



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원 저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리와 책임은 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)



석사학위 논문

제주도 자생 왕벚나무의
계통 및 유전구조 분석

Phylogenetic and Population Structure Analysis of
Wild *Prunus yedoensis* (Rosaceae) in Jeju Island, Korea

명지대학교 대학원

생명과학정보학과

조아라

지도교수 문정환

2017년 2월

제주도 자생 왕벚나무의
계통 및 유전구조 분석

Phylogenetic and Population Structure Analysis of Wild
Prunus yedoensis (Rosaceae) in Jeju Island, Korea

이 논문을 석사학위 논문으로 제출함.

2017년 2월

명지대학교 대학원
생명과학정보학과
조아라

제주도 자생 왕벚나무의
계통 및 유전구조 분석

Phylogenetic and Population Structure Analysis of Wild
Prunus yedoensis (Rosaceae) in Jeju Island, Korea

명지대학교 대학원
생명과학정보학과
조아라

상기자의 이학석사 학위논문을 인준함.

심사위원장 최상봉 (인)
심사위원 문정환 (인)
심사위원 송상익 (인)

2017년 2월

감사의 글

설레는 마음으로 명지대학교 생명과학정보학과에 입학하여 어느새 4년의 학부과정과 2년의 석사과정을 마치게 되었습니다. 지난 시간들을 뒤돌아보니 이 논문이 나오기 까지 많은 분들의 도움을 받았습니다. 그 분들께 감사한 마음을 가지며 이 자리를 빌려 감사의 인사를 드리고자 합니다.

제가 맡은 일들을 스스로 할 수 있게 기다려 주시며 부족한 점들을 언제나 섬세하고 꼼꼼하게 지적해 주신 문정환 교수님께 진심으로 감사드립니다. 연구 과정뿐만 아니라 연구자로서의 태도와 인생을 살아가는데 있어 많은 조언을 해주셔서 감사합니다. 앞으로 A+ 학생이 될 수 있도록 항상 노력하겠습니다. 더하여 바쁜 와중에도 석사학위 논문에 따뜻한 격려와 많은 조언을 해주신 최상봉 교수님과 송상의 교수님께도 감사드립니다.

연구 과정에 있어서 모르는 부분들을 함께 토론해 주시고 문제들을 해결하는데 도움을 주신 김군보 박사님과 백승훈 선생님께 감사드립니다. 그리고 지칠 때마다 노래를 부르며 힘차게 응원해준 장호열 오빠와 김문진 언니, 함께 해줘서 너무 고마워요. 지금은 회사에 취직했지만, 졸래졸래 쫓아다며 이것저것 물어봐도 귀찮다는 기색 없이 대답해준 저의 사수, 권영은 언니도 고마워요.

마지막으로 ‘후회 없는 삶’을 살 수 있도록 제 선택을 존중해 주신 아버지 어머니께 감사의 인사를 드립니다. 언제나 변함없이 저를 믿고 기다려 주셔서 감사합니다. 그 믿음에 보답할 수 있는 큰 사람으로 성장할 수 있도록 항상 노력하겠습니다. 그리고 내 동생 철영이, 전화 용건이 없어 매번 ‘그냥’ 이라는 단어로 시작되는 통화를 묵묵히 받아줘서 고맙다.

또한 일일이 나열할 수는 없지만 늘 저를 응원해주시고 격려해주시는 많은 분들께 감사드립니다.

조아라 올림

목 차

목차.....	i
그림 목차.....	iii
표 목차.....	v
국문초록.....	vi

제 1 장 서 론

제 1 절 연구의 배경 및 필요성.....	1
제 2 절 연구의 목표.....	6

제 2 장 연구 재료 및 방법

제 1 절 연구 재료.....	7
1. 식물 재료	7
2. Genomic DNA 추출.....	8
제 2 절 연구 방법.....	8
1. 왕벚나무 유전자 서열 및 벚나무속 COS 유전자 선발.....	8
2. COS 분자 마커 개발 및 선발.....	9
3. 계통 및 유전 분석.....	13
4. PrCOS 분자 마커 목표 영역의 서열 분석.....	14

제 3 장 연구 결과 및 고찰

제 1 절 제주도 자생 왕벚나무 및 근연종 채집 및 집단 구성.....	15
1. 제주도 자생 왕벚나무 기념목 지정 개체.....	15
2. 제주도 자생 왕벚나무.....	19
3. 제주도 자생 왕벚나무 근연종 벚나무류.....	23
제 2 절 PrCOS 분자 마커의 개발 및 검증.....	27

1. 벚나무속 PrCOS 유전자 선발.....	27
2. 왕벚나무 이형성 PrCOS 마커 설계.....	30
3. PrCOS 마커 검증 및 단계적 선발.....	38
제 3 절 제주도 왕벚나무 자생 집단의 유전형 및 유전구조 분석.....	47
1. 왕벚나무 자생 집단의 개체별 유전형 분석.....	47
2. 제주도 왕벚나무 자생 집단의 유전구조 및 계통 비교 분석.....	54
제 4 절 관음사 및 봉개동 왕벚나무 자생 집단의 개체별 유전형 분석.....	62
1. 벚나무 근연종의 유전구조 분석 및 왕벚나무와의 계통 비교 분석.....	62
제 5 절 제주도 자생 왕벚나무의 부계 추정.....	69
1. 왕벚나무 이형성 PrCOS 분자 마커의 목표 영역 서열 추출.....	69
2. 서열 일치도 분석.....	70
 제 4 장 결 론.....	72
 참고 문헌.....	74
 부 록.....	78
 Abstract.....	109

그림 목차

Figure 1. Photographs of native wild <i>P. yedoensis</i> in Jeju island, Korea	3
Figure 2. Composite geographical locations of wild <i>P. yedoensis</i> and its closely relative species plants used in this study	16
Figure 3. Locations of the wild <i>P. yedoensis</i> plants registered as the Natural Monument in Jeju island, Korea.....	17
Figure 4. Geographical locations of the wild <i>P. yedoensis</i> populations in Jeju island.....	22
Figure 5. Geographical locations of closely relative <i>Prunus</i> populations of wild <i>P. yedoensis</i> in Jeju island.....	26
Figure 6. Workflow to select <i>Prunus</i> Conserved Ortholog Sets (PrCOS) candidate genes using stepwise BLASTP searches of <i>P. mume</i> and <i>P. persica</i> genomes.....	28
Figure 7. Workflow to select heterozygous Conserved Ortholog Set candidate genes in <i>P. yedoensis</i>	29
Figure 8. A diagram representing strategy of InDel marker development and expected result of marker application.....	33
Figure 9. Examples of PCR amplification of PrCOS markers developed from transcript unigenes of <i>P. yedoensis</i>	35
Figure 10. Examples of PCR amplification of PrCOS markers developed from Illumina short read assemblies (A) or predicted gene models of a draft genome of <i>P. yedoensis</i> (B).....	36
Figure 11. Locations of 156 PrCOS genes mapped to the <i>P. mume</i> chromosomes.....	37
Figure 12. Examples of PCR amplification for selection of dimorphic PrCOS markers in <i>P. yedoensis</i> . Step 1: Application for <i>P. yedoensis</i> and <i>P. pendula</i> for. <i>ascendens</i>	40
Figure 13. Examples of PCR amplification for selection of dimorphic PrCOS markers in <i>P. yedoensis</i> . Step 2: Application for <i>P. yedoensis</i> and its closely relative species	41

Figure 14. Examples of PCR amplification for selection of dimorphic PrCOS markers in <i>P. yedoensis</i> . Step 3: Application for <i>P. yedoensis</i> and its all available related species	42
Figure 15. Summary of selected PrCOS markers.....	43
Figure 16. Examples of applied of the dimorphic PrCOS markers for genotyping of wild <i>Prunus</i> populations	50
Figure 17. A maximum Likelihood tree of <i>P. yedoensis</i> and its closely related species using 20 PrCOS markers.....	55
Figure 18. Genetic structure analysis of populations of wild <i>P. yedoensis</i> population.....	58
Figure 19. Haplotype analysis of each wild <i>P. yedoensis</i> population. Pie charts were presented based on percentages of maternal genotype.....	59
Figure 20. Reconstruction of genetic structure of wild <i>P. yedoensis</i> in Jeju island.....	61
Figure 21. Examples of PCR amplification of PrCOS markers using <i>Prunus</i> populations in Kwaneumsa temple area and Bonggae-dong area.....	65
Figure 22. A Maximum Likelihood tree of <i>Prunus</i> species in Kwaneumsa temple and Bonggae-dong populations based on 20 PrCOS markers.....	67
Figure 23. Characterization of genetic structure of wild <i>Prunus</i> species in Kwaneumsa temple area and Bonggae-dong area based on population structure analysis using 20 PrCOS markers.....	68

표 목차

Table 1. Conditions for PCR amplification of PrCOS markers.....	12
Table 2. List of the wild <i>P. yedoensis</i> plants that are registered as the Natural Monuments in Jeju island.....	18
Table 3. List of wild <i>P. yedoensis</i> plants in Jeju island used in this study.....	20
Table 4. List of <i>Prunus</i> species used in this study as closely relatives of <i>P. yedoensis</i> in Jeju island.....	24
Table 5. Summarized statistics of 156 primer sets designed from candidate dimorphic <i>P. yedoensis</i> PrCOS genes.....	34
Table 6. Summarized statistics of 38 primer sets designed from dimorphic PrCOS markers.....	44
Table 7. Summary of 38 primer sets for dimorphic PrCOS markers in <i>P. yedoensis</i>	45
Table 8. Summary of 20 primer sets for dimorphic PrCOS markers used in phylogenetic and population structure analysis of <i>P. yedoensis</i>	48
Table 9. Genotyping matrix of wild <i>P. yedoensis</i> plants using 20 PrCOS markers.....	51
Table 10. Determination of optimum K value for <i>P. yedoensis</i> populations using Evanno method.....	57
Table 11. List of populations and individual wild <i>Prunus</i> species collected from Kwaneumsa temple region.....	63
Table 12. List of populations and individual wild <i>Prunus</i> species collected from Bonggae-dong region.....	64
Table 13. Sequence identity of the paternal and maternal allele between <i>P. yedoensis</i> and its closely relative species.....	71

제주도 자생 왕벚나무의 계통 및 유전구조 분석

조아라

명지대학교 대학원 생명과학정보학과

지도교수 문정환

왕벚나무(*Prunus yedoensis* Mat.)는 전 세계에서 유일하게 제주도에서 자생지가 발견되었지만 현재까지 왕벚나무 종의 기원과 분류학적 기준에 대한 명확한 근거가 제시되지 않아 많은 논란이 있어 왔다. 본 연구에서는 벚나무속 범용 유전자 분자 마커를 이용하여 왕벚나무의 유전형과 근연종 벚나무류와의 계통분류학적 관계를 분석함으로써 제주도 자생 왕벚나무의 유전구조를 파악하고 왕벚나무의 양친 계통을 추정하고자 하였다. 중국 매실(*P. mume*)과 복숭아(*P. persica*) 그리고 왕벚나무의 유전체 정보를 이용하여 벚나무속 Conserved Ortholog Set 유전자(*Prunus* COS) 1,303개를 발굴하였다. 발굴한 COS 유전자를 이용하여 왕벚나무 이형성 PrCOS의 Insertion/Deletion Polymerase Chain Reaction (InDel PCR) 마커 38개를 개발하였다. 이를 자생 왕벚나무 기념목 21개체와 자생 왕벚나무 52개체 그리고 왕벚나무 근연종 자생 벚나무류 중 올벚나무(*P. pendula* for. *ascendens*) 19개체, 벚나무(*P. jamasakura*) 16개체, 산벚나무(*P. sargentii*) 7개체 그리고 사옥(*P. jamasakura* var. *quelpaertensis*) 3개체에 적용하여 자생 왕벚나무와 근연종 벚나무류를 아가로스 젤 전기영동상의 밴드 패턴으로 구분하였다. 또한 왕벚나무 종 구분 마커 38개 중 20개를 이용하여 자생 왕벚나무의 유전형을 분석한 결과, 자생 왕벚나무와 재배 왕벚나무 78개체 중 70개체(90%)는 모계 추정종인 올벚나무와 부계 후보종인 벚나무, 산벚나무 또는 사옥의 유전형을 50%씩 가지고 있는 이형 접합체였다. 나머지 8개체는 이형 접합체에 모계 혹은 부계의 유전형이 역교배 되어 모계 또는 부계의 유전형을 75%씩 가지고 있었다. 한편 제주도 자생 왕벚나무와 근연종 벚나무류 중 모계 추정종인 올벚나무와 부계 후보종인 벚나무, 산벚나무 그리고 사옥과의 계통을 비교 분석하였다. 그 결과, 왕벚나무와 올벚나무는 각각 하나의 집단으로 묶여 다른 벚나무류와 구분이 되었지만 벚나무와 산벚나무 그리고 사옥은 혼합되어 있는 하나의 집단을 형성하였다. 제주도 자생 왕벚나무의 양친

