resources of PC
- **b.** To develop and evaluate novel mitochondrial microsatellites (mtSSRs) of PC
- **c.** To assess genetic diversity and population structure of PC populations

Materials and methods

Sample collection

Leaf samples of different *Pyrus* species, including Asian specimens of *Pyrus calleryana* and the original US cultivar selections of PC used by Sapkota et al. (2021) in their study, were re-examined in part, during this study. This was supplemented by samples of four different PC cultivars maintained in UT's bowhouse and leaf samples of confirmed cultivars from Arnold Arboretum. In total, 72 specimens were from Asian collection, 67 from US cultivar selections representing 12 unique cultivars (UScult), and 33 specimens representing 14 different species of *Pyrus,* plus two samples of *Malus rockii* Rehder. Further, the documented accessions of PC and

other relevant species were requested from Germplasm Resources Information Network (GRIN USDA); herbaria and arboreta (Supplementary Table 1). Altogether, 89 specimens of other 16 *Pyrus* species and two specimens of *M. rockii* were also used to evaluate the cross-amplification potential among the developed mtSSRs. Likewise, leaf samples of naturalized escaped trees of PC collected from Tennessee, Georgia, and South Carolina (Sapkota et al., 2022) were also used. The total of 180 samples included 90 from Tennessee (TNesc) and 90 from Georgia/South Carolina (SNesc).

DNA extraction

The genomic DNA of the collected specimens was extracted using approximately 100 mg of dried leaf per sample. The samples were homogenized using a Bead Mill 24 (Fisher Scientific, Pittsburgh, Pennsylvania, USA) and subsequently used for DNA extraction using EZNA DNA DS Mini Kit (Omega Bio-Tek, Georgia, USA), following the manufacturer's protocol. The concentration and purity of extracted DNA were assessed using Nanodrop Spectrophotometer (Thermo Fisher Scientific, Delaware, USA).

Development of mitochondrial genome and its phylogenetic analysis

The draft mtDNA genome of PC was assembled using NOVOwrap version 3.7.2 (Dierckxsens et al., 2016). The available whole-genome paired-end Illumina sequencing data of PC accession from China (Jinshan Fruit Tree Test Station in Shanghai Academy of Agricultural Sciences; GenBank SRR16505594) was used for reference-guided assembly of mitochondrial genome. The mitochondrial genome *P. communis* (GenBank NC_065229.1) was used as a reference to guide the assembly of SRR16505594. NOVOwrap assembly yielded 16 possible genome architectures, however, only six were deemed to be circular or complete; these were further evaluated bioinformatically and confirmed using PCR and Sanger sequencing. To the latter goal, first the six genome assemblies were aligned in mVISTA (Frazer, et al., 2004), with the longest available mtDNA of *Pyrus, P. betulifolia* mtDNA (NCBI: KY563267) used as the baseline reference genome. The alignment allowed for the identification of gaps or discrepancies present among the different putative mtDNA PC assemblies compared to that reference. Those observed gaps in the assemblies indicated potential areas of variation or errors in the assembly. Missing point primers were designed using Integrated DNA Technologies (IDT) (Morrisville, NC) PrimerQuest, based on the identified gaps. Then, the PCR conditions were optimized using touch-down and gradient

PCR and final PCR reactions were performed using the optimized conditions. For smaller expected products, with sizes up to 3 kb, PCR was done in a 10 µl reaction mixture consisting of 5 µl of AccuStart II PCR supermix (2×), 1 µl each of 10 µM forward and reverse primers, 2 µl of sterile water, and 1 µl of 5ng/µl DNA. For expected PCR products of > 3 kb, Platinum SuperFI PCR mastermix was used instead of AccuStart II PCR supermix, otherwise the reaction composition remained the same. The 3 kb PCR program used for temperature optimization was: initial denaturation at 94 ℃ for 2 min, followed by 10 cycles of denaturation at 94 ℃ for 15 s, gradient annealing at 53 - 60 ℃ with a touch-down of 0.7 ℃/cycle for 30 sec and an extension at 72 ℃ for 3 min 30 s, followed again by 30 cycles of denaturation at 94 ℃ for 15 s, gradient annealing at 48 - 55 ℃ for 30 s, and an extension at 72 ℃ for 3 min 30 s, with final extension at 72 ℃ for 1 min. Everything remaining same, extension time was adjusted as per the expected product size for different primer pairs, accordingly. The amplified PCR products were electrophoresed in 2% w/v Agarose Gel stained with Ethidium bromide, with a 100 bp Invitrogen ladder (Thermo Fisher Scientific) for smaller PCR products and Lambda DNA-HindIII Digest (New England Biolabs) as a reference for larger PCR products. The optimized thermal profile for < 3 kb products was: initial denaturation at 94 ℃ for 2 min., followed by 10 cycles of denaturation at 94 ℃ for 15 s, annealing at 56.4 ℃ for 30 s and an extension at 72 ℃ for 3 min. 30 s, followed again by 30 cycles of denaturation at 94 ℃ for 15 s, annealing at 51.5 ℃ for 30 s, and an extension at 72 ℃ for 3 min. 30 s, with final extension at 72 ℃ for 1 min. The solution that exhibited amplification of all investigated gaps in the final PCR was selected as the correct assembly. The selected genome underwent further validation through Sanger sequencing of the amplified PCR products below 3 kb and the mtDNA F3B3 sequence data (Chapter 1). The final draft assembly of mtDNA was then polished using the PILON v.1.23 (Walker et al., 2014) and the SRR16505594 fastq data. Finally, annotation was done using GeSeq v.2.03 from the Chlorobox suite (Tillich et al., 2017).

The phylogenetic analysis of PC mtDNA was based on the alignment with available mtDNA of other *Pyrus* spp., related Rosaceae spp., and the monocot *Oryza sativa* L. to root the tree. The mtDNA sequences identified by their respective GenBank accession numbers (Figure 2.2) were aligned using the default parameters in MAFFT v7 (Kuraku et al. 2013). The resultant sequence matrix was analyzed for the best substitution model (GTR + G), and the tree was visualized in SeaView (Guoy et al., 2010).

Mitochondrial Microsatellite markers

The final mitochondrial genome draft was used to generate only the perfect short sequence repeats (SSRs) using USDA SSR finder v1.00 (Stieneke and Eujayl, 2007). The mitochondrial SSRs with minimum of five repeats and a motif length of 2 to 4 bp were selected. The 19- to-28 bp long SSR primers were designed using PrimerQuest tool of Integrated DNA Technologies (IDT) (Morrisville, NC) with 35 to 50% GC content, 53 to 55 °C melting temperature, and 100 to 500 bp of expected product sizes. PCR was done in a 10 µl reaction mixture consisting of 5 µl of AccuStart II PCR supermix (2×), 1 µl each of 10 µM forward and reverse primers, 2 µl of sterile water, and 1 µl of 2ng/µl DNA. The DNA extracted from *P. calleryana* sample PC_A_054 (Arnold Arboretum, catalog number: 156,099) was used as a positive control and sterile distilled water was used as a negative control. The gradient and touch-down PCR program used for temperature optimization of mtSSR primers was: initial denaturation at 94 ℃ for 2 min., followed by 10 cycles of denaturation at 94 ℃ for 15 s, gradient annealing at 50 - 60 ℃ with a touch-down of 0.5 ℃/cycle 30 sec and an extension at 72 ℃ for 30 s, followed again by 30 cycles of denaturation at 94 ℃ for 15 s, gradient annealing at 45 - 55 ℃ for 30 s, and an extension at 72 ℃ for 30 s, with final extension at 72 ℃ for 1 min. The amplified PCR products were subjected to Agarose Gel Electrophoresis with 2% w/v agarose gel stained with ethidium bromide and a 100 bp Invitrogen ladder (Thermo Fisher Scientific) as a reference. The visualization was done under UV light using UVP GelStudio PLUS (Analytikjena, Upland, California, U.S.) and documented using VisionWorks 8.22.18309.10577. The optimized thermal profile for final PCR was: initial denaturation at 94 ℃ for 2 min., followed by 10 cycles of denaturation at 94 ℃ for 15 s, annealing at 59 ℃ for 30 s with a touch-down of 0.9 ℃/cycle and an extension at 72 ℃ for 30 s, followed again by 30 cycles of denaturation at 94 ℃ for 15 s, annealing at 50 ℃ for 30 s, and an extension at 72 ℃ for 30 s, with final extension at 72 ℃ for 1 min. Finally, after genotyping in this manner the collection of 321 samples of *Pyrus* and outgroup specimens using five mtSSR markers, the PCR products were visualized, sized, and cleaned using QIAxcel ScreenGel software v1.2 (QIAGEN, Germantown, Maryland, US). The allelic sizes were determined using a 15/600 bp alignment marker and 25 to 500 bp DNA size marker (QIAGEN).

Data binning

The raw allelic data obtained from QIAxcel were binned using FlexiBin excel macro (Amos et al., 2007) and used for downstream analysis. The data was transformed into call allelic classes based on the motif lengths for each mtSSRs, respectively. The poppr package version 2.9.6 (Kamvar et al., 2014) in RStudio v4.3.2 (Team, 2013) was used for clone correction at the population level. Exclusively the clone corrected data was used thereafter.

Genetic Diversity

The genetic diversity indices across the five mtSSRs and 4 different populations i.e. Asian, TNesc, SNesc, and UScult were calculated using R studio v4.3.2 with package poppr version 2.9.6. For each mtSSRs, the number of alleles amplified (N), genetic diversity or expected heterozygosity (He), Stoddard and Taylor index (G) (Stoddart & Taylor, 1988), and allelic richness were calculated using R package hierfstat v0.04-22 (Goudet, 2005). For each analyzed population, number of individuals observed (N), number of multi-locus genotypes (MLG) observed, number of expected MLG (eMLG), standard error based on eMLG (SE), Shannon-Wiener Index of MLG diversity (H) (Shannon, 2001) were calculated. Analysis of Molecular Variance (AMOVA) was also performed to evaluate the molecular variance within and between populations using excel macro add-in GenAIEx v6.5 (Peakall and Smouse, 2012).

Population Structure using STRUCTURE and DAPC

STRUCTURE v2.3.4 (Pritchard et al., 2000) was used to analyze the population structure of the genotyped PC collection based on the Bayesian clustering algorithm. Genetic clusters among the PC individuals were inferred using 20 independent Monte Carlo Markov Chains (MCMC) with 500,000 MCMC repetitions that followed 500,000 generations of burn-in period for each used number of inferred clusters (K = 1 to 5). STRUCTURE results were then visualized in POPHELPER structure web app v1.0.10 (Francis, 2017). Evanno's method (Evanno et al., 2005) was used to estimate the optimal number of K and the plot was created for optimal K by merging across 20 MCMCs.

The model free multivariate clustering approach, Discriminant Analysis of Principal Components (DAPC), was also performed using the R package *adegenet* (Jombart, 2008) to analyze the population structure of the PC dataset. The DAPC analysis was cross validated for selecting the optimum number of principal components (PCs) based on cumulative variance. The

analysis was further confirmed using a dendrogram of unrooted neighbor-joining tree of pairwise genetic distances (Nei, 1978) among the studied PC populations. FigTree v1.4.4 (Rambaut, 2018) was used to visualize and edit the phylogenetic tree.

Results

Characteristics of Pyrus calleryana mtDNA

The assembly resulted in a circular genome of 458,892 bp in length with 45.21% GC content (Figure 2.1). This was deemed a complete assembly with a single contig, as per the NOVOwrap program. The mitochondrial genome of PC encompassed 14 core genes, including five ATP synthase genes (*atp1, atp4, atp6, atp8, atp9*), four cytochrome C biogenesis genes (*ccmB, ccmC, ccmFc, ccmFn*), three cytochrome c oxidase genes (*cox1, cox2, cox3*), one ubiquinol cytochrome c reductase gene (*cob*), and one maturase gene (*matR*). There were three ribosomal RNAs (*rrnL, rrnS, rrn5*) detected. Among the 17 transfer RNAs (tRNAs), *trnF-GAA* and *trnM-CAU* had three copies, *trnS-UGA*, *trnK-UUU*, *trnH-GUG*, *trnG-GCC, trnC-GCA*, and *trnY-GUA* had two copies, and the remaining ones had a single copy each. The genome also had nine NADH dehydrogenase genes (*nad1, nad2, nad3, nad4, nad4L, nad5, nad6, nad7, nad9*), two large ribosome protein subunits (*rpl5, rpl10*), five small ribosome protein subunits (*rps1, rps12, rps13, rps3, rps14*), one transport membrane protein (*mttB*), and one succinate dehydrogenase gene (sdh4). The mtDNA of PC appeared to be typical in terms of gene content compared to other *Pyrus* species, and when placed among the other genomes of the related Rosaceae species, clustered closely with other *Pyrus* spp. mtDNA (Figure 2.2).

mtSSRs and cross-amplification

The USDA SSR finder generated five mtSSRs (Figure 2.1; Table 2.1) from the polished mtDNA assembly of PC. The SSRs included di- and trinucleotides with six or seven repeats and total repeats length ranging from 12 to 21 bp. Cross-amplification test was performed using four mtSSRs and all of the mtSSRs showed 100% amplification of all the tested *Pyrus* species as well as the outgroup *Malus rockii* (Supplementary Table 2).

Genetic diversity

After performing clone correction at the population level, 190 individuals represented unique multi-locus genotype (MLGs) out of total 321 individuals that were genotyped. The number of individuals reduced drastically in each population, highest reduction being in UScult after clone correction, indicating the presence of clonal replicates which could inflate the measures of genetic diversity. As such, the genotypic dataset representative of those 190 individuals was used for remainder of the analyses. There was 2.32% missing data throughout the dataset. The locus 417k and the Asian population had the highest portion missing data of 3.68% and 33.85%, respectively. Other populations had no missing data. The average Shannon-Wiener Index of MLG diversity (H) among four studied populations was 5.15, with the Asian population having relatively high value of 4.17, which indicated overall high genetic diversity in the genotypic dataset (Table 2.2). Similarly, the overall allelic richness (A_r) of 5.83 was observed, with UScult having the highest value of 7.49 and TNesc the lowest of 3.80. A total of 149 private alleles were found in the dataset, with Asian population having the majority of private alleles (n = 123), whereas the TNesc had no private alleles. The individual inbreeding coefficient (F_{IS}) was negative for all populations as was expected for haploid mtSSRs.

All five mtSSRs were needed to capture all of the MLGs present (Figure 2.3), which indicates that each marker carries essential and unique genetic information necessary for distinguishing between the genotypes. This also highlights the complexity and diversity of the PC populations.

An average of about 11 alleles per locus ranging from 7 in 67k to 20 in 353k were detected (Table 2.3). The allelic richness (A_r) across the loci ranged from 4.74 in 67k to 8.26 in 353k with an average of 5.82 which again suggests the high genetic diversity and evolutionary potential of PC populations.

Analysis of Molecular Variance (AMOVA) indicated the 17% of the genetic variation is among populations and 83% is within populations, suggesting high genetic differentiation among different individuals within a population (Table 2.4). ΦPT value of 0.167 with P<=0.001 support suggests the presence of genetic differentiation or population structure among the populations.

Population Structure

The Bayesian clustering of populations using STRUCTURE indicated two genetically distinct clusters (ΔK = 2) by Evanno's method in the genotyped PC dataset (Figure 2.4a). Cluster one mainly encompassed Asian and TNesc populations, whereas cluster two comprised of SNesc and UScult populations (Figure 2.4b). Additionally, three genetically distinct clusters of the genotyped PC also showed SNesc and UScult in the same cluster (Figure 2.4c). There was admixture in all four populations, with relatively more gene flow and admixture in SNesc and UScult populations.

The multivariate model-free analysis, DAPC, showed the clustering of PC dataset similar to STRUCTURE, however, the Asian population represented a complete separate cluster without admixture with the other inferred genetic cluster (Figure 2.5). DAPC indicated two clusters for the given *P. calleryana* dataset, with Asian populations in one cluster and all other populations in the second cluster. The DAPC result indicated that all other populations were diverged from Asian populations which was also supported by the genetic distance dendrogram (Figure 2.5; upper right). These results indicated the possibility of divergence of US commercial cultivars (UScult) being the intermediary between Asian population and US escaped populations (TNesc and SNesc).

Discussion

We developed the novel mitochondrial genomic resource of PC and investigated the genetic diversity and population structure of Asian populations, naturalized US-escapes (TNesc and SNesc), and US commercial cultivar collections using the newly developed mtSSRs. Our comprehensive analysis yielded several important insights into the genetic structure and diversity of PC populations, thus amending our previous investigations of the same plant collection using gSSRs (Sapkota et al., 2021; Sapkota et al., 2022).

The mtSSRs proved to be effective in capturing genetic variation within and between populations. The designed primers exhibited amplification success across different *Pyrus* species, which demonstrates their utility in cross-species amplification. This feature is particularly valuable for studying genetic relationships and diversity in the broader *Pyrus* genus. Our analysis of genetic diversity revealed overall high Shannon-Wiener Index of MLG diversity and Nei's unbiased gene diversity of PC population. This result is similar to the results obtained

from genetic diversity study of PC using nuclear SSRs (Liu et al., 2012; Kato et al., 2013; Sapkota et al., 2021; Sapkota et al., 2022). We observed average allelic richness in our datasets as compared to other related species such as *Malus* and *Prunus.* In a genetic diversity study of *Malus domesticus* and *M. sylvestris* using nuclear microsatellite markers (Bitz et al., 2019; Schnitzler et al., 2014), A_r was reported as ranging from 2 to 3.09. Similarly, a study by Bourguiba et al., 2020 in *Prunus* revealed A_r ranging from 8 to 13.

The high genetic diversity and private alleles in the Asian population of PC suggests a broad base for the genetic diversity with unique genetic variants in the species native range. The study of gSSRs identified the highest genetic diversity in Asian populations (Sapkota et al., 2021), whereas mtSSRs showed the highest genetic diversity in UScult. The highest value of He in the UScult collection may be due to the use of genetic material from a broad geographic range (Sapkota et al., 2021; Vincent 2005) and distinct gene pools and possibly outcrossing with other *Pyrus* species (Stewart, 1999; Vincent 2005). This practice can introduce diverse maternal lineages from various sources into the breeding pool, resulting in relatively high mitochondrial genetic diversity. The inter- and intraspecific hybridization and development of new cultivars are directly linked to increase in the genetic diversity and evolution of invasiveness (Culley and Hardiman, 2008; Culley at al., 2011). The absence of private alleles in the TNesc population and only two private alleles in SNesc population suggests a high gene flow among escaped PC individuals. The result signifies the ability of PC to produce viable seeds by cross pollination which are then dispersed by birds and humans to nearby locations. Our results of high gene flow among escaped PC populations are consistent with previous studies on escaped population using gSSRs (Sapkota et al., 2022). In a self-incompatible species, for example, PC, high gene flow helps in the exchange of genetic material and results in abundant fruit set (Dunphy and Hamrick, 2005). Similarly, the expected negative value of F_{IS} suggests random mating or outbreeding in PC populations, which is a common phenomenon in pears, again resulting in high level of genetic diversity and contributing to their evolutionary potential and adaptability. However, in other studies, a positive F_{IS} was observed in PC indicating the common ancestors of different populations (Liu et al., 2012; Sapkota et al., 2021; Sapkota et al., 2022).

The presence of 41 multi-locus genotypes (MLGs) in 41 individuals comprising of 12 different cultivars in our dataset challenges the expectation that the plants grafted from the same rootstock should be genetically identical. The presence of these unique MLGs may be linked to

genetic variation in scions used, and hybridization events, yet it also is possible that they may reflect specimen mislabeling or record-keeping errors from original collecting efforts. Our results align with the previous records of cultivar mishandling in PC (Culley and Hardiman, 2007; Santamour et al., 1980; Sapkota et al., 2021). That each individual tree possesses a unique MLG suggests a high level of genetic diversity and their invasive potential within the sampled individuals of US commercial cultivars. But the potential biases introduced by the limited resolution of QIAxcel Capillary Electrophoresis system and data binning process could have affected the accuracy of allele detection to a certain extent, which would have but a miniscule chance to modify the overall conclusions stemming from several independently analyzed PC cultivars.

The results from AMOVA showed high within-population diversity, which highlighted the genetic differentiation and structured population. These findings align with other studies on PC, where high genetic variability within populations has been documented (Liu et al., 2012; Sapkota et al., 2021). The differentiation of populations into two genetic clusters by STRUCTURE indicates historical and contemporary gene flow patterns with high admixture in escaped populations. The admixture observed, especially in the SNesc and UScult populations, suggests ongoing hybridization and gene flow between escaped and cultivated individuals. The DAPC further supported the population structure revealed by STRUCTURE, with added historical context, i.e., the Asian population formed a distinct cluster without admixture. This separation might indicate that the Asian populations have retained their unique genetic identity, likely due to geographical and reproductive isolation. Considering an alternative scenario, the Asian populations represent the original genetic pool and the subsequent bottleneck events led to the genetic structure of US populations. Our findings highlight the complexity and genetic richness within PC populations, which is essential for their conservation in at least a part of their native range where they are considered threatened (Kato et al., 2013) and management in the US escaped areas.