계통을 추정하기 위하여 자생 왕벚나무와 근연종 벚나무류의 핵 유전체 단편 서열을 이용하여 모계 추정 공통 서열과 부계 추정 공통 서열을 제작한 후 비교 분석하였다. 그 결과, 왕벚나무의 모계 대립 유전자 서열과 올벚나무의 공통 서열의 일치도는 평균 0.99로 매우 높았다. 한편 왕벚나무의 부계 대립 유전자 서열과 부계 후보종 대립 유전자의 서열의 일치도는 벚나무 평균 0.99, 사오은 평균 0.98 그리고 산벚나무 평균 0.93로 벚나무의 서열 일치도가 매우 높았다. 이러한 결과로 볼 때 왕벚나무는 동배수성 잡종(homoploid hybrid)으로서 모계는 올벚나무로 확인하였으며 부계는 후보종 중에서 벚나무가 가장 유력한 것으로 추정하였다. 또한 자생 왕벚나무에 부계와 모계의 유전형이 역교배를 통하여 자연적으로 유입되고 있는 것으로 사료된다.

주제어

자생 왕벚나무, COS, 길이 다형성, 분자표지, 유전구조, 계통분석, 유전체 서열 일치도, 동배수성 잡종

제 1 장 서 론

제 1 절 연구의 배경 및 필요성

벚나무류는 장미과(Rosaceae) 벚나무속(*Prunus* L.)에 속하는 현화식물로서 북반구 온대 및 아열대에 약 250여종이 분포하고 있다(Cho et al. 2014; Gray 1867; Lee and Wen 2001; Potter et al. 2007; Kim et al. 1998). 한국에는 재배종을 포함하여 약 22종의 벚나무류가 분포하고 있으며 특히 제주도에는 왕벚나무(*P. yedoensis*), 벚나무(*P. jamasakura*), 산벚나무(*P. sargentii*), 사옥(*P. jamasakura* var. *quelpaertensis*), 산개벚지나무(*P. maximowiczii*), 개벚나무(*P. leveilleana*) 등 11종이 분포해 있다(Cronquist 1981; Jung et al. 1998; Kim 1985; Kim et al. 1998). 벚나무류 수종들은 주로 일본에서 다양한 품종으로 개발되어 전 세계적으로 식재되고 있다. 벚나무류의 많은 품종들이 관상수, 조경수 및 과수 등으로 활용되고 있어 경제적으로도 그 가치가 높다(Cho et al. 2016; Harn et al. 1977; Jung et al. 1998; Kim et al. 1998). 매년 한국과 일본 그리고 미국에서는 벚꽃 축제가 열리고 있으며 특히, 미국 Washington D.C의 벚꽃 축제는 세계적인 꽃 축제의 하나로 자리 잡았고 매년 축제를 보러 많은 관광객이 방문하고 있다. 이처럼 벚나무류는 문화적·사회적으로도 귀중한 수목 자원 중 하나이다.

제주도 자생 왕벚나무(*P. yedoensis* Mat.)는 벚나무아속(Subgenus *Cerasus*) 벚나무 절(Section *Pseudocerasus*)에 속하는 관목이다(Kim et al. 1998). 왕벚나무는 형태학적으로 꽃자루(peduncle)와 암술대(style)에 텔이 있고 꽃받침통(calyx tube)의 중간 부분이 부푼 원통형인 특징을 기준으로 다른 벚나무류 종과 구분된다(Cho et al. 2016; Harn et al. 1977; Jung et al. 1998; Kim et al. 1998; Koidzumi 1932; Roh et al. 2007; Takenaka 1963). 그러나 이러한 특성은 일본의 재배 왕벚나무인 소메이 요시노 벚나무(*Prunus* × *yedoensis*)와 매우 비슷하다. 왕벚나무는 1908년에 프랑스 선교사였던 Taquet 신부가 처음으로 한라산에서 채집하여 유럽에 보고하였다(Matsumura 1901; Kim 1985; Kim et al. 1998; Cho et al. 2016). 이후 1912년에 베를린 대학교의 Koehne

교수가 Taquet 신부의 표본을 근거로 *P. yedoensis* Matsumura var. *nudiflora* Koehne라는 학명으로 발표하였다(Koehne 1912; 이창복 1996). 한편, 1916년에는 일본 Nakai 교수가 조선삼림식물편 5집에서 이를 *P. yedoensis* Matsumura의 이명으로 처리하였다(Cho et al. 2016).

왕벚나무는 세계에서 유일하게 제주도에서 자생지가 발견되었다. 제주도에 자생하고 있는 왕벚나무는 한라산 사면 해발 450 m - 850 m에 분포해 있으며 개체의 직경과 수고 연구를 통해 어린 나무부터 약 200년생으로 추정되는 자생 개체까지 발견되었다(Kim et al. 1998). 이러한 이유로 제주도 왕벚나무 자생지내 개체들을 국가 또는 지방의 기념목으로 지정하여 보호하고 있다. 주요 기념목들은 천연 기념물 제 159호 제주 봉개동 왕벚나무 자생지, 천연기념물 제 156호 제주 신예리 왕벚나무 자생지, 제주 시도 기념물 제 51호 제주 관음사 왕벚나무 자생지 등에 분포하고 있다. 특히 2015년 국립산림과학원, 한국식물분류학회, 제주특별자치도는 공동으로 왕벚나무 유전자원을 관리하고 자원화하기 위하여 생육상태와 나무의 모양, 개화 형질 등이 우수하여 학술적, 자원적 가치가 높은 관음사 야영장 지구 기념목(높이 15 m, 밑동 둘레 3.45 m)을 ‘왕벚나무 기준 어미나무’로 지정하여 보호하고 있다(Figure 1). 그러나 왕벚나무 자생지가 전 세계적으로 제주도에서 유일하게 발견되고 있지만 왕벚나무 종의 기원과 분류학적 기준에 대해서는 아직까지 명확하게 제시된 바가 없어 많은 논란이 있어 왔다.

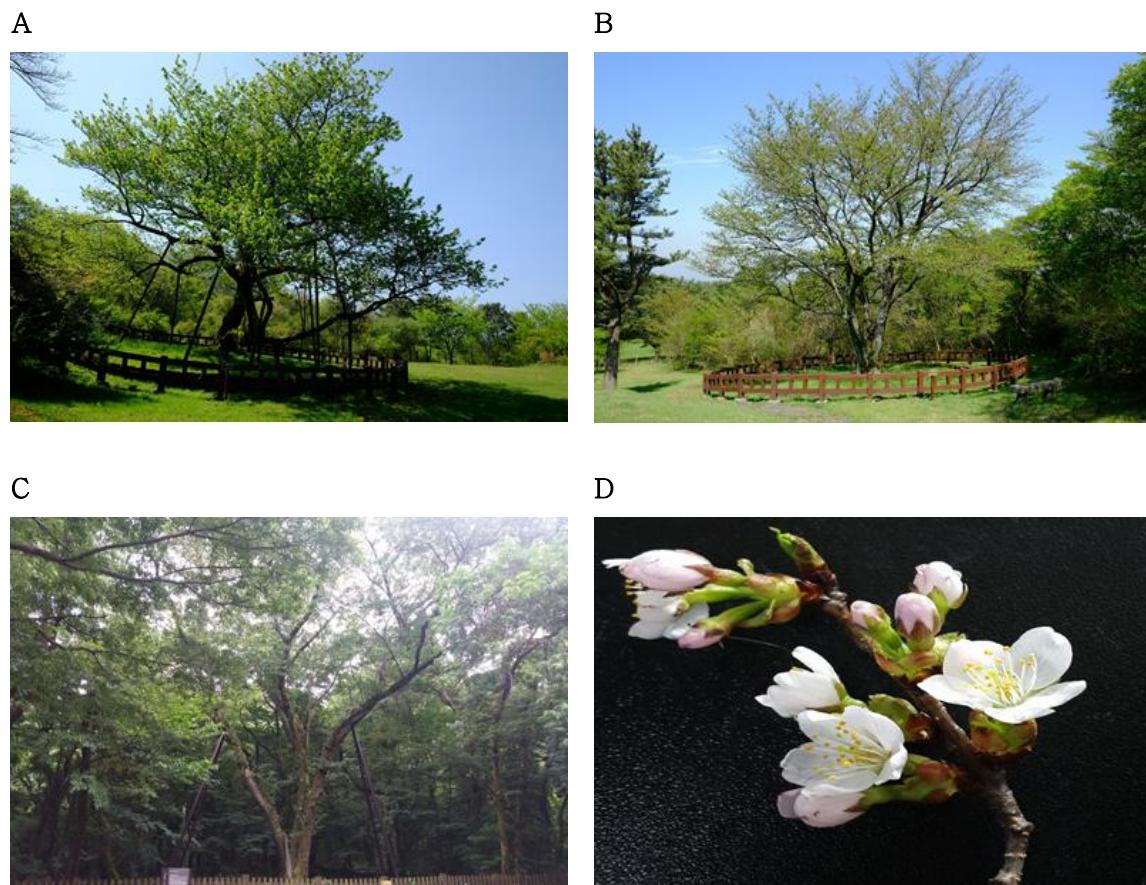


Figure 1. Photographs of native wild *P. yedoensis* in Jeju island, Korea

A and B. Wild *P. yedoensis* trees in the area of the National Monument No. 159 in Bonggae-dong, Jeju island. Plant identification number were designated as 159-1 (A) and 159-2 (B). **C.** Wild *P. yedoensis* in the area of the Kwaneumsa temple, Jeju island. This plant was designated as the standard reference plant of wild *P. yedoensis* by Jeju Special Self-Governing Province on 2015. Plant identification number is PY128. **D.** Blossom of the wild *P. yedoensis* 159-2 tree.

왕벚나무 종에 대한 기원은 벚나무류의 종간 잡종에 의하여 형성된 교잡종이라는 ‘왕벚나무 잡종기원설’이 가장 유력하게 받아 들여 지고 있다. 1916년 Wilson 교수가 처음으로 참벚나무(*P. lannesiana*)와 올벚나무(*P. pendula* for. *ascendens*)의 교잡실험을 통하여 왕벚나무는 두 종의 잡종이며 발생지가 일본의 이즈 반도(Idz peninsula)일 것으로 추정하였다(Wilson 1916). 이후 1964년에 한라산에 자생하는 벚나무류의 개화기와 분포지를 조사하고, 외부 형태학적 형질과 양친 교잡실험을 비교하여 왕벚나무는 산벚나무(*P. sargentii*)와 올벚나무(*P. pendula* for. *ascendens*)의 종간 잡종에 의하여 형성된 교잡종이라고 추정하였다(한창열 1964; 1965). 뿐만 아니라 일본에서 도입된 재배 왕벚나무와 산벚나무 그리고 한라산에 자생하는 올벚나무를 대상으로 6가지 동위효소를 전기영동법으로 비교하여 재배 왕벚나무가 산벚나무와 올벚나무 두 종의 잡종일 가능성이 제시되었다(한창열 등. 1977). 그러나 재배 왕벚나무와 자생 왕벚나무의 Randomly Amplified Polymorphic DNA (RAPD)를 분석한 결과 재배 왕벚나무와 자생 왕벚나무는 기원이 서로 다른 종이라고 추정되기도 하였다(Jeong et al. 1998). 이처럼 자생 왕벚나무는 벚나무류의 종간 잡종에 의하여 형성된 교잡종이라고 추정하는 연구가 많은 학자들에 의해 다양하게 제시되고 있지만 아직까지 왕벚나무의 종의 기원에 대한 정확한 근거는 밝혀지지 않았다.

왕벚나무의 종간 잡종에 대한 연구뿐만 아니라 부계와 모계를 추적하기 위한 연구들도 계속되고 있다. 왕벚나무에 대하여 화분학적 연구와 RAPD, Inter-Simple Sequence Repeat (ISSR), Internal Transcribed Spacer of ribosomal RNA genes (rRNA ITS) 그리고 엽록체 서열의 Intergenic spacer와 microsatellites 등을 이용한 분자유전학적인 연구들이 수행되었다. 화분학적 연구 및 PCR-RAPD를 바탕으로 한 연구에서는 자생 왕벚나무와 근연종 벚나무류의 화분 크기와 표면 무늬 그리고 유연관계를 비교 분석하여 자생 왕벚나무와 사옥이 가장 가까운 유연관계를 가진다고 밝혔다(Jung et al. 1998; Park et al. 1984). 반면 ISSR과 엽록체 서열에 의하면 자생 왕벚나무의 부계는 벚나무 또는 산벚나무로 추정되며 왕벚나무와 높은 유연관계를 가진다고 밝혔다(Roh et al. 2007). 한편, 왕벚나무와 올벚나무의 엽록체 DNA를 비교한 결과 두 종의 엽록체 유전체가 동일한 것으로 나타났다(Roh et al. 2007). 이러한 연구들을 종합하면 제주도 자생 왕벚나무의 모계는 올벚나무(*P. pendula* for. *ascendens*) 계통으로, 부계는

벚나무(*P. jamasakura*), 산벚나무(*P. sargentii*) 또는 사옥(*P. jamasakura* var. *quelpaertensis*)등 다수의 계통일 것으로 추정된다. 그러나 아직까지 제주 자생 왕벚나무의 부계를 정확하게 추정할 수 없으며 왕벚나무와 근연 벚나무류와의 계통 분류학적 관계도 명확하게 밝혀지지 않았다. 특히, 기존 연구에서 사용된 rRNA ITS, ISSR나 엽록체 서열 같은 변이가 적은 바코드 유형의 마커로서 근연종간을 구분할 수 있는 분류학적 해상도가 부족하다. 또한 대부분의 연구에서 사용된 벚나무류 수목의 개체수가 적거나 분류군이 한정되어 있어 자생 왕벚나무의 고유한 유전적 특성을 분석하는데 한계가 있으며 제주 자생 왕벚나무와 근연 벚나무간 계통을 비교 분석할 수 없었다(Cho et al. 2014; Iketani et al. 2007; Jeong et al. 1998; Jung et al. 2002; Roh et al. 2007). 따라서 왕벚나무의 고유한 유전적 특성을 분석하기 위하여 진화적 변화 과정을 반영하고 있는 유전체 서열 정보를 기반으로 한 분자 마커를 개발해야 한다.

다양한 식물 종에 대한 유전체 연구가 진행됨에 따라 최근에는 동일 조상에서 기원한 단일 진화 계통에 속하는 종들 간에 진화적으로 보전된 구조를 가지면서 단일 카페로 존재하는 유전자 세트인 Conserved Ortholog Sets (COS)를 활용한 연구가 활발하다. COS 유전자들은 근연종간에 유전자 구조와 염기서열이 높은 수준으로 보전되어 있어서 근연종간에 범용적으로 사용할 수 있어 재현성이 높고 직관적인 해석이 가능한 장점이 있다. 따라서 COS 유전자를 활용하면 종내 또는 종간 분석이 용이한 유전자 기반 분자 마커를 개발할 수 있다(Doganlar et al. 2002; Fulton et al. 2002; Krutovsky et al. 2006; Kuhn et al. 2012; Quraishi et al. 2009; Wu et al. 2006). 이러한 이유로 COS 분자 마커는 근연 분류군에 속하는 다양한 종의 진화적 계통 분석에 적합한 마커로 평가되어 비교유전체 연구에 핵심적으로 활용되고 있다. 예를 들어, 벼과(Quraishi et al. 2009), 콩과(Choi et al. 2004), 가지과(Doganlar et al. 2002)등에서는 모델식물과 작물의 유전체 서열을 비교하여 COS 마커를 개발하고, 이를 비교유전체 연구, 유전지도 작성, 유용 유전자 클로닝 등에 활용하고 있다. 임목의 유전 분석에서도 포플러(Brunner et al. 2004)와 소나무(Echt and Nelson 1997)등 유전체가 해독된 수종을 중심으로 COS 분자 마커가 개발되어 활용되고 있다. 소나무과에서 가계 분석에 COS 마커를 적용한 경우, 미토콘드리아나 엽록체 DNA 마커에 비해 유전자 정보량이 많고 재현성과 범용성이 뛰어난 장점이 있는 것으로 평가되고 있다(Liewlaksaneeyanawin et al.

2009). COS 유전자 기반 분자 마커를 개발하기 위해서는 연구 대상 분류군의 복수의 종에서 핵 유전체 정보가 필요하다. 2016년 현재 장미과 벚나무속에는 중국 매실(*P. mume*)과 복숭아(*P. persica*)의 전체 유전체의 초안 정보가 각각 2012년과 2013년에 보고되어 있어 장미과 벚나무속 COS (*Prunus* COS, PrCOS) 유전자 동정이 가능하다. 또한 명지대학교 식물분자유전학 연구실에서 자체적으로 해독한 왕벚나무의 핵 유전체 정보를 이용하여 왕벚나무 및 근연 벚나무류에서 범용적으로 적용할 수 있는 COS 유전자 기반 분자 마커를 개발할 수 있고 이를 이용하여 제주도 자생 왕벚나무의 개체 간 유전형을 분석하고 집단 내의 유연관계를 비교할 수 있다.

제 2 절 연구의 목표

본 연구에서는 제주도 자생 왕벚나무(*Prunus yedoensis*)와 올벚나무(*P. pendula* for. *ascendens*), 벚나무(*P. jamasakura*), 산벚나무(*P. sargentii*) 그리고 사옥(*P. jamasakura* var. *quelpaertensis*)등 제주도 고유 자생 벚나무류의 유전구조를 분석하고 각 집단별 계통 관계를 파악하고자 하였다. 이를 위하여 이미 해독이 완료된 중국 매실과 복숭아의 유전체 정보와 자체적으로 확보하고 있는 왕벚나무의 핵 유전체 정보들을 바탕으로 제주도 벚나무류 집단에 적용 가능한 유전자 기반 COS 분자 마커를 개발하고 제주도 한라산 사면에 분포하고 있는 자생 벚나무류 집단에 적용하였다. 최종적으로 자생 왕벚나무와 근연 벚나무류에 유전자 기반 COS 분자 마커를 적용하여 유전구조 및 계통 관계를 분석하고 부계와 모계를 추정함으로써 왕벚나무 종 보전을 위한 유전적 지표를 제시하는 것을 연구의 목표로 하였다.

제 2 장 연구 재료 및 방법

제 1 절 연구 재료

1. 식물 재료

가. 제주도 자생 왕벚나무 기념목 개체

제주도에는 왕벚나무 자생지가 산재해 있으며 자생지내 개체들은 국가 또는 지방 기념물로 지정되어 보호되고 있다. 본 연구에서는 천연 기념물 제 159호 제주 봉개동 왕벚나무 자생지, 천연기념물 제 156호 제주 신예리 왕벚나무 자생지, 제주 시도 기념물 제 51호 제주 관음사 왕벚나무 자생지의 개체들과 제주시 향토유산 기념물 제 3호 제주도 자생 왕벚나무 지정 어미나무를 연구 재료로 설정하였다. 각 개체들의 어린잎과 잎눈 등을 채집하여 사용하였다.

나. 제주도 자생 왕벚나무

제주도 자생 왕벚나무는 국립산림과학원 난대·아열대산림연구소에서 50개체의 생체 재료를 제공받았으며 자체적으로 18개체를 채집하였다. 채집한 생체 재료 중 수목의 수령과 생체 재료의 상태를 고려하여 연구소에서 제공한 50개체와 자체적으로 채집한 2개체를 포함한 총 52개체를 제주도 자생 왕벚나무 집단으로 구성하여 사용하였다.

다. 제주도 자생 왕벚나무 근연종 벚나무류

제주도 자생 왕벚나무의 근연종 벚나무류는 국립산림과학원 난대·아열대산림연구소에서 올벚나무(*P. pendula* for. *ascendens*) 9개체를 제공받았으며 자체적으로 근연종 벚나무류 45개체를 채집하였다. 채집한 개체들은 모두 국립산림과학원 난대·아열대산림연구소에서 종 동정을 검증받았다. 근연종 벚나무류로는 모계 추정종인 올벚나무(*P. pendula* for. *ascendens*)와 부계 추정종인 벚나무(*P. jamasakura*), 산벚나무(*P. sargentii*), 사옥(*P. jamasakura* var. *quelpaertensis*)을 포함하여 총 6종을 사용하였다.

연구에 사용한 벚나무류는 올벚나무 19개체, 벚나무 16개체, 산벚나무 7개체, 사옥 3개체이며 각 개체의 어린잎을 채집하여 사용하였다.

2. Genomic DNA 추출

채집된 식물 재료는 액체 질소에 담가 얼린 후 -80°C 의 deep freezer에서 보관하였다. 채집한 자생 왕벚나무와 근연종 벚나무류의 생체 재료(건조된 잎과 냉동 보관한 잎눈) 중 일부(100 mg - 1 g)를 이용하여 genomic DNA (gDNA)를 추출하였다. 식물 재료는 액체 질소로 마쇄한 후 CTAB buffer (2% hexadecyltrimethylammonium bromide, 1.4 M NaCl, 0.2% 2- β mercaptoethanol, 20 mM EDTA, 100 mM Tris-HCl, 2% Polyvinylpyrrolidone K-40)를 사용하여 gDNA를 추출하고 RNase 처리 후 Isopropanol로 -20°C 에서 8시간 동안 침전시켜 추출하였다(Khanuja et al. 1999; Murray and Thompson 1980). 시료 당 평균 140 μg 의 gDNA를 추출하였으며 최종 농도 50 ng/ μL 로 희석하여 이후 분석에 사용하였다.