The genetic distance dendrogram and DAPC suggest the possibility of UScult being the intermediary between the Asian populations and the US escaped populations. This evolutionary pathway may reflect the introduction history of PC in the US, where initial introductions were followed by cultivation and escapes and naturalization events, leading to distinct but related

genetic clusters. This result is in contrast with the findings of previous study where SNesc populations appeared to have first diverged from the Asian population (Sapkota et al., 2022).
In conclusion, we observed high genetic diversity of PC. The observed genetic diversity among PC cultivars illustrates the intricate genetic landscape shaped by multiple factors, including potential, highly likely unintended, mislabeling of the US-released cultivars. Mislabeling could have led to a mix of genetically distinct plants being grouped under the same cultivar names. This mislabeling could occur at various stages, from initial breeding and propagation *en masse* to commercial distribution and planting. This study underscores the need for broader genomic studies to further elucidate the genetic architecture of cultivated PC composition and supports the development of robust method for cultivar truthing and invasion management. Despite the valuable insights gained from mtSSRs, they may cover only a small fraction of the genome and may not capture all the relevant genetic variations. Future research should aim to employ high-throughput sequencing technologies to provide a more comprehensive and detailed understanding of genetic diversity and to help formulate effective management strategies for the escaped populations. Specifically, a population genomics study can help identify the genetic markers associated with the invasive traits of PC, which in turn will help prioritize management efforts, develop molecular detection tools, and control the spread of PC.

APPENDIX: Tables and figures

Table 2.1. List of Short Sequence Repeats (SSRs) generated by USDA SSR finder and their respective primers generated by PrimerQuest tool.

SSR ID	Motifs	Start position in genome	End position in genome	Repeat total (bp)	Forward (F) and Reverse (R) Primers	Tm of primers	Expected PCR product size
					F: GTATATGTGGCACAGAGGCTAC	55 ℃	200 bp
PC_ssr68001	$(at)_6$	68001	68013	12	R: TTTCATACTGTCGTATTT	53.9 ℃	
					F: GCAGGTATCGTATGACCCAAA	54.6 ℃	190 bp
PC_ssr150549	$(tat)_7$	150549	150570	21	R: TGCCAGTAGCTTCGTAGAATG	54.5 ℃	
					F: GGCCTGTAGAGCACATGTAA	54.5 ℃	180 bp
PC_ssr333964	$(tc)_6$	333964	333976	12	R: CGGCTAGGATAATTGGGAAAGA	54.3 ℃	
					F: GAGTAGTGAAGAATTATAGTGAGTTGTG	53.3 ℃	290 bp
PC_ssr353226	$(tc)_6$	353226	353238	12	R: GCCTTACCGGAGAGTAAGAAT	54.3 ℃	
					F: TTGCCTTCGCTTCCTCTTT	54.6 ℃	165 bp
PC_ssr417915	$(tc)_7$	417915	417929	14	R: TCTCTCCCATCAGGGTCATT	55.1 ℃	

Tm calculated by IDT

Table 2.2. Genetic diversity indices observed among different *Pyrus calleryana* populations.

Population	Before CC	After CC	MLG	eMLG	% Missing	No. of Alleles	H	G	F_{IS}	A_r	λ	H_e	$\overline{r}_d$	E.5	P_a
Asian	72	65	65	36	33.85	31	4.17	65	–0.4709	5.45	0.985	0.663	0.097***	1	123
TNesc	90	36	36	36	0	19	3.58	36	–0.4945	3.80	0.972	0.629	0.033**	1	0
SNesc	90	48	48	36	0	34	3.87	48	–0.4223	6.59	0.979	0.773	0.271***	1	2
UScult	67	41	41	36	0	38	3.71	41	–0.3655	7.49	0.976	0.848	0.329***	1	24
Summary statistics	***321***	***190***	***177***	***35.5***	***11.57***		***5.15***	***164***	***–0.4321***	***5.83***	***0.994***	***0.842***	***0.163******	***0.95***	***149***

Before CC: Number of individual samples before clone correction; After CC: Number of individual samples after clone correction; MLG: Multi-locus genotype; eMLG: expected multi-locus genotype; % Missing: percent of genotypes missing i.e. those genotypes that failed to amplify; No. of Alleles: Number of alleles; H: Shannon-Wiener Index of MLG diversity (Shannon, 2001); G: Stoddart and Taylor's Index of MLG diversity (Stoddart & Taylor, 1988); F_{IS}: Individual inbreeding coefficient; A_r: Allelic richness; λ: Simpson's Index (Simpson, 1949); H_e: Nei's unbiased gene diversity (Nei, 1978); $\overline{r}_d$: Standardized index of association taking into account for the number of loci sampled (Kamvar et al., 2014); E.5 : Allelic richness; Pa: Number of private alleles in each population. Significance was assessed by using 1000 permutations among all individuals, *** = $P < 0.001$ and ** = $P > 0.01$.

Table 2.3. Genetic diversity indices of *P. calleryana* based on microsatellite loci.

mtSSR Locus	No. of Alleles	% Missing	A_r	H_e	D_{st}	D_{est}	G	F_{ST}	F_{IS}
67k	7	1.58	4.74	0.80	0.27	0.11	4.88	0.039	–0.483
150k	11	1.58	4.75	0.85	0.05	0.24	6.37	0.078	–0.472
333k	10	1.58	6.34	0.83	0.02	0.15	5.71	0.036	–0.328
353k	20	3.16	8.26	0.93	0.11	0.51	13.39	0.134	–0.405
417k	8	3.68	5.05	0.81	0.03	0.12	5.02	0.043	–0.484
Average	***11.2***	***2.32***	***5.82***	***0.84***	***0.09***		***7.07***		

No. of Alleles: Number of alleles detected; % Missing: % of samples that failed to amplify in the given locus; A_r: Allelic richness; H_o: Observed heterozygosity (Frequency of heterozygous individuals per locus averaged over the number of sampled loci); H_e: Expected heterozygosity (Nei's unbiased gene diversity; Nei 1978); Dest: Jost's differentiation estimate (Jost, 2008); G: Stoddard and Taylor index (Stoddard & Taylor, 1988); F_{ST}: a measure of sub-population genetic structure; F_{IS}: a measure of deviations from Hardy-Weinberg equilibrium in terms of heterozygote deficiency if <0 or homozygote excess if >0.

Table 2.4. Analysis of Molecular Variance (AMOVA) for the *Pyrus calleryana* dataset.

Source of variation	Degree of freedom	Sum of squares	Mean squares	Est. variance	% variance	ΦPT
Among Populations	3	57.520	19.173	0.371	17%	
Within Populations	186	343.575	1.847	1.847	83%	
Total/Average	***189***	***401.095***		***2.219***	***100%***	***0.167****

Degree of freedom (df) given by no. of samples - 1; Sum of squares: sum of squares of the deviations of all the observations from their mean; Mean squares: mean squares is the sample variance obtained by dividing the sum of squares by the respective df; Est. variance.: estimated variance; % Variance: percent of the total variance for each hierarchical level; PhiPT: partitioning of total genetic variation is the statistics calculated by the test; Significance was assessed by using 1,000 permutations of the dataset, ** = $P < 0.001$.

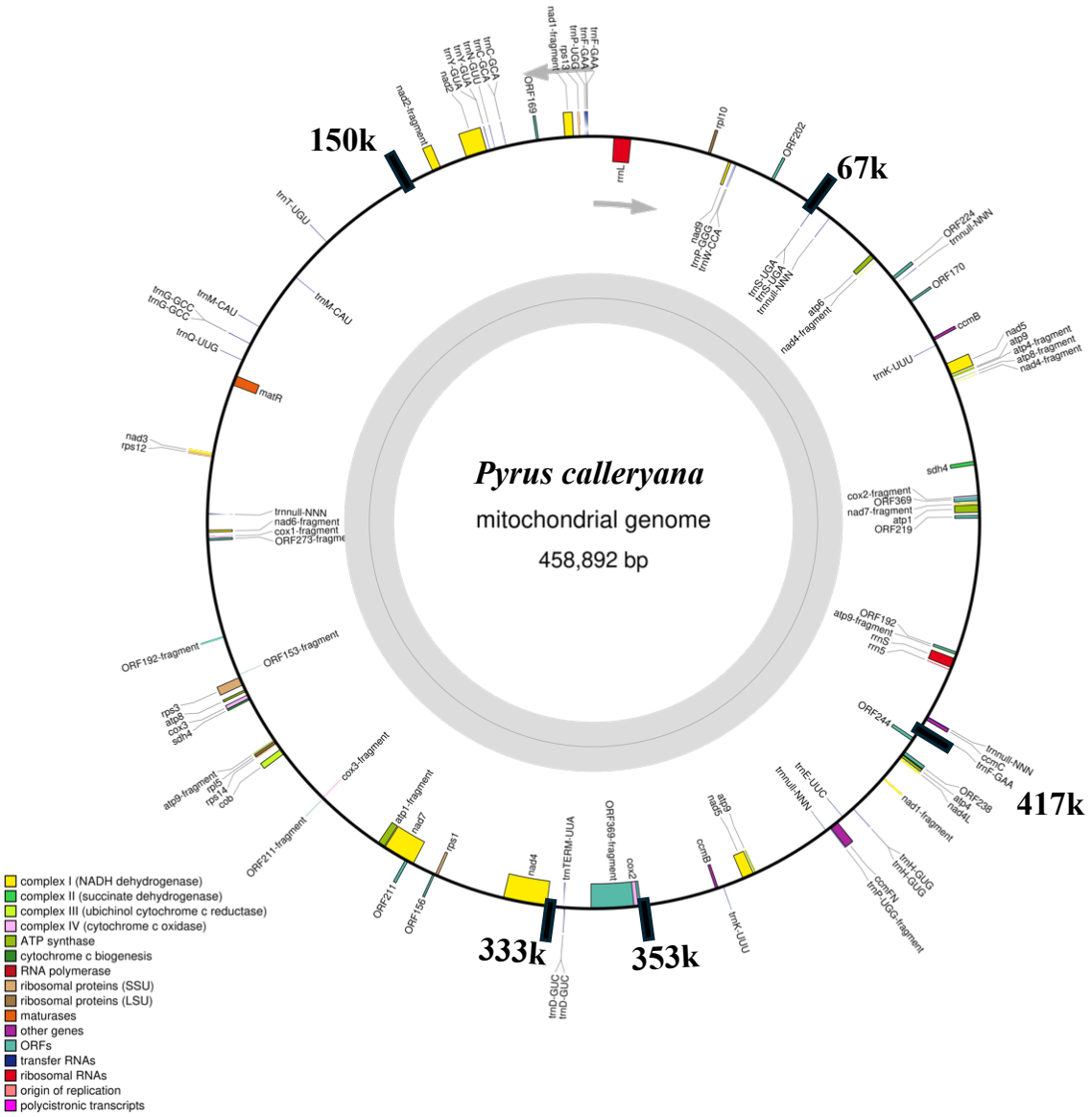


Figure 2.1. The circular map of the complete mitochondrial genome of *P. calleryana*. The charcoal grey colored region in the inner circle represents the GC content (45.21% overall) in a sliding window manner using the default settings of Chlorobox (Tillich et al., 2017). Different functional gene groups are color coded and explained in the legend (bottom-left). The approximate positions of the developed five mtSSRs: 67k, 150k, 333k, 353k, 417k are marked by black rectangles. The arrows represent the direction of transcripts.