제 2 절 연구 방법

1. 왕벚나무 유전자 서열 및 벚나무속 COS 유전자 선발

가. 왕벚나무 유전자 서열

본 연구에서는 명지대학교 식물분자유전학연구실에서 확보하고 있는 왕벚나무 전사체 조립 서열(unigene)과 핵 유전체 서열(Illumina short read와 PacBio RS II reads)을 사용하였다. Illumina short read를 이용한 핵 유전체 서열 조립은 3가지 방법을 이용하여 생성한 각각의 조립 서열(scaffold)을 사용하였다. 왕벚나무 핵 유전체 조립 서열은 각각 Allpaths-LG, SOAPdenovo, Platanus assembler를 이용하여 조립되었다. Allpaths-LG를 이용하여 조립된 단편은 14,011개, 총 길이 417,639,250 bp, 평균 길이

29,807 bp, N50 73,233 bp이다. SOAPdenovo를 이용하여 조립된 단편은 871,246개, 총 길이 703,896,073 bp, 평균 길이 807 bp, N50 290,781 bp이다. Platanus를 이용하여 조립된 단편은 17,551개 총 길이 32,803,7571 bp, 평균 길이 18,690 bp, N50 266,739 bp이다.

한편, PacBio RS II reads를 이용한 왕벚나무 핵 유전체 초안은 Falcon assembler를 이용하여 조립된 서열을 사용하였다. Falcon을 이용하여 조립된 단편은 4,016개, 총 길이 319,209,792 bp, 평균 길이 79,484 bp, N50 145,140 bp이다.

나. 벚나무속(*Prunus*) 속 COS 유전자 선발

COS 유전자를 발굴하기 위하여 유전체 해독이 완료된 중국 매실과 복숭아의 유전체 초안에서 유전자 서열 정보를 추출하였다. 중국 매실(Zhang et al. 2012)과 복숭아(Verde et al. 2013)의 유전체 서열은 NCBI (www.ncbi.nlm.nih.gov/genome)에서 내려 받아 사용하였다.

벚나무속 COS 유전자(PrCOS)를 선발하기 위하여 먼저 중국 매실과 복숭아의 유전체 초안에서 유전자 서열 정보를 추출하였다. 중국 매실과 복숭아의 유전자 모델에 대하여 E-value cutoff 1E⁻¹⁰의 조건에서 self-BLASTP를 수행하여 각각의 단일 카페 유전자를 구성하였다. 이후 E-value 1E⁻¹⁰, query coverage 50% 이상의 조건에서 중국 매실과 복숭아의 coding sequence (CDS)의 아미노산 서열에 대하여 상보적 상동성 검색(reciprocal-BLASTP)을 수행하였다.

2. COS 분자 마커 개발 및 선발

가. COS 분자 마커 개발 및 선발

왕벚나무의 모계와 부계를 동시에 반영하는 이형성 PrCOS 분자 마커는 왕벚나무 전사체 서열과 핵 유전체 단편 서열 그리고 왕벚나무 유전자 모델을 이용하여 설계하였다.

우선 왕벚나무 전사체 서열을 중국 매실의 PrCOS 유전자 모델에 매핑하고 Spidey 프로그램(Wheelan et al. 2001)을 이용하여 왕벚나무 전사체 서열과 PrCOS 유전자 서열을 정렬하였다. PrCOS 유전자와 정렬된 왕벚나무 전사체 서열 영역을 엑손, 정렬되지 않은

영역을 인트론이라고 추정하였다. 이후 중국 매실 염색체 8개상의 PrCOS 유전자의 분포 양상을 기준으로 무작위적인 엑손 영역을 선발하고 프라이머 세트를 설계하였다.

왕벚나무 핵 유전체 단편 서열의 경우 각각 3가지 방법으로 조립된 서열을 모두 이용하였다. 각각 조립된 핵 유전체 단편 서열을 중국 매실의 PrCOS 유전자의 아미노산 서열에 대하여 TBLASTN (E-value cutoff 1E⁻¹⁰)으로 분석한 후 선발된 서열을 ClustalW2 (www.ebi.ac.uk, Larkin et al. 2007)로 정렬하였다. 이후 하나의 PrCOS 유전자 모델에 두 개의 왕벚나무 핵 유전체 단편 서열이 길이 다형성을 가지며 대응되는 영역의 서열을 추출하였다. 중국 매실 PrCOS 유전자의 아미노산 서열과 정렬되는 왕벚나무 핵 유전체 서열 영역을 엑손, 정렬되지 않는 영역을 인트론이라고 추정하였다. 길이 다형성을 보이는 인트론의 양쪽 측면 엑손 영역에 Polymerase Chain Reaction (PCR) 프라이머를 설계하였다.

왕벚나무 핵 유전체 초안으로부터 예측한 유전자 모델의 경우, PrCOS 유전자와 왕벚나무 유전자 모델을 ClustalW2로 정렬한 후 왕벚나무 PrCOS 유전자의 이형성 서열 영역을 탐색하였다. 왕벚나무 이형성 PrCOS 유전자 서열은 엑손과 인트론 구역을 구분하여 길이 다형성을 나타내는 인트론 부위에서 Insertion/Deletion (InDel) 영역만 선택적으로 발굴하였다.

왕벚나무 PrCOS 유전자의 InDel 영역의 서열을 주형으로 하는 프라이머는 FastPCR Professional v6.0 프로그램(Kalendar et al. 2009)을 이용하여 길이 범위 20-25 bp, Tm(°C) 범위 51°C - 65°C, GC(%) 범위 45 - 65%의 조건으로 설계하였다. 설계된 프라이머 세트 중에서 증폭 산물의 길이를 고려하여 InDel 영역의 길이가 예상되는 PCR 결과 크기의 10% 이상이 되는 프라이머 세트 156개를 선발하였다. 선발된 156개의 프라이머 세트는 왕벚나무 핵 유전체 서열을 대상으로 PCR 증폭 여부, 예상되는 증폭 산물의 길이 등을 FastPCR Professional v6.0 프로그램의 *in silico* PCR 방법으로 예측하여 검증하였다.

나. PCR 증폭 및 전기영동

PCR은 Bio-Rad Thermal Cycler (C1000)를 사용하였으며, PCR 수행 조건은 95°C에서 3분간 predenaturation 후 95°C에서 30초간 denaturation, 58°C에서 1분 30초간 annealing, 72°C에서 1분간 extension하는 과정을 35 cycle로 반복 시행한 후 72°C에서 최종적으로 5분간 유지하였다(Table 1). 증폭된 PCR 산물은 아가로스 젤 전기영동으로 확인하였다. 전기영동은 2 - 3% MP-Bio 아가로스 젤(QA-Agarose TM, USA)을 사용하여 0.5X TBE buffer에서 150 V로 2-3 시간 진행하였다. 전기영동시 사용한 마커는 100 bp DNA ladder(LPS solution, Korea)를 사용하였으며 전기영동이 끝난 아가로스 젤은 DNA 증폭 밴드의 UV 형광사진을 촬영하였다.

Table 1. Conditions for PCR amplification of PrCOS markers

Reaction set up for PCR	
DNA (50 ng/μL)	1 μL
primer (1 μM)	1 μL
dNTP (2.5 mM, ea.)	2 μL
PCR buffer (10X)	2.5 μL
DW	18.3 μL
Taq polymerase (1 unit/μL)	0.2 μL
Total	25 μL

Temperature cycling progress	
95°C	3 min
95°C	30 sec
58°C	1 min 30 sec
72°C	1 min
72°C	5 min
12°C	∞

3. 계통 및 유전분석

가. 개체별 유전형 분석

제주도 자생 왕벚나무와 근연종 벚나무류(올벚나무, 벚나무, 산벚나무, 사옥)에 선발된 PrCOS 분자 마커를 적용하여 PCR과 전기영동을 수행하였다. 이후 각각의 allele에 대한 전기영동 밴드 패턴을 육안으로 판독하였으며 마이크로소프트사의 엑셀 프로그램으로 유전형 매트릭스를 작성했다. 판독한 밴드 패턴은 codominant data 형식으로 재구성한 후 개체별 계통 분석과 집단별 유전구조 분석에 사용하였다.

나. 계통 분석

왕벚나무와 근연종 벚나무 집단의 각 개체별 계통을 비교하기 위하여 MEGA7의 Maximum Likelihood (ML) 분석을 사용하였으며, Maximum Composite Likelihood (MCL) 옵션을 사용한 Tamura-Nei mode 방식으로 계통수를 작성하였다(Kumar et al. 2016). 계통 분석의 신뢰도를 높이기 위하여 형질 재추출 과정을 반복하는 부트스트랩(bootstrap) 분석을 1,000회 지정하여 반복 수행하였다.

다. 유전구조 분석

왕벚나무와 근연종 벚나무 집단의 각 개체별 모계 및 부계 유전형의 유전적 조성을 파악하기 위하여 STRUCTURE v2.3.4 프로그램(Pritchard et al. 2000)을 이용하였다. K값을 1부터 15까지 연속적으로 설정하였으며 각 K값 당 10회 반복 분석하였다. 이때, 분석 조건은 Length of Burnon Perion 100,000, Number of MCMC Reps after Burnin 5000, Frequency of Metropolis update for Q 10로 설정하였다. 10번의 반복 수행 결과는 Evanno method를 이용한 ad hoc criterion (ΔK)로 분석하여 최종적으로 K 값을 결정하였다(Falush et al. 2003; Hubisz et al. 2009).

4. PrCOS 분자 마커 목표 영역의 서열 분석

가. 왕벚나무 및 근연종 벚나무류의 공통서열(consensus sequence) 추출

PrCOS 마커의 목표 영역을 탐색하고 왕벚나무 부계를 추정하기 위하여 제주도 자생 왕벚나무 기념목 지정 4개체(Identification No. 159-1, 159-2, 159-3, PY128)와 왕벚나무 근연종 벚나무류인 올벚나무(Sample ID: A05), 벚나무(Sample ID: MJ_101), 산벚나무(Sample ID: MJ_111) 그리고 사옥(Sample ID: MJ_163) 각각 1개체의 핵 유전체 Illumina short read 서열을 이용하였다. 각 개체별 핵 유전체 단편서열(read)을 왕벚나무 유전자 모델에 Bowtie2 (Langmead B et al. 2012)를 이용하여 기본 조건으로 매핑하였고 Geneious v7.0 (Kearse M et al. 2012)을 이용하여 각 개체별 핵 유전체 단편서열을 기반으로 한 공통서열(consensus sequence)로 만들어 추출하였다.

나. 서열 일치도 분석

왕벚나무의 모계 대립 유전자 서열은 올벚나무의 공통 서열과 정렬하였으며, 왕벚나무의 부계 대립 유전자 서열은 부계 후보종인 벚나무와 산벚나무 그리고 사옥의 공통서열과 각각 정렬하였다. 왕벚나무와 근연종 벚나무류의 정렬한 공통서열을 모계와 부계로 구분하고 각각 BioEdit (Hall TA 1994) 프로그램의 Sequence Identify Matrix를 이용하여 각 개체별 서열 일치도를 비교하였다.

제 3 장 연구 결과 및 고찰

제 1 절 제주도 자생 왕벚나무 및 근연종 채집 및 집단 구성

채집한 제주도 자생 왕벚나무는 총 73개체로서 기념목 지정 21개체와 야생종 52개체를 채집하였다. 또한 근연종 벚나무류는 총 45개체를 채집하였다. 채집한 왕벚나무와 근연종 벚나무류 개체의 채집지는 Figure 2에 나타내었다.

1. 제주도 자생 왕벚나무 기념목 지정 개체

제주도 자생 왕벚나무 기념목 지정 개체는 총 21개체를 사용하였다. 천연 기념물 제 159호 제주 봉개동 왕벚나무 자생지의 3개체, 천연기념물 제 156호 제주 신예리 왕벚나무 자생지의 14개체, 제주 시도 기념물 제 51호 제주 관음사 왕벚나무 자생지의 3개체, 제주시 향토유산 기념물 제 3호 제주 왕벚나무 지정 어미목을 연구 재료로 설정하였으며 각 개체의 어린잎과 잎눈 등을 실험재료로 사용하였다. 지정 기념목의 경우 개체가 보존되어 있기 때문에 확증표본은 제작하지 않았다. 기념목 개체의 위치 및 특징은 Figure 3과 Table 2에 나타내었다.

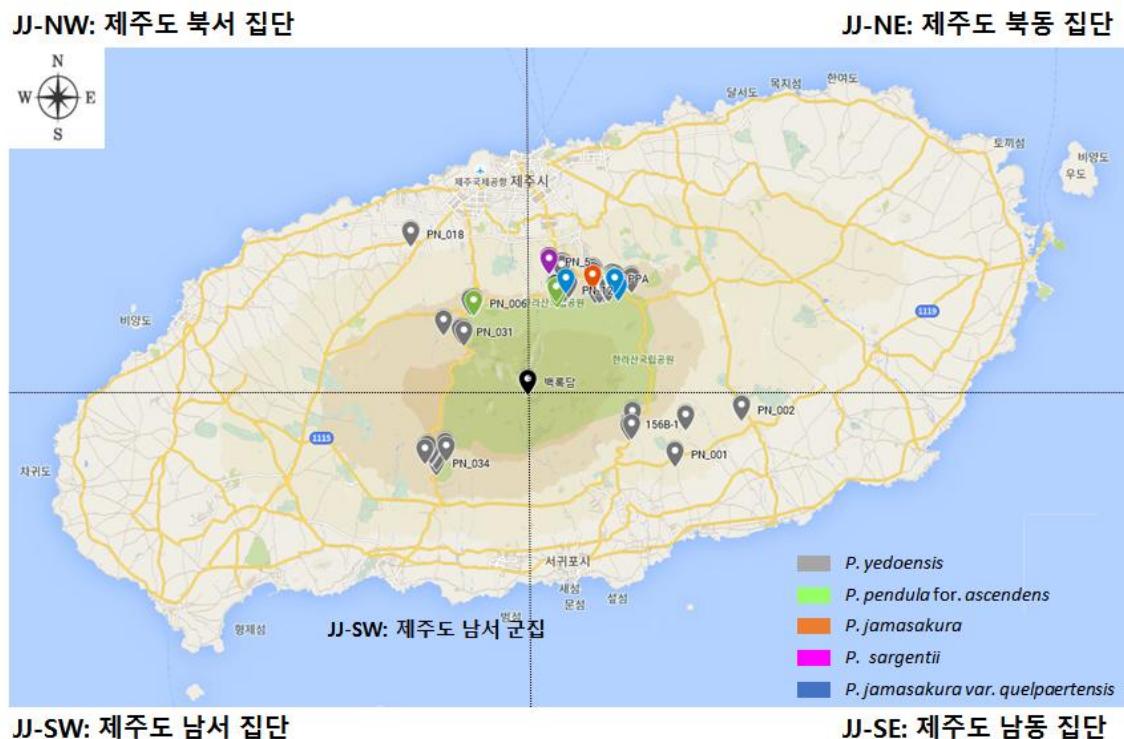


Figure 2. Composite geographical locations of wild *P. yedoensis* and its closely relative species plants used in this study

P. yedoensis and its closely relative species plants are marked according to GPS coordinates which were designated by the National Institute of Forest Science.

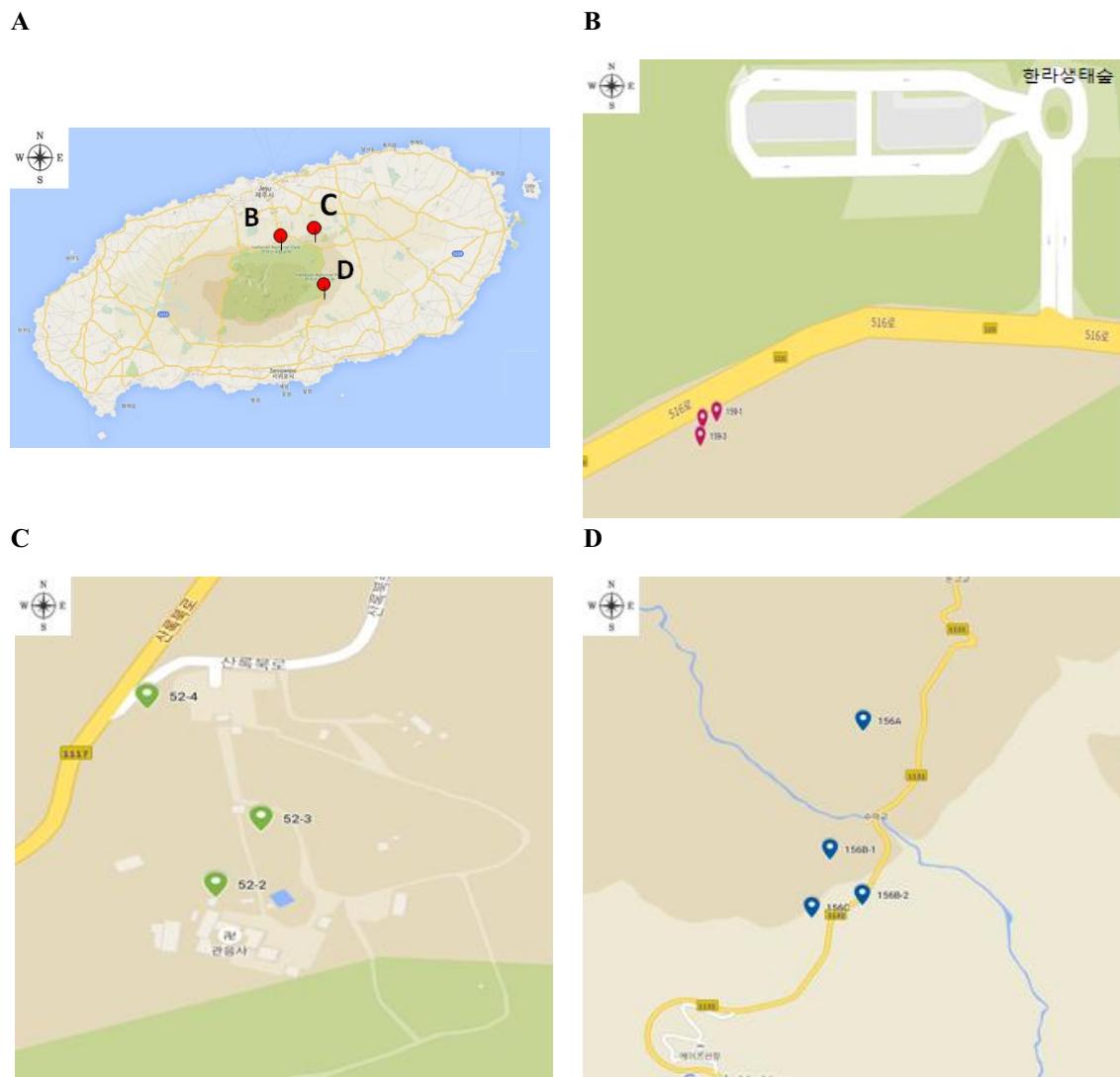


Figure 3. Locations of the wild *P. yedoensis* plants registered as the Natural Monument in Jeju island, Korea

A. Locations of 3 populations of wild *P. yedoensis* plants that are registered as the Natural monuments and Province monuments in Jeju island; **B.** The Natural Monument trees (159-1, 159-2, 159-3) located at Bonggae-dong, Jeju island; **C.** The Jeju provincial monument trees (51-2, 51-3, 51-4) located at Kwaneumsa temple, Jeju island **D.** The Natural Monument trees (156-1, 156-A, 156-B, 156-C) located at Sinyeri, Jeju island

Table 2. List of the wild *P. yedoensis* plants that are registered as the Natural Monuments in Jeju island

Location ¹⁾	Identification No. ²⁾	Sample ID ³⁾	Feature
Bonggae -dong	159-1	15-Pr001	The biggest plant
	159-2	15-Pr002	The reference tree for genome sequencing
	159-3	15-Pr003	
Kwaneumsa temple	51-2	15-Pr004	
	51-3	15-Pr005	
	51-4	15-Pr006	
	PY128 ⁴⁾	16-Py128	Designated as the standard reference plant of wild <i>P. yedoensis</i>
Sinyeri-A	156-A 가	15-Pr007	
	156-A 나	15-Pr008	
	156-A 다	15-Pr009	
Sinyeri-B	156-1	15-Pr010	
	156-B 가	15-Pr011	
	156-B 나	15-Pr012	
	156-B 다	15-Pr013	
Sinyeri-C	156-C 가	15-Pr014	
	156-C 나	15-Pr015	
	156-C 다	15-Pr016	
	156-C 라	15-Pr017	
	156-C 마	15-Pr018	
	156-C 바	15-Pr019	
	156-C 사	15-Pr020	

¹⁾Populations of wild *P. yedoensis*

²⁾The registered number of 'Natural Monument' and 'Province Monument'

³⁾Sample IDs were given according to an order of collection in this study

⁴⁾Sample Py128 was designated as the stand reference plant of wild '*P. yedoensis*' by Jeju Province and National Institute of Forest Science

2. 제주도 자생 왕벚나무

제주도 자생 왕벚나무는 총 52개체를 4개의 집단으로 구성하여 분석하였다. 52개체는 국립산림과학원 난대아열대산림연구소에서 제공한 50개체와 자체적으로 채집한 2개체로 구성되어 있다. 자생 왕벚나무의 집단은 한라산 백록담을 기준으로 지형적, 지리적 위치를 고려하여 북동/북서/남동/남서 4개의 지역으로 구분하였다. 구분한 4개의 지역을 각각 제주-북동(JJ-NE), 제주-북서(JJ-NW), 제주-남동(JJ-SE), 제주-남서(JJ-SW)로 명명하였다. 이에 따라 왕벚나무 집단은 제주-북동(JJ-NE) 34개체, 제주-북서(JJ-NW) 6개체, 제주-남동(JJ-SE) 4개체, 제주-남서(JJ-SW) 8개체로 구분하였다. 각 집단별 채집된 생체 재료는 확증표본으로 제작하였고 샘플 명(Sample ID)과 확증표본 번호(Voucher No.)는 Table 3에 나타냈으며 채집지 GPS 좌표 정보는 별도로 기록하였다(Table 3). 자생 왕벚나무의 보존을 위하여 채집지 GPS 좌표 정보는 따로 기술하지 않았으며 대신 채집지 GPS 좌표 정보를 기반으로 한 개체의 위치는 Figure 4에 나타내었다.