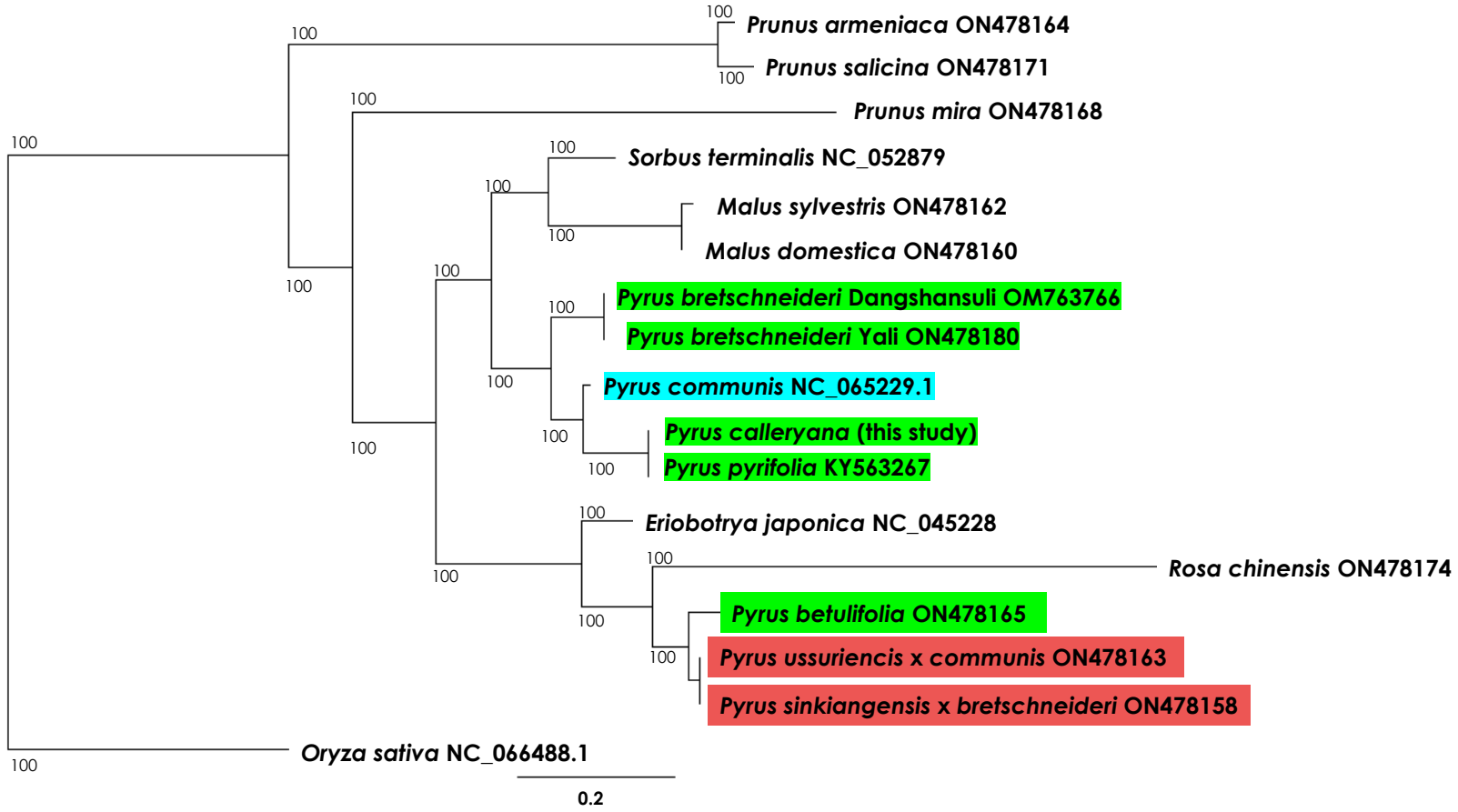


Figure 2.2. Phylogenetic placement of the newly developed mitochondrial genome of *Pyrus calleryana*. All mitochondrial genomes used for analysis are labeled with their respective GenBank IDs. MAFFT-aligned genomes were assessed for the nucleotide substitution model (GTR + G) and bootstrapped 100 times. The tree was re-rooted using the single monocot species, *Oryza sativa*. For reference, the genetic distance legend is placed at the bottom. Colored nameplates denote the geographic origins of given *Pyrus* species: East Asia — green; Europe — blue; *Pyrus* cultivars — dark red.

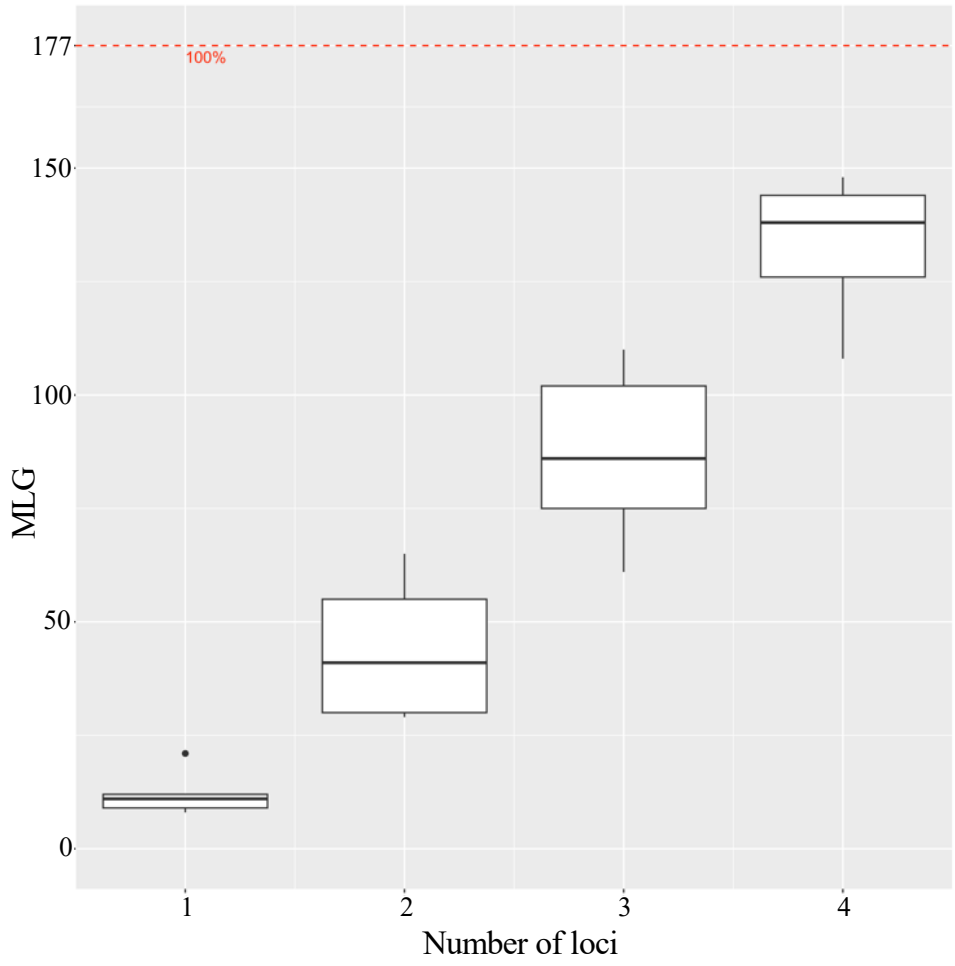


Figure 2.3. Genotype Accumulation Curve (GAC) for populations included within the *P. calleryana* dataset. It represents the number of MLG detected (Y-axis) in relation to the number of loci (X-axis) used for genotyping. The boxes represent the interquartile ranges of the number of unique genotypes detected.