Table 3. List of wild *P. yedoensis* plants in Jeju island used in this study

Populations	Sample ID ¹⁾	Voucher No. ²⁾
JJ-NE	JJ_NP_4	63557
	JJ_NP_5	63559
	JJ_NP_8	63566
	JJ_NP_9	63567
	JJ_NP_10	63568
	JJ_NP_11	63569
	JJ_NP_12	63570
	JJ_NP_13	63571
	JJ_NP_14	63572
	JJ_NP_16	63574
	JJ_NP_17	63575
	JJ_NP_19	63582
	JJ_NP_20	63583
	JJ_NP_21	63584
	JJ_NP_22	63585
	JJ_NP_23	63586
	JJ_NP_24	63587
	JJ_NP_25	63588
	JJ_NP_26	63579
	JJ_NP_27	63580
	JJ_NP_28	63581
	JJ_NP_29	63597
	JJ_NP_30	63598
	JJ_NP_40	63608
	JJ_NP_41	63609
	JJ_NP_42	63610
	JJ_NP_43	63611
	JJ_NP_44	63612
	JJ_NP_45	63613
	JJ_NP_46	63614
	JJ_NP_47	63615
	JJ_NP_48	63616
	JJ-MJ_127	63401
	JJ-MJ_157	63431

Table 3. (Continued)

Populations	Sample ID ¹⁾	Voucher No. ²⁾
JJ-SE	JJ_NP_1	63530
	JJ_NP_2	63531
	JJ_NP_3	63532
	JJ_NP_15	63573
JJ-SW	JJ_NP_34	63602
	JJ_NP_35	63603
	JJ_NP_36	63604
	JJ_NP_37	63605
	JJ_NP_38	63606
JJ-NW	JJ_NP_39	63607
	JJ_NP_49	63577
	JJ_NP_50	63578
	JJ_NP_6	63561
	JJ_NP_7	63563
	JJ_NP_18	63576
	JJ_NP_31	63599
	JJ_NP_32	63600
	JJ_NP_33	63601

¹⁾Sample IDs were given according to an order of collection in this study. JJ_NP means JeJu Natural Population

²⁾Vouchers were deposited at the herbarium of Warm-Temperature and Subtropical Forest Research Center, NIFS

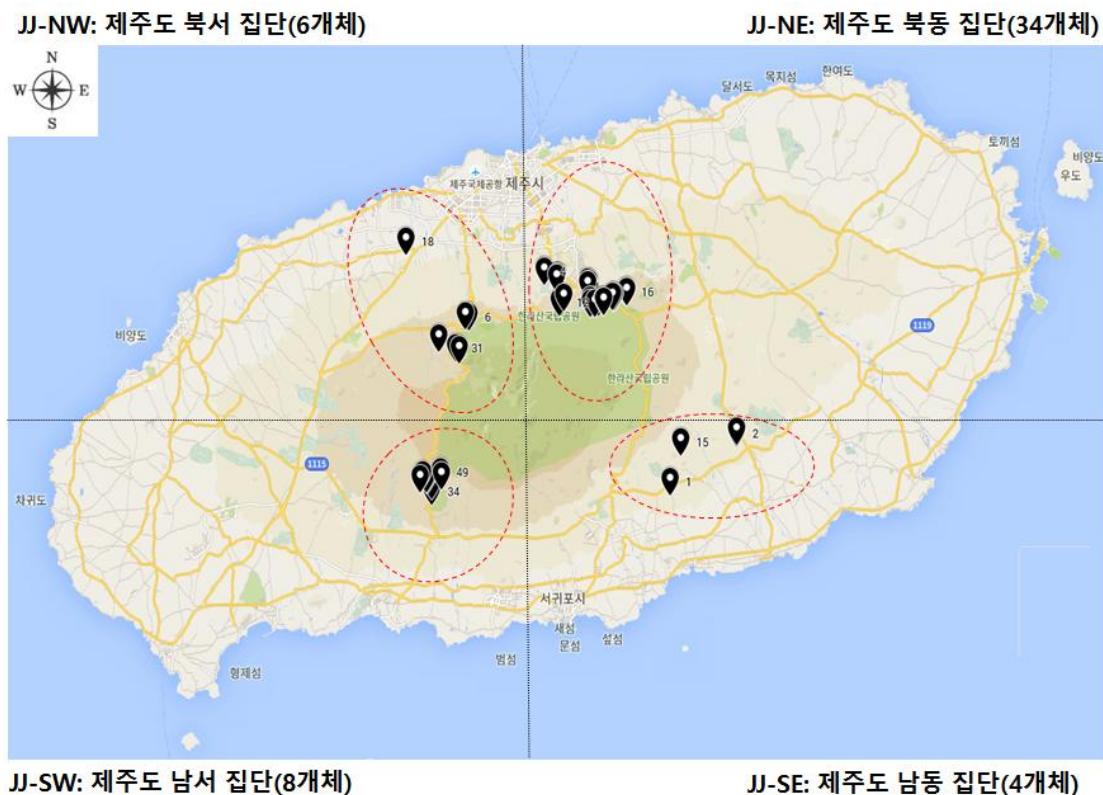


Figure 4. Geographical locations of the wild *P. yedoensis* populations in Jeju island

Wild *P. yedoensis* plants divided into 4 poulations North-East (JJ-NE), North-West (JJ-NW), South-East (JJ-SE) and South-West(JJ-SW) according to the topographic and geographic locations to Mt. Halla. Each population includes 34, 6, 4 and 8 plants, respectively.

3. 제주도 자생 왕벚나무 근연종 벚나무류

국립산림과학원 난대·아열대산림연구소에서 제공받은 올벚나무 9개체는 개발한 PrCOS 마커의 적용성을 평가하기 위하여 사용하였으며 제주도 자생 왕벚나무와 유연 관계를 분석하는 데는 사용하지 않았다. PrCOS 마커의 적용성 평가에 사용한 올벚나무 9개체는 Table 4에 별표(*)로 따로 표시하였다. 제주도 자생 왕벚나무와 유연관계를 분석하기 위한 제주도 자생 왕벚나무 근연 벚나무류는 자체적으로 채집한 45개체를 사용하였다. 왕벚나무와 유연관계를 분석하기 위하여 올벚나무 19개체, 벚나무 16개체, 산벚나무 7개체, 사옥 3개체를 사용하였다. 채집한 개체 중 올벚나무 18개체, 벚나무 15개체, 산벚나무 7개체, 사옥 3개체는 제주-북동(JJ-NE) 지역에 위치해 있으며, 올벚나무 1개체, 벚나무 1개체는 제주-북서(JJ-NW) 지역에 위치해 있다. 채집된 생체 재료는 개체별로 확증표본으로 제작하였으며 샘플 명(Sample ID)과 확증표본 번호(Voucher No.)를 Table 4에서 나타내었다. 자생 벚나무류의 보존을 위하여 채집지 GPS 좌표 정보 역시 따로 기술하지 않았으며 대신 채집지 GPS 좌표 정보를 기반으로 한 개체의 위치는 Figure 5에 나타내었다.

Table 4. List of *Prunus* species used in this study as closely relatives of *P. yedoensis* in Jeju island

Scientific name	Sample ID ¹⁾	Voucher No. ²⁾
	A04*	60570
	A05*	60571
	A06*	60573
	A07*	60574
	A08*	60575
	A09*	60576
	A10*	60578
	A11*	60579
	A12*	60581
	MJ-103	63377
	MJ-104	63378
	MJ-107	63381
	MJ-108	63382
	MJ-113	63387
<i>P. pendula</i> for. <i>ascendens</i>	MJ-116	63390
	MJ-119	63393
	MJ-120	63394
	MJ-121	63395
	MJ-126	63400
	MJ-133	63407
	MJ-138	63412
	MJ-148	63422
	MJ-152	63426
	MJ-154	63428
	MJ-158	63432
	MJ-160	63434
	MJ-162	63436
	NP-008	63562
	MJ-101	63375
	MJ-102	63376
	MJ-106	63380
	MJ-110	63384
	MJ-112	63386
	MJ-117	63391
	MJ-118	63392
	MJ-123	63397
<i>P. jamasakura</i>	MJ-125	63399
	MJ-135	63409
	MJ-150	63424
	MJ-151	63425
	MJ-155	63429
	MJ-161	63435
	NP-010	63564
	NP-029	63592

Table 4. (Continued)

Scientific name	Sample ID ¹⁾	Voucher No. ²⁾
<i>P. sargentii</i>	MJ-153	63427
	MJ-105	63379
	MJ-111	63385
	MJ-114	63388
	MJ-115	63389
<i>P. jamasakura</i> var. <i>quelpaertensis</i>	NP-034	63591
	NP-006	63560
	MJ-163	63437
	MJ-159	63433
	NP-033	63590

¹⁾Sample IDs were given according to an order of collection in this study

²⁾Vouchers were deposited at the herbarium of Warm-Temperature and Subtropical Forest Research Center, NIFS

**P. pendula* for. *ascendens* plants used in test application of PrCOS markers

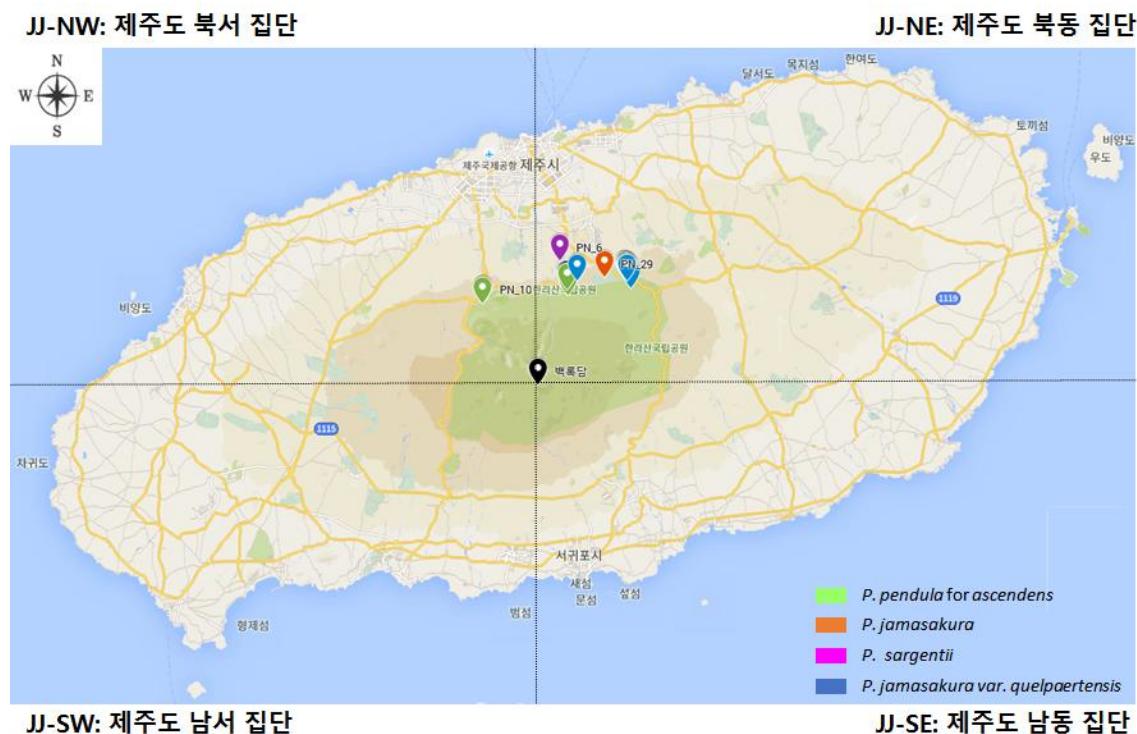


Figure 5. Geographical locations of closely relative *Prunus* populations of wild *P. yedoensis* in Jeju island

JJ-NE population consists of 18 *P. pendula* for. *ascendens*, 15 *P. jamasakura*, 7 *P. sargentii* and 3 *P. jamasakura* var. *quelpaertensis*. JJ-NW consists of 1 *P. pendula* for. *ascendens* and 1 *P. jamasakura*

제 2 절 PrCOS 분자 마커의 개발 및 검증

1. 벚나무속 PrCOS 유전자 선발

PrCOS 유전자를 선발하기 위하여 먼저 중국 매실과 복숭아의 유전체 초안에서 각각 31,390개와 28,702개의 유전자 서열 정보를 추출하였다. 중국 매실과 복숭아의 유전자 모델에 대하여 각각 self-BLASTP (E-value cutoff $1E^{-10}$)로 분석하여 중국 매실의 3,055개, 복숭아의 3,772개의 단일 카피 유전자를 구성하였다. 이후 각각의 단일 카피 유전자를 E-value cutoff $1E^{-10}$, query coverage $>50\%$ 조건에서 정렬하여 중국 매실에는 1,954개, 복숭아에는 1,981개의 벚나무속 유전체간에 보존된 단일 카피 유전자를 선발하였다. 이후 E-value cutoff $1E^{-10}$, query coverage $>50\%$ 조건에서 상보적 상동성 검색 방법(reciprocal BLASTP search)으로 비교하여 벚나무속 COS 후보 유전자 1,801개를 1차 선발하였다(Figure 6).

자생 왕벚나무는 양친교잡실험과 분자유전학적 실험들을 통하여 올벚나무와 다른 벚나무류의 종간 잡종에 의하여 기원되었을 것으로 추정되며, 명지대학교 식물분자유전학 연구실의 연구 결과로 볼 때 전체 유전체의 이형성(heterozygosity) 또한 매우 높은 것으로 나타났다. 왕벚나무의 이형성이 매우 높은 점을 고려하여 중국 매실 및 복숭아의 COS 유전자 모델을 이용하여 왕벚나무의 이형성 COS 유전자 서열을 추출하였다. 우선 1차 선발한 1,801개 PrCOS 후보 유전자와 왕벚나무의 유전자 모델을 E-value cutoff $1E^{-50}$, query coverage $>50\%$ 의 조건에서 상보적 상동성 검색 방법으로 분석하였다. 그 결과, 1차 선발한 PrCOS 유전자 1,801개 중 1,501개의 유전자가 왕벚나무 유전자 2,179개와 정렬되었다. 이때 1,501개의 PrCOS 유전자 중 왕벚나무에 2개 이상의 동위형 유전자를 갖는 유전자만 고려할 경우 631개 PrCOS 유전자가 1,301개의 왕벚나무 유전자에 대응되어 1,301개의 왕벚나무 유전자를 왕벚나무 이형성 PrCOS 유전자로 최종 선발하였다(Figure 7).

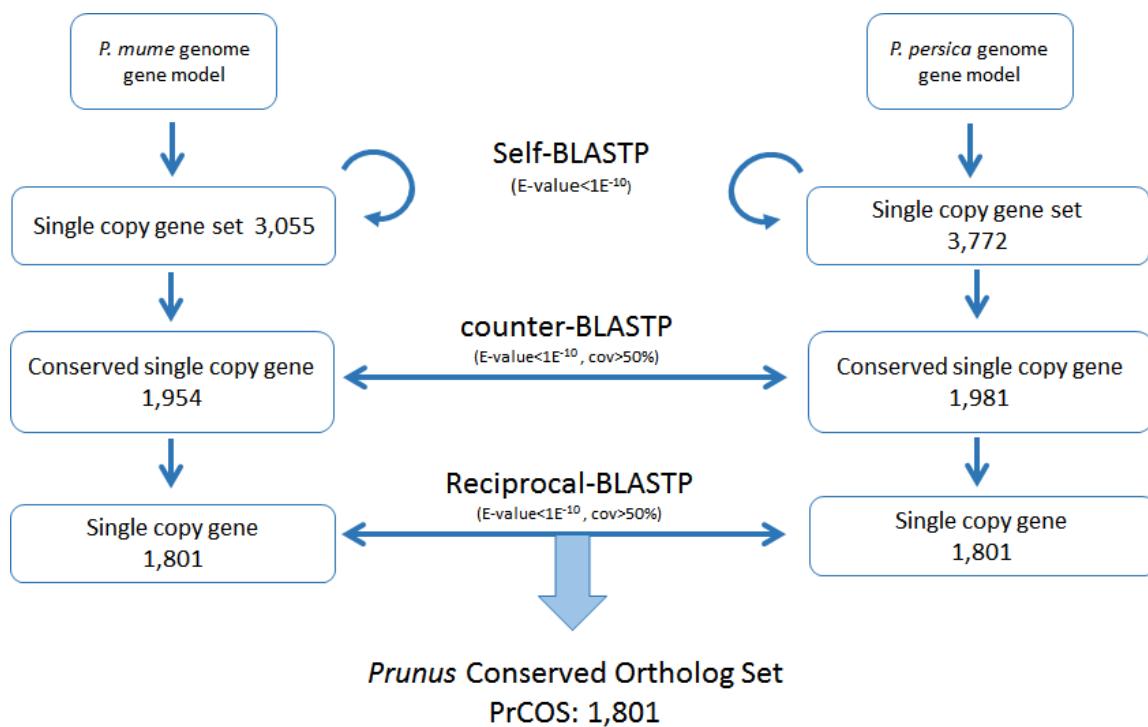


Figure 6. Workflow to select *Prunus* Conserved Ortholog Sets (PrCOS) candidate genes using stepwise BLASTP searches of *P. mume* and *P. persica* genomes

Counter-BLAST are successfully identified PrCOS genes; Reciprocal-BLAST search of single copy genes in each genome.

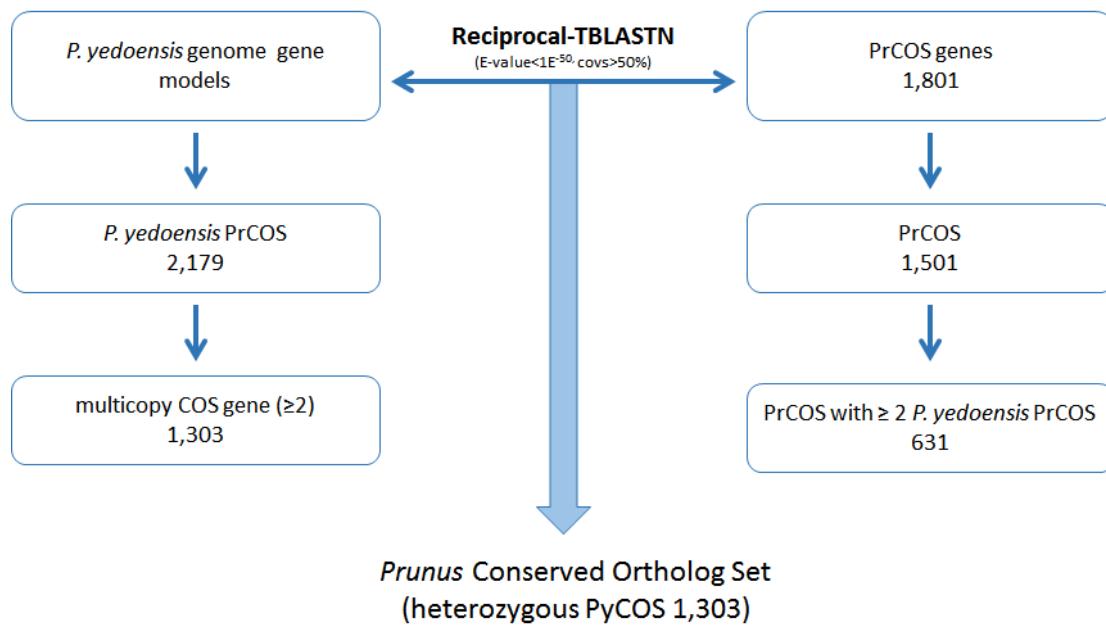


Figure 7. Workflow to select heterozygous Conserved Ortholog Set candidate genes in *P. yedoensis*

2. 왕벚나무 이형성 PrCOS 마커 설계

가. 왕벚나무 유전체 중 이형성 PrCOS 유전자 서열 추출

왕벚나무 이형성 PrCOS 분자 마커는 왕벚나무 전사체 서열과 핵 유전체 Illumina short read 조립 서열(scaffold) 그리고 왕벚나무 유전체 초안으로부터 예측한 유전자 모델을 이용하여 설계하였다.

왕벚나무 전사체 서열을 이용한 분자 마커를 개발하기 위하여 우선 중국 매실과 복숭아에서 추출한 PrCOS 유전자 1,801개에 왕벚나무 전사체 서열을 정렬하였다. 이후 전사체 서열 중에서 PrCOS 유전자와 정렬된 영역을 엑손이라고 추정하고 서열을 추출하였다.

왕벚나무 핵 유전체 Illumina short read 조립 서열(scaffold)을 기반으로 한 분자 마커를 개발하기 위하여 3가지 방법으로 각각 조립된 서열을 이용하였다. 왕벚나무 핵 유전체 조립 서열에서 엑손과 인트론 영역을 구분하기 위하여 우선적으로 중국 매실 PrCOS 유전자의 아미노산 서열을 추출하고, 왕벚나무 핵 유전체 조립 서열에 TBLASTN(E-value cutoff 1E⁻¹⁰)을 이용하여 매핑하였다. 왕벚나무 핵 유전체 서열 중에서 Allpath-LG로 조립된 서열 14,011개 중 2,966개, SOAPdenovo로 조립된 서열 871,246개 중 3,140개, Platanus로 조립된 서열 17,551개 중 141개가 각각 중국 매실의 PrCOS 유전자 1,711개, 1,802개, 65개에 매핑되었으며 일부 영역에서 정렬되었다. 정렬된 왕벚나무 핵 유전체 조립 서열의 일부 영역을 엑손, 정렬되지 않은 영역을 인트론이라고 추정하였다. 이후 하나의 중국 매실 PrCOS 유전자 서열에 두 개의 왕벚나무 핵 유전체 단편 서열이 인트론 길이 차이를 보이며 정렬되는 서열을 선발하였다. PacBio Rs II reads를 이용한 왕벚나무 핵 유전체 초안은 FALCON assembler를 이용하여 조립되었으며 이를 기반으로 한 유전자 모델을 사용하였다. 유전자 모델 서열 4,016개를 중국 매실의 PrCOS 유전자 서열에 BLASTN (E-value cutoff 1E⁻¹⁰)로 매핑한 결과, 왕벚나무 유전자 모델 서열 1,669개가 중국 매실의 PrCOS 유전자 서열 1,693개에 매핑되었다. 이후, 왕벚나무 유전자 모델 서열 1,669개를 추출하여 유전자 서열 내 Insertion/Deletion (InDel) 영역을 선별하였다.

나. 왕벚나무 이형성 PrCOS의 InDel PCR 마커 개발

왕벚나무 전사체 서열에서는 인트론 길이 차이나 InDel 영역을 확인할 수 없다. 따라서 PrCOS 유전자에 정렬된 왕벚나무 전사체 서열을 중국 매실 염색체 8개에 매핑하여 각 염색체에 골고루 분포되어있게 무작위로 왕벚나무 전사체 서열을 추출하고 그 서열을 기반으로 한 프라이머 27세트(PrC_v1.1)를 설계하였다. 또한 왕벚나무 핵 유전체 Illumina short read 조립 서열을 기반으로 한 프라이머 93세트를 개발하였다. 왕벚나무 핵 유전체 조립 서열 중 Allpath-LG, SOAPdenovo, Platanus로 조립된 서열을 기반으로 한 문자 마커는 각각 34세트(PrC_v1.2), 14세트(PrC_v1.3), 45세트(PrC_v1.4)이다. 한편 왕벚나무 유전체 초안으로부터 예측한 유전자 모델을 기반으로 한 프라이머는 36세트(PrC_v2.1)를 개발하였다. 각각 개발된 프라이머 세트를 왕벚나무와 근연종 벚나무류에 적용할 경우, 왕벚나무에서는 이형성을 나타내는 두 개의 밴드가 나타나야 하며, 그 중 하나의 밴드는 왕벚나무의 모계인 올벚나무와, 나머지 하나의 밴드는 부계로 추정되는 벚나무류와 공유할 것으로 기대하였다(Figure 8). 전체 156개 PCR 프라이머 세트에 대한 정보는 Table 5에 나타내었다.