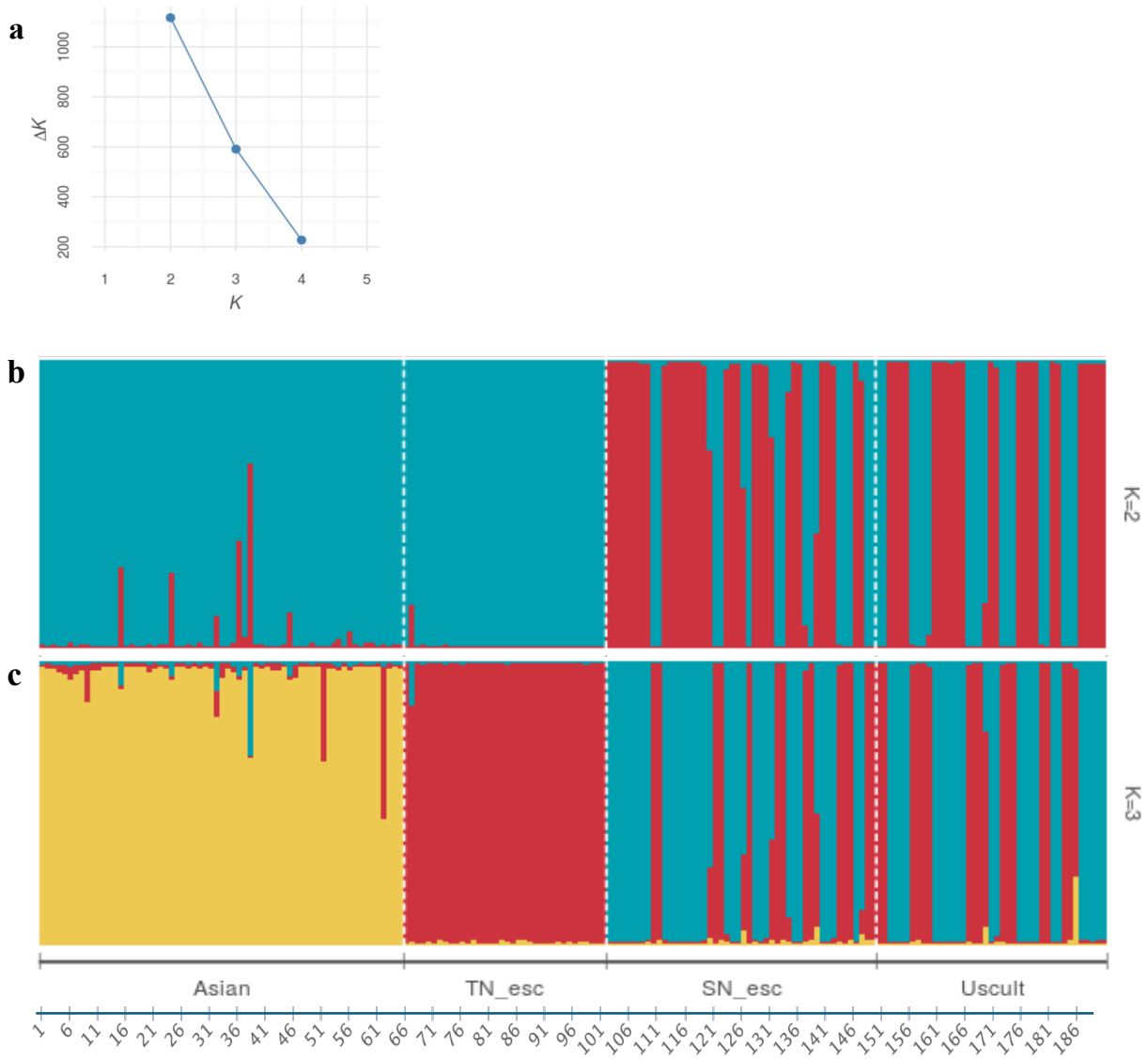


Figure 2.4. Bayesian clustering using STRUCTURE for *P. calleryana* populations. STRUCTURE Bayesian clustering analyzed with (a) the Evanno method and visualized using (b) 2 genetic clusters (K=2) and (c) 3 genetic clusters (K=3). Each vertical bar represents an individual sample, and the blue and red bar color indicates the probability of an individual to get assigned to one of the clusters identified. ΔK is a metric used to identify the optimal K value that best explains the genetic structure observed in the data.

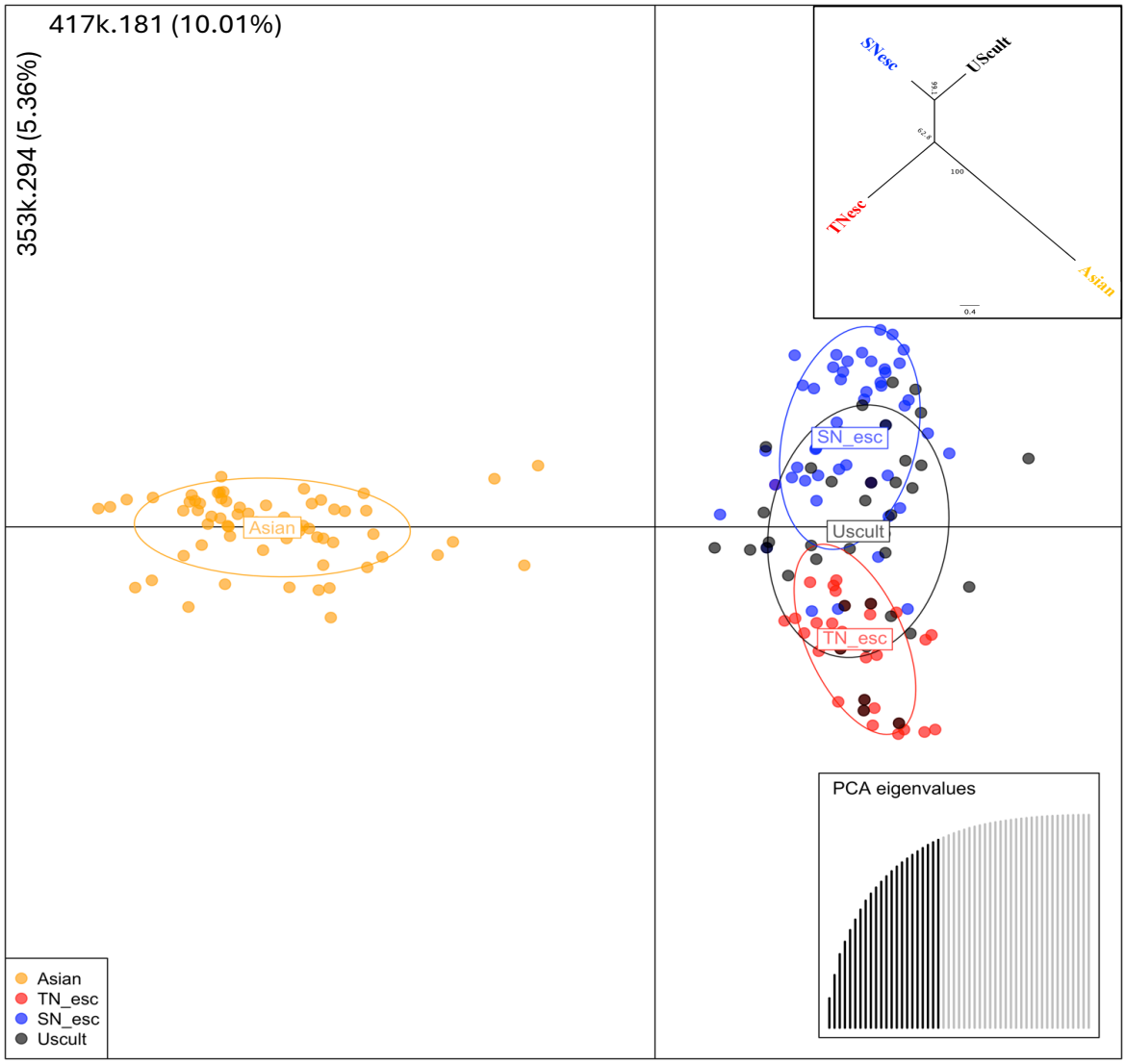


Figure 2.5. Discriminant Analysis of Principal Components (DAPC) of *P. calleryana* dataset. The four populations are represented by four different colors (Orange: Asian, Red: TN_esc, Blue: SN_esc and Black: Uscult). The 22 PCAs, as shown in the bottom right insert, that covered 80% of the cumulative variance were selected to visualize the dataset. The two alleles (353k.291 and 417k.181) that explained relatively most of the variance within the populations are indicated along the x and y axes. The top right insert represents the genetic distance tree: Unrooted neighbor-joining tree of pairwise genetic distances (Nei) among the sampled *P. calleryana* datasets.

CHAPTER III
FUTURE DIRECTIONS

Molecular Detection of PC

Field validation and optimization

Extensive field trials must be done to validate the LAMP protocol under diverse locations. This will involve testing the assay in various locations and seasons to ensure its reliability and robustness in real-world applications. Field validations will help identify any potential limitations and if any adjustments are required for the protocol, such as optimizing sample preparation methods, ensuring the stability of reagents used, and refining storage conditions.

Expansion of LAMP tool to other invasive species

LAMP assay for detecting any invasive plant species have not been documented in the literatures. As such, expanding the LAMP protocol to detect other invasive species presents a valuable opportunity to broaden its utility and contribute to invasive species management. This process involves adapting and validating the existing LAMP assay to identify the genetic markers unique to other invasive plants. Developing multiplex LAMP assay by identifying and including multiple target sequences from different species or genetic variants in the assay could detect multiple invasive plants, providing a rapid and comprehensive assessment of biodiversity and ecosystem health.

Genetic Diversity study of PC

Whole genome sequencing and Transcriptome analysis

Unlike mtSSRs, which cover only a fraction of the genome, whole genome sequencing (WGS) aims to capture the entire genome, providing a comprehensive view if genetic diversity and identifying specific genetic variations associated with invasive traits. By comparing the genomes of invasive and non-invasive native populations of PC, genetic markers correlated with invasive traits such as rapid growth, reproductive viability, resistance to biotic and abiotic stresses could be identified. Also, WGS can provide information on the evolutionary history and adaptive mechanism of PC helping to understand how certain configurations contribute to invasiveness. Similarly, RNA-seq can be done to see which genes are upregulated or downregulated in response to specific environmental stimuli, helping to identify key regulatory pathways involved in invasiveness. Understanding PC's response to biotic and abiotic stresses can inform strategies

to manage its spread. Insights from transcriptome analysis can guide targeted gene editing approaches to develop less invasive cultivars.

Population genomics study and SNP genotyping assays

Population genomics could be utilized to identify SNPs linked to invasive traits and SNP genotyping assays could be developed for rapid and accurate detection of PC cultivars. The tool could aid in cultivar truthing, ensuring accurate labeling and tracking of genetic lines. The population genomics study could also inform the extent of gene flow and hybridization.

Entanglement analysis for comparing nuclear SSRs and mtSSRs

Entanglement analysis will inform how genetic information from different sources interacts or overlaps. It helps to unravel the complexity in genetic data by identifying correlations and differences between the markers that could be due to evolutionary processes, gene flow, or other factors. High correlations might indicate similar evolutionary pressures or gene flow patterns affecting both genomes. Finally, invasion route could be traced.

REFERENCES

[EDDMapS] Early Detection and Distribution Mapping System (2024). Callery pear (Bradford pear). University of Georgia–Center for Invasive Species and Ecosystem Health. http://www.eddmaps.org. Accessed: June 24, 2024.

Barrett S. C. (2015). Foundations of invasion genetics: the Baker and Stebbins legacy. *Molecular ecology*, *24*(9), 1927–1941. https://doi.org/10.1111/mec.13014.

Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology letters*, *12*(2), 20150623. https://doi.org/10.1098/rsbl.2015.0623.

Bhattacharya, S., Hernández, F., Ferreira, M.A., Machado, R.M., Sun, Y., Wang, M., Zhang, C., Hao, J. (2022). Genetic diversity and population structure of invasive and native populations of *Erigeron canadensis* L., *Journal of Plant Ecology*, 15(4): 864–876, https://doi.org/10.1093/jpe/rtac016

Bi, A. , Nakajima, C. , Fukushima, Y. , Tamaru, A. , Sugawara, I. , Kimura, A. , Kawahara, R. , Hu, Z. *et al* (2012) A rapid loop-mediated isothermal amplification assay targeting *hspX* for the detection of *Mycobacterium tuberculosis* complex. *Jpn J Infect Dis* 65, 247–251.

Bi, C., Paterson, A. H., Wang, X., Xu, Y., Wu, D., Qu, Y., Jiang, A., Ye, Q., & Ye, N. (2016). Analysis of the Complete Mitochondrial Genome Sequence of the Diploid Cotton *Gossypium raimondii* by Comparative Genomics Approaches. *BioMed research international*, *2016*, 5040598. https://doi.org/10.1155/2016/5040598.

Bi, C., Paterson, A.H., Wang, X., Xu, Y., Wu, D., Qu, Y., Jiang, A., Ye, Q., Ye, N. (2016). Analysis of the Complete Mitochondrial Genome Sequence of the Diploid Cotton *Gossypium raimondii* by Comparative Genomics Approaches. *BioMed Research International*. 1–18. doi:10.1155/2016/5040598

Bitz, L., Heinonen, M., Moisander, S., Tanhuanpää, P., Sarvarinne, J. (2019). Differential Allelic Richness between *Malus sylvestris* L. and *Malus* × *domestica* Borkh. from Finland as a Measure of Genetic Loss. *Sustainability* **.** *11*, 6949. https://doi.org/10.3390/su11246949.

Bourguiba, H., Scotti, I., Sauvage, C., Zhebentyayeva, T., Ledbetter, C., Krška, B., Remay, A., D'Onofrio, C., Iketani, H., Christen, D., Krichen, L., Trifi-Farah, N., Liu, W., Roch, G., & Audergon, J. M. (2020). Genetic Structure of a Worldwide Germplasm Collection

of *Prunus armeniaca* L. Reveals Three Major Diffusion Routes for Varieties Coming From the Species' Center of Origin. *Frontiers in plant science*, *11*, 638. https://doi.org/10.3389/fpls.2020.00638.

Bowen, L. (2023). 11 reasons to avoid planting Bradford pear trees. Epic gardening. https://www.epicgardening.com/avoid-bradford-pear/

Brudno, M., Malde, S., Poliakov, A., Do, C.B., Couronne, O., Dubchak, I., and Batzoglou, S. 2003. Bioinformatics, 19S1: i54-i62.

Chatzidimopoulos, M., Ganopoulos, I., Madesis, P., Vellios, E., Tsaftaris, A., & Pappas, A. C. (2014). High-resolution melting analysis for rapid detection and characterization of Botrytis cinerea phenotypes resistant to fenhexamid and boscalid. *Plant pathology*, *63*(6), 1336-1343.

Chatzidimopoulos, M., Ganopoulos, I., Moraitou-Daponta E., Lioliopoulou, F., Ntantali, O., Panagiotaki, P., and Vellios, E.K. (2019). High-Resolution Melting (HRM) Analysis Reveals Genotypic Differentiation of Venturia inaequalis Populations in Greece. Frontiers in Ecology and Evolution, 7. https://www.frontiersin.org/articles/10.3389/fevo.2019.00489

Chung, H. Y., Won, S. Y., Kang, S. H., Sohn, S. H., & Kim, J. S. (2017). The complete mitochondrial genome of Wonwhang (*Pyrus pyrifolia*). *Mitochondrial DNA. Part B, Resources*, *2*(2), 902–903. https://doi.org/10.1080/23802359.2017.1413300

Coyle, D. R., Williams, B. M., & Hagan, D. L. (2021). Fire Can Reduce Thorn Damage by the Invasive Callery Pear Tree. *HortTechnology hortte*, *31*(5), 625-629. Retrieved May 17, 2024, from https://doi.org/10.21273/HORTTECH04892-21

Creech, J. L. (1973). Ornamental Plant Introduction—Building On the Past. *Arnoldia*. 33(1): 13–25.

Cuizhi, G., & Spongberg, S. A. (2003). Pyrus. *Flora of China*. *9*:173-179.

Culley, T.M & Hardiman, M.N. (2008). The role of intraspecific hybridization in the evolution of invasiveness: A case study of the ornamental pear tree *Pyrus calleryana*. *Biological Invasions. 11*(5): 1107-1119. http://dx.doi.org/10.1007/s10530-008-9386-z

Culley, T.M. (2017). The rise and fall of the ornamental Callery pear tree. *Arnoldia, 74*(3), 2-11.

Culley, T.M. and Hardiman, N.A. (2007). The Beginning of a New Invasive Plant: A History of the Ornamental Callery Pear in the United States. *BioScience, 57*(11), 956-964. https://doi.org/10.1641/B571108

Dhama, K., Karthik, K., Chakraborty, S., Tiwari, R., Kapoor, S., Kumar, A., & Thomas, P. (2014). Loop-mediated isothermal amplification of DNA (LAMP): a new diagnostic tool lights the world of diagnosis of animal and human pathogens: a review. *Pakistan journal of biological sciences: PJBS*, *17*(2), 151–166. https://doi.org/10.3923/pjbs.2014.151.166

Diagne, C., Leroy, B., Vaissière, A.-C., Gozlan, R. E., Roiz, D., Jarić, I., et al. (2021). High and Rising Economic Costs of Biological Invasions Worldwide. *Nature.* 592 (7855):571–576. doi:10.1038/s41586-021-03405-6

Dickson, E. E., Arumuganathan, K., Kresovich, S., & Doyle, J. J. (1992). Nuclear DNA content variation within the Rosaceae. *American journal of botany*, *79*(9), 1081-1086.

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology, 17*(1), 431-449.

Domingo, R., Perez, C., Klair, D. *et al.* (2021). Genome-informed loop-mediated isothermal amplification assay for specific detection of *Pectobacterium parmentieri* in infected potato tissues and soil. *Sci Rep* **11**, 21948. https://doi.org/10.1038/s41598-021-01196-4.

Duan, Y., Zhang, X., Ge, C. *et al.* (2014). Development and application of loop-mediated isothermal amplification for detection of the F167Y mutation of carbendazim-resistant isolates in *Fusarium graminearum*. *Sci Rep.* **4**:7094. https://doi.org/10.1038/srep07094

Dumolin, S., Demesure, B., & Petit, R. J. (1995). Inheritance of chloroplast and mitochondrial genomes in pedunculate oak investigated with an efficient PCR method. *TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik*, *91*(8), 1253–1256. https://doi.org/10.1007/BF00220937

Dunphy, B., & Hamrick, J. (2005). Gene flow among established Puerto Rican populations of the exotic tree species, *Albizia lebbeck*. *Heredity, 94*(4), 418-425.

E. Barnard. (2002). *New York City trees: a field guide for the metropolitan area*. Columbia University Press, New York.

Emms, D.M., Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**, 238 https://doi.org/10.1186/s13059-019-1832-y.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology, 14*(8), 2611-2620.

Fischbach, J., Xander, N. C., Frohme, M., & Glökler, J. F. (2015). Shining a light on LAMP assays--a comparison of LAMP visualization methods including the novel use of berberine. *BioTechniques*, *58*(4), 189–194. https://doi.org/10.2144/000114275

Fleury, D., Baumann, U. and Langridge, P. (2012). Plant genome sequencing: Models for developing synteny maps and association mapping. *Plant Biotechnology and Agriculture*. 83-97. Doi: 10.1016/B978-0-12-381466-1.00006-7

Focke, F., Haase, I., & Fischer, M. (2013). Loop-mediated isothermal amplification (LAMP): methods for plant species identification in food. *Journal of agricultural and food chemistry*, *61*(12), 2943–2949. https://doi.org/10.1021/jf304295b

Forsythe, E. S., Grover, C. E., Miller, E. R., Conover, J. L., Arick, M. A., Chavarro, M. C. F. et Sloan, D. B. (2022). Organellar transcripts dominate the cellular mRNA pool across plants of varying ploidy levels. *Proceedings of the national academy of sciences*, *119*(30). doi:10.1073/pnas.2204187119

Francis, R. M. (2017). pophelper: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources, 17*(1), 27-32.

Frazer, K. A., Pachter, L., Poliakov, A., Rubin, E. M., & Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic acids research*, *32*(Web Server issue), W273–W279. https://doi.org/10.1093/nar/gkh458.

Fulcher, B. (2002). Trouble on Main Street. *Tennessee Conservationist*. *68*:12-15.

Gawkins K. (2019). The Prickly Problems of Pear Potential Effects of Native Orthopteran Herbivory on an Invasive Woody Plant. University of Dayton.

Goto, M., Eiichi Honda, Atsuo Ogura, Akio Nomoto, and Ken-Ichi Hanaki. (2009). Colorimetric detection of loop-mediated isothermal amplification reaction by using hydroxy naphthol blue. *BioTechniques*. 46:3: 167-172.

Goudet, J. (2005). Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Notes, 5*(1), 184-186.

Gouy M., Guindon S. & Gascuel O. (2010) SeaView version 4 : a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27(2)**:221-224.

Gouy M., Guindon S. & Gascuel O. (2010) SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Molecular Biology and Evolution 27(2):221-224.

Guillaume Collet (2012) https://github.com/gcollet/MstatX.

Gurevitch J. and Padilla D.K. (2004). Are invasive species a major cause of extinctions?**.** *Trends in Ecology and Evolution*. *19*: 470-474.

Hadziabdic, D., Wadl, P. A., Vito, L. M., Boggess, S. L., Scheffler, B. E., Windham, M. T., & Trigiano, R. N. (2012). Development and characterization of sixteen microsatellite loci for *Geosmithia morbida*, the causal agent of thousand canker disease in black walnut (*Juglans nigra*). *Conservation Genetics Resources, 4*(2), 287-289.

Hall, T.A. (1999). **BioEdit**: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT.

Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings. Biological sciences*, *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218

Hebert, P. D., Cywinska, A., Ball, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings. Biological sciences*, *270*(1512), 313–321. https://doi.org/10.1098/rspb.2002.2218.

Higgins, A. Scientists thought they had created the perfect tree. But it became a nightmare. The Washington Post, 2018. https://www.washingtonpost.com/how-we-turned-the-bradford-pear-into-a-monster

Hsieh, K., Mage, P. L., Csordas, A. T., Eisenstein, M., & Soh, H. T. (2014). Simultaneous elimination of carryover contamination and detection of DNA with uracil-DNA-glycosylase-supplemented loop-mediated isothermal amplification (UDG-

LAMP). *Chemical communications (Cambridge, England)*, *50*(28), 3747–3749. https://doi.org/10.1039/c4cc00540f

IUCN (2000). IUCN GUIDELINES FOR THE PREVENTION OF BIODIVERSITY LOSS CAUSED BY ALIEN INVASIVE SPECIES

Jansen, R.K., Raubeson, L.A., Boore, J.L., and dePamphilis, C.W. (2005). Methods for Obtaining and Analyzing Whole Chloroplast Genome

Jia, B., Li, X., Liu, W., Lu, C., Lu, X., Ma, L., Li, Y. Y., & Wei, C. (2019). GLAPD: Whole Genome Based LAMP Primer Design for a Set of Target Genomes. *Frontiers in microbiology*, *10*, 2860. https://doi.org/10.3389/fmicb.2019.02860.

Kamvar, Z. N., Tabima, J. F., & Grünwald, N. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *2*, e281.

Kariyawasam, C.S., Kumar, L., and Ratnayake, S.S. (2019). Invasive Plant Species Establishment and Range Dynamics in Sri Lanka under Climate Change. *Entropy (Basel)*, 21(6),571. doi: 10.3390/e21060571. PMID: 33267285; PMCID: PMC7515060.

Kato, S., Imai, A., Rie, N., & Mukai, Y. (2013). Population genetic structure in a threatened tree, *Pyrus calleryana* var. *dimorphophylla* revealed by chloroplast DNA and nuclear SSR locus polymorphisms. *Conservation Genetics Resources, 14*(5), 983-996.

Katoh et al., (2002). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization.

Kim, S.H., Lee, S.Y., Kim, U., Oh, S.W. (2023). Diverse methods of reducing and confirming false-positive results of loop-mediated isothermal amplification assays: a review. *Anal. Chim. Acta.* 1280: 341693.

Kitamura M, Aragane M, Nakamura K, Adachi T, Watanabe K, Sasaki Y. (2018). Improved On-Site Protocol for the DNA-Based Species Identification of *Cannabis sativa* by Loop-Mediated Isothermal Amplification. *Biol Pharm Bull*.;41(8):1303-1306. doi: 10.1248/bpb.b18-00272. PMID: 30068883.

Kurosaki Y, Takada A, Ebihara H, Grolla A, Kamo N, Feldmann H, Kawaoka Y, Yasuda J. (2007). Rapid and simple detection of Ebola virus by reverse transcription-loop-mediated

isothermal amplification. *J Virol Methods*. 141(1):78-83. doi: 10.1016/j.jviromet.2006.11.031. Epub 2006 Dec 27. PMID: 17194485.

Lapidos, J. (2013). Your Street Probably Smells Like Semen Right Now but It Might Not Next Spring. The Awl. https://theawl.com/your-street-probably-smells-like-semen-right-now-but-it-might-not-next-spring-90880b2d941c?gi=f9d143a1124

Leigh, F. J., Mackay, I., Oliveira, H. R., Nicholas, E., Richard, G. A., Huw, H. J., et al. (2013). Using diversity of the chloroplast genome to examine evolutionary history of wheat species. Genet. Resour. Crop Evol. 60, 1831–1842. doi: 10.1007/s10722-013-9957-4.

Li WH, Zhang CB, Jiang HB, Xin GR, Yang ZY (2006) Changes in soil microbial community associated with invasion of the exotic weed, *Mikania micrantha* H.B.K. *Plant Soil* 281: 309–324. doi: 10.1007/s11104-005-9641-3.

Li, J.J., Xiong, C., Liu, Y., Liang, J.S., Zhou, X.W. (2016). Loop-Mediated Isothermal Amplification (LAMP): Emergence As an Alternative Technology for Herbal Medicine Identification. *Frontiers in Plant Science*. 7:1956. DOI: 10.3389/fpls.2016.01956. PMID: 28082999; PMCID: PMC5183589.

Li, R. , Wang, C. , Ji, L. , Zhao, X. , Liu, M. , Zhang, D. and Shi, J. (2015). Loop-Mediated Isothermal Amplification (LAMP) Assay for GMO on: Recent Progresses and Future Perspectives. *Open Access Library Journal*, **2**, 1-8. doi: 10.4236/oalib.1101264.

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y. T., and Chen, S. (2015). Plant DNA barcoding:from genes to genome. Biol. Rev. 90, 157–166. doi: 10.1111/brv.12104.

Liu, J., Zheng, X., Potter, D., Hu, C., & Teng, Y. J. B. s. (2012). Genetic diversity and population structure of Pyrus calleryana (Rosaceae) in Zhejiang province, China. *Biochemical Systematics Ecology, 45*, 69-78.

Maloney, M.E., E.B. Borth, G. Dietch, M.C. Lloyd and R.W. McEwan**.** (2023). A trial of fire and ice: assessment of control techniques for *Pyrus calleryana* stems during grassland restoration in southwestern Ohio, USA. *Ecological Restoration*. 41: 25-33. *Open Access:* https://muse.jhu.edu/pub/19/article/883911

Margaret E. Maloney, Abby Hay, Eric B. Borth, Ryan W. McEwan. (2022). Leaf phenology and freeze tolerance of the invasive tree *Pyrus calleryana* (Roseaceae) and potential native competitors. *The Journal of the Torrey Botanical Society*. 149(4): 273-279.

Matheson, P. & McGaughran, A. (2022). Genomic data is missing for many highly invasive species, restricting our preparedness for escalating incursion rates. *Scientific Reports.* 12, 13987.

Maxwell, S., Fuller, R., Brooks, T. *et al.* (2016). Biodiversity: The ravages of guns, nets and bulldozers. *Nature. 536*:143–145 https://doi.org/10.1038/536143a

McNeely JA, Mooney HA, Neville LE, Schei PJ and Waage JK (eds). (2001). Global Strategy on Invasive Alien Species. IUCN, Gland, Switzerland

McNeely, J. (2001). Invasive Species: A Costly Catastrophe for Native Biodiversity. IUCN Biodiversity Programme, Gland, Switzerland. *Land Use and Water Resources Research. 2*: 1-10.

Meyer, F. N. (1918). South China Explorations: Typescript, July 25, 1916–Setpember 21, 1918. The National Agricultural Library. Available online at: https://archive.org/details/CAT10662165MeyerSouthChinaExplorations.

Morewood, W.D., Hoover, K., Neiner, P.R., McNeil, J.R. and Sellmer, J.C. (2004), Host tree resistance against the polyphagous wood-boring beetle *Anoplophora glabripennis*. Entomologia Experimentalis et Applicata, 110: 79-86. https://doi.org/10.1111/j.0013-8703.2004.00120.x.

Mori, Y., Nagamine, K., Tomita, N., & Notomi, T. (2001). Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. *Biochemical and biophysical research communications*. *289*(1): 150–154. https://doi.org/10.1006/bbrc.2001.5921

Morley, S.A., Nielsen, B.L. (2017). Plant mitochondrial DNA. *Front. Biosci. (Landmark Ed)*, 22(6), 1023–1032. https://doi.org/10.2741/4531

Nagamine, K., Hase, T., & Notomi, T. (2002). Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Molecular and cellular probes*, *16*(3), 223–229. https://doi.org/10.1006/mcpr.2002.0415.

Nagamine, K., Watanabe, K., Ohtsuka, K., Hase, T., & Notomi, T. (2001). Loop-mediated isothermal amplification reaction using a nondenatured template. *Clinical chemistry*. *47*(9): 1742-1743.

Nebhut, A. N., & Dukes, J. S. (2023). Invasion by Callery pear (Pyrus calleryana) does not affect understory abundance or diversity in early-successional meadows. *Invasive Plant Science and Management*, *16*(4), 215–224. doi:10.1017/inp.2023.28

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics, 89*(3), 583-590.

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N. and Hase, T. (2000). Loop-mediated isothermal amplification of DNA, *Nucleic Acids Research*, *28*(12), e63 https://doi.org/10.1093/nar/28.12.e63

Nowicki M, Boggess SL, Saxton AM, Hadziabdic D, Xiang Q-YJ, Molnar T, et al. (2018). Haplotyping of *Cornus florida* and *C. kousa* chloroplasts: Insights into species-level differences and patterns of plastic DNA variation in cultivars. PLoS ONE 13(10): e0205407. https://doi.org/10.1371/journal.pone.0205407

Nowicki, M., Huff, M. L., Staton, M. E., & Trigiano, R. N. (2022). Chloroplast genome of the invasive Pyrus calleryana underscores the high molecular diversity of the species. *Journal of applied genetics*, *63*(3), 463–467. https://doi.org/10.1007/s13353-022-00699-

Nowicki, M., Nowakowska, M., Wrzesińska, M., and Kozik, E. U. (2015). "Development of LAMP-HRM for sensitive and specific detection of Phytophthora infestans," in Brasov, Romania: DLO Foundation, ed. H. T. A. M. Schepers 237–238.

Oscorbin, I.P., Belousova, E.A, Boyarskikh, U.A, Zakabunin, A.I., Khrapov, E.A. and Filipenko, M.L. (2017). Derivatives of *Bst*-like Gss-polymerase with improved processivity and inhibitor tolerance. *Nucleic Acids Research*, *45*(16), 9595–9610 https://doi.org/10.1093/nar/gkx645

Parida M, Sannarangaiah S, Dash PK, Rao PV, Morita K. (2008). Loop mediated isothermal amplification (LAMP): a new generation of innovative gene amplification technique; perspectives in clinical diagnosis of infectious diseases. *Rev Med Virol*. 18(6):407-21. doi: 10.1002/rmv.593. PMID: 18716992; PMCID: PMC7169140.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes, 6*(1), 288-295.

Pejchar, L. and Mooney, H. (2009). Invasive species, ecosystem services and human well-being. *Trends in Ecology and Evolution.* 24:497-504. 10.1016/j.tree.2009.03.016.

Phillips, W. D., Ranney, T. G., Touchell, D. H., & Eaker, T. A. (2016). Fertility and Reproductive Pathways of Triploid Flowering Pears (Pyrus sp.), *HortScience,* 51(8), 968-971. Retrieved Dec 5, 2022, from https://journals.ashs.org/hortsci/view/journals/hortsci/51/8/article-p968.xml

Pimental, D., Zuniga, R, and Morrison, D. (2005). Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics, 52(*3), 273-288.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945-959.

Pyšek, P. and D. M. Richardson (2010). Invasive Species, Environmental Change and Management, and Health. *Annual Review of Environment and Resources, 35*(1), 25-55.

Ranjbar R, Afshar D. (2015). Development of a loop-mediated isothermal amplification assay for rapid detection of Yersinia enterocolitica via targeting a conserved locus. *Iran J Microbiol*. 7(4):185-90. PMID: 26697156; PMCID: PMC4685161.

Ravan H, Amandadi M, Sanadgol N. (2016). A highly specific and sensitive loop-mediated isothermal amplification method for the detection of Escherichia coli O157:H7. *Microb Pathog*. 91:161-5. doi: 10.1016/j.micpath.2015.12.011. Epub 2015 Dec 25. PMID: 26724736.

Renz, M., Gibson, K., Hillmer, J., Howe, K., Waller, D., & Cardina, J. (2009). Land Manager and Researcher Perspectives on Invasive Plant Research Needs in the Midwestern United States. *Invasive Plant Science and Management, 2*(1), 83-91. doi:10.1614/IPSM-08-109.1

Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. and West, C.J. (2000), Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, 6: 93-107. https://doi.org/10.1046/j.1472-4642.2000.00083.x

Roth, S.A. (2021). Landscape Tree Gone Wild and Rogue. BLUE RIDGE PARTNERSHIP FOR REGIONAL INVASIVE SPECIES MANAGEMENT.

Sahoo, P. R., Sethy, K., Mohapatra, S., & Panda, D. (2016). Loop-mediated isothermal amplification: An innovative gene amplification technique for animal diseases. *Veterinary world*, *9*(5), 465–469. https://doi.org/10.14202/vetworld.2016.465-469.

Santamour, F.S., & Demuth, P. (1980). Identification of Callery pear cultivars by peroxidase isozyme patterns, *Journal of Heredity*, 71(6):447-449, https://doi.org/10.1093/oxfordjournals.jhered.a109412

Sapkota, S., Boggess, S. L., Trigiano, R. N., Klingeman, W. E., Hadziabdic, D., Coyle, D. R., Olukolu, B. A., et al. (2021). Microsatellite Loci Reveal Genetic Diversity of Asian Callery Pear (Pyrus calleryana) in the Species Native Range and in the North American Cultivars. *Life*, *11*(6), 531. MDPI AG. Retrieved from http://dx.doi.org/10.3390/life11060531

Sapkota, S., Boggess, S.L., Trigiano, R.N., Klingeman, W.E., Hadziabdic, D., Coyle, D.R., and Nowicki, M. (2022). Microsatellite Loci Reveal High Genetic Diversity, Mutation, and Migration Rates as Invasion Drivers of Callery Pear (*Pyrus calleryana*) in the Southeastern United States. *Front Genet.13*:861398. doi: 10.3389/fgene.2022.861398.

Semerikova, S. A., and Semerikov, V. L. (2014). Molecular phylogenetic analysis of the genus Abies (Pinaceae) based on the nucleotide sequence of chloroplast DNA. Russ J. Genet. 50, 7–19. doi: 10.1134/S1022795414010104.

Sequences. *Methods in Enzymology, 395*, 348-384. https://doi.org/10.1016/S0076-6879(05)95020-9.

Serota, T. and T.M. Culley. 2019. Seedgermination and seedling survival of inva-sive callery pear (Pyrus calleryanaDecene.) 11 years after fruit collection.Castanea 84:47–52.

Shezi, Z. and Poona K.N. (2010). An investigation into using different satellite remote sensors and techniques to identify, map, monitor and predict the spread and distribution of some of the major current and emerging invasive alien plant species in KwaZulu-Natal. Invasive Alien Species Project. UKZN/DAEA.

Smith, D.R. (2015). Mutation rates in plastid genomes: they are lower than you might think. *Genome biology and evolution*, 7(5): 1227-34. doi:10.1093/gbe/evv069

Smith, D.R., and Keeling, P.J. (2015). Mitochondrial and plastid genome architecture: Reoccurring themes, but significant differences at the extremes. Proceedings of the National Academy of Sciences of the United States of America, 112(33): 10177-84. doi:10.1073/pnas.1422049112

Srivastava, S., Avvaru, A. K., Sowpati, D. T. & Mishra, R. K. (2019). Patterns of microsatellite distribution across eukaryotic genomes. *BMC Genom*. 20(153). https://doi.org/10.1186/s12864-019-5516-5 (2019).

Srivastava, S., Avvaru, A.K., Sowpati, D.T. *et al.* (2019). Patterns of microsatellite distribution across eukaryotic genomes. *BMC Genomics* **20**, 153. https://doi.org/10.1186/s12864-019-5516-5.

Stewart B. (1999). The coming plague of pears. Green Industry News. 5: 1(29 October 2007; http://ipmnet.umd.edu/5-8art1.htm).

Stoddart, J. A., & Taylor, J. F. (1988). Genotypic diversity: estimation and prediction in samples. *Genetics, 118*(4), 705-711.

Swearingen, J., Reshetiloff, K., Slattery, B., & Zwicker, S. (2002). Plant invaders of mid-Atlantic natural areas. *National Park Service and US Fish and Wildlife Service (Washington, DC, 168pp)*.

Team, R. C. (2013). R: A language and environment for statistical computing.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R and Greiner S (2017) GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* 45: W6-W11.

Tobin P. C. (2018). Managing invasive species. *F1000Research*, *7*, F1000 Faculty Rev-1686. https://doi.org/10.12688/f1000research.15414.1.

Tobin, P.C. (2018). Managing invasive species. 7:F1000 Faculty Rev-1686. doi: 10.12688/f1000research.15414.1.

Tomita, N., Mori, Y., Kanda, H., & Notomi, T. (2008). Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. *Nature protocols*, *3*(5), 877–882. https://doi.org/10.1038/nprot.2008.57

Torres, C., Vitalis, E. A., Baker, B. R., Gardner, S. N., Torres, M. W., & Dzenitis, J. M. (2011). LAVA: an open-source approach to designing LAMP (loop-mediated isothermal amplification) DNA signatures. *BMC bioinformatics*, *12*, 240. https://doi.org/10.1186/1471-2105-12-240.

Vieira, M. L. C., Santini, L., Diniz, A. L., & Munhoz, C. (2016). Microsatellite markers: what they mean and why they are so useful. *Genetics Molecular Biology, 39*(3), 312-328.

Vilà, M, Espinar, JL, Hejda, M, Hulme, PE, Jarošík, V, Maron, JL, Pergl, J, Schaffner, U, Sun, Y , Pyšek, P. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecol Lett*, *14*,702–708.

Vincent, M.A. (2005). On the spread and current distribution *of Pyrus calleryana* in the United States. *Castanea, 70*, 20-31.

Vogt JT, Coyle DR, Jenkins D, Barnes C, Crowe C, Horn S, Bates C, and Roesch FA (2020). Efficacy of five herbicide treatments for control of Pyrus calleryana. *Invasive Plant Sci*. Manag 13: 252–257. doi: 10.1017/inp.2020.36

Wang FP, Wang ZJ, Li YX. (2019). Comparison of GC contents in genomes of the three model microbes. Genom Appl Biol. 38:2215–2220. doi:10.13417/j.gab.038.002215.

Wang, R., Luo, Y., Lan, Z. et al. (2024). Insights into structure, codon usage, repeats, and RNA editing of the complete mitochondrial genome of Perilla frutescens (Lamiaceae). *Sci Rep* 14, 13940. https://doi.org/10.1038/s41598-024-64509-3

Wang, S.K., Xuezhu Liao, Jiawei Zhou, Luke R. Tembrock, Henry Daniell, Shuangxia Jin, Zhiqiang Wu. (2024). Plant organellar genomes: much done, much more to do. *Trends in Plant Science.*

Warrix, A., & Marshall, J. (2018). Callery pear (Pyrus calleryana) Response to Fire in a Managed Prairie Ecosystem. *Invasive Plant Science and Management, 11*(1), 27-32. doi:10.1017/inp.2018.4

Warrix, A.R., A.L. Myers, and J.M.Marshall. (2017). Estimating invading Callery Pear (*Pyrus calleryana*) age and flowering probability in an Indianamanaged prairie. *Proc. Indiana Acad.Sci*. 126:153–157.

Whitehouse, W.E., J.L. Creech, and G.A. Seaton (1963) Bradford Ornamental Pear—A Promising Shade Tree. *American Nurseryman* 117: 7–8, 56–61.

Wolfe, K. H., Li, W. H., & Sharp, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(24), 9054–9058. https://doi.org/10.1073/pnas.84.24.9054

Wong, Y.-P., Othman, S., Lau, Y.-L., Radu, S. and Chee, H.-Y. (2018). Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. *J Appl Microbiol*, 124: 626-643. https://doi.org/10.1111/jam.13647

Wong, Y.P., Othman, S., Lau, Y.L., Radu, S. and Chee, H.Y. (2018). Loop-mediated isothermal amplification (LAMP): a versatile technique for detection of micro-organisms. J Appl Microbiol, 124: 626-643. https://doi.org/10.1111/jam.13647.

Woods, M., Grace, A. and Ryan, M. (2021). Resprouting of the woody plant Pyrus calleryana influences soil ecology during invasion of grasslands in the American Midwest. *Applied Soil Ecology*, 166. 10.1016/j.apsoil.2021.103989.

Woods, Michaela J et al. (2023). *Pyrus calleryana* extracts reduce germination of native grassland species, suggesting the potential for allelopathic effects during ecological invasion. *PeerJ*, 11 e15189. doi:10.7717/peerj.15189

Wu, P., Xu, C., Chen, H., Yang, J., Zhang, X. and Zhou, S. (2021). NOVOWRAP: An automated solution for plastid genome assembly and structure standardization. Mol Ecol Resour, 21: 2177-2186. https://doi.org/10.1111/1755-0998.13410

Xia, X. (2013). What is Comparative Genomics? In: Comparative Genomics. SpringerBriefs in Genetics. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-37146-2_1

Xiang Q-H, Zhang D-X, Wang Q, Wang X-R, Guan W-B. (2019). The complete chloroplast genome sequence of *Pyrus phaeocarpa* Rehd. *Mitochondrial DNA Part B.* 4:1370–1371. doi: 10.1080/23802359.2019.1598804.

Zhang, J., Qian, J.Y., Bian, Y.H., Xiao, L., and Wang, C.L. (2022). Transcriptome and Metabolite Conjoint Analysis Reveals the Seed Dormancy Release Process in Callery

Pear. *International Journal of Molecular Sciences* 23, 4: 2186. https://doi.org/10.3390/ijms23042186.

Zhang, M., Liu, Y.N., Chen, L.L., Quan, S., Jiang, S.M., Zhang, D.B. and Yang, L.T. (2013). One Simple DNA Extraction Device and Its Combination with Modified Visual Loop-Mediated Isothermal Amplification for Rapid On-Field Detection of Genetically Modified Organisms. Analytical Chemistry, 85, 75-82. http://dx.doi.org/10.1021/ac301640p.

Zhu, B.R., Barrett, S.C.H., Zhang, D.Y., et al. (2017). Invasion genetics of Senecio vulgaris: loss of genetic diversity characterizes the invasion of a selfing annual, despite multiple introductions. *Biol Invasions,* 19:255–267.

Zimin, A.V., Guillaume Marçais, Daniela Puiu, Michael Roberts, Steven L. Salzberg, James A. Yorke, (2013). The MaSuRCA genome assembler, *Bioinformatics*, 29(21): 2669–2677, https://doi.org/10.1093/bioinformatics/btt476.

Zuntini, A.R., Carruthers, T., Maurin, O. et al. (2024). Phylogenomics and the rise of the angiosperms. *Nature*, 629: 843–850. https://doi.org/10.1038/s41586-024-07324-0.

VITA

Alina Pokhrel was born in Arghakhanchi, Nepal on July 18, 1995. She had her Bachelor of Science (BSc) in Institite of Agriculture and Animal Science (IAAS), Tribhuvan University, Lamjung, Nepal in May 2019. After some years of professional experience in her home country, she moved to the United States to pursue her master's degree in the University of Tennessee. She joined the Department of Entomology and Plant Pathology in August 2022 under the supervision of Dr. Marcin Nowicki.