프라이머 세트별 PCR 조건을 수립하고 벚나무류에서의 적용성을 확인하기 위하여 왕벚나무 2개체(Identification NO: 159-1, 159-2)와 서로 다른 집단에서 1개체씩 선발된 올벚나무 4개체(Sample ID: A04, A06, A09, A11)에 우선적으로 적용하였다. PrC_v1.1 27세트의 경우, 왕벚나무에서 두 개의 밴드 패턴이 나타나는 이형성 밴드를 띠지 않으며 올벚나무와 같은 크기의 하나의 밴드만 나타났다(Figure 9). 이러한 결과로 볼 때 PrC_v1.1 27세트는 왕벚나무 이형성 유전자의 엑손영역에서 무작위로 선발한 유전자를 기반으로 설계한 프라이머 세트로서 왕벚나무의 유전적 이형성을 나타내지 못한 것으로 판단되었다. 반면 왕벚나무 핵 유전체 조립 서열과 유전체 초안으로부터 예측한 유전자 모델 서열을 기반으로 한 문자 마커의 경우에는 왕벚나무에서는 두 개의 밴드를 나타내는 이형성 밴드 패턴이 나타났으며 그 중 하나의 밴드는 올벚나무와 공유하는 형태의 밴드 패턴이 나타났다(Figure 10). 이러한 밴드 패턴을 통하여 왕벚나무 핵 유전체 조립 서열 및 왕벚나무 유전자 모델로부터 개발한 마커들은 왕벚나무의 유전적 이형성을 명확하게 나타내는 것을 확인할 수 있었다. 이상의 결과를 종합하여 총 156개의 프라이머 세트를 개발하였고 각각의 프라이머 세트를 설계한 PrCOS 유전자

156개를 중국 매실의 염색체에 매핑한 결과, 8개의 염색체에 고르게 분포하고 있는 것을 확인하였다(Figure 11). 또한 156개의 PrCOS 유전자에 대한 프라이머 세트를 왕벚나무 핵 유전체 서열을 대상으로 PCR 증폭 여부, 예상되는 증폭 산물의 길이 등을 *in silico* PCR 방법으로 예측하여 검증하였다.

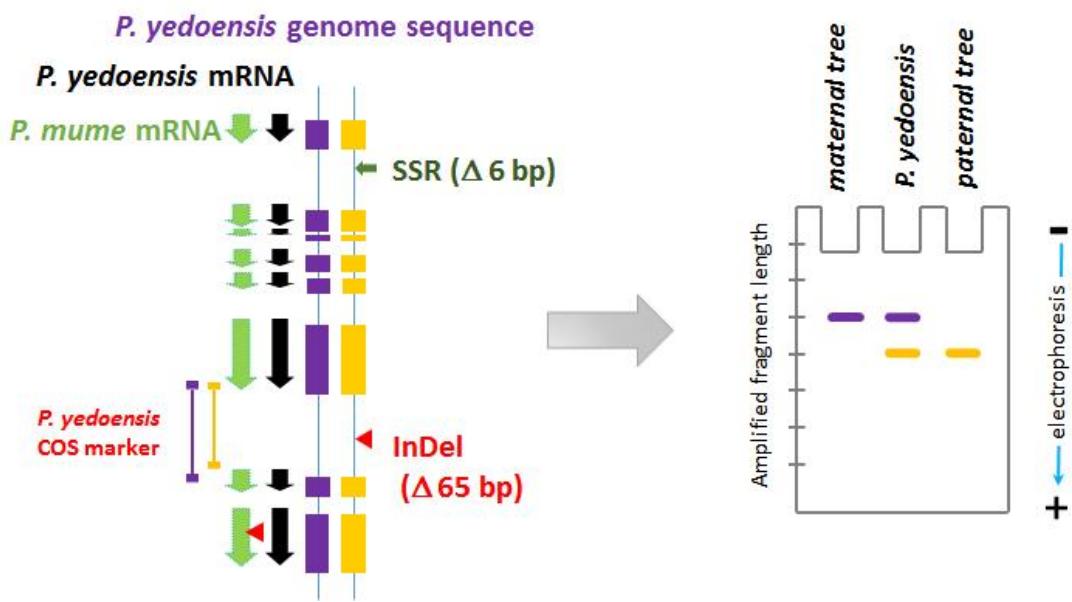


Figure 8. A diagram representing strategy of InDel marker development and expected result of marker application

Table 5. Summary statistics of 156 primer sets designed from candidate dimorphic *P. yedoensis* PrCOS genes

	Tm (°C)	GC (%)	Primer length (bp)	Predicted amplicon size in <i>P. yedoensis</i> (bp)	
Average	57.6	50.9	22	461.1	564.5
Maximum	65.6	66.7	26	1117.0	1639.0
Minimum	53.1	40.0	28	153.0	190.0
Median	57.4	50.0	22	422.0	535.0

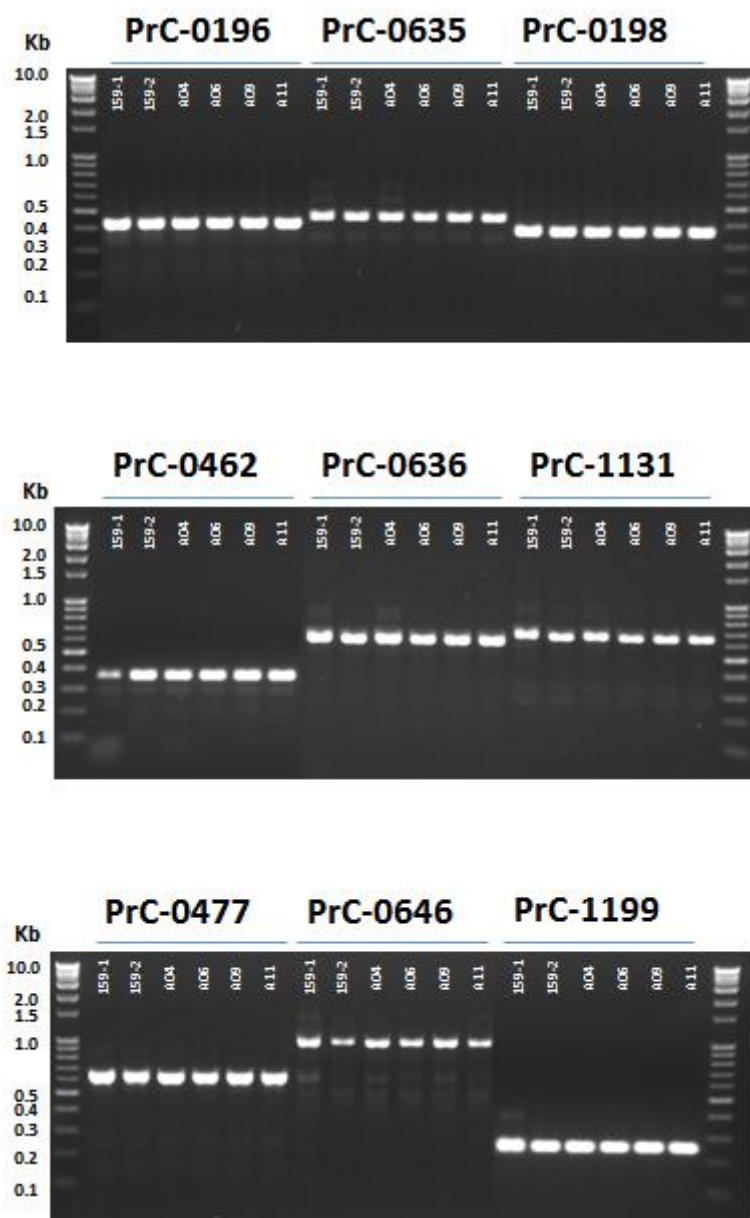


Figure 9. Examples of PCR amplification of PrCOS markers developed from transcript unigenes of *P. yedoensis*

Two *P. yedoensis* (159-1, 159-2) and four *P. pendula* for. *ascendens* (A04, A06, A09, A11) plants were tested to verify dimorphic amplification of the markers.

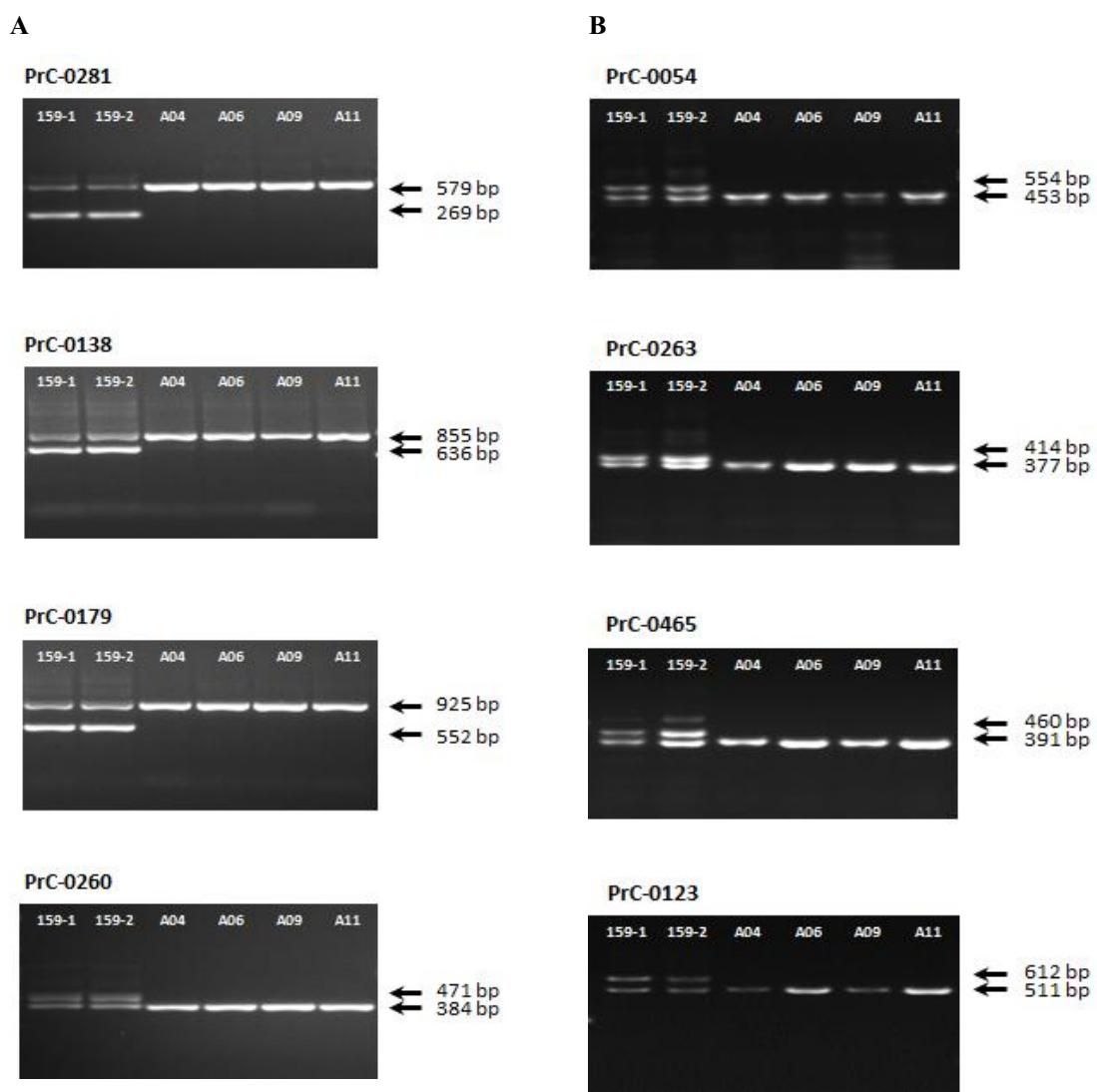


Figure 10. Examples of PCR amplification of PrCOS markers developed from Illumina short read assemblies (A) or predicted gene models of a draft genome of *P. yedoensis* (B)

Two *P. yedoensis* (159-1, 159-2) and four *P. pendula* for. *ascendens* (A04, A06, A09, A11) plants were tested to verify dimorphic amplification of the markers.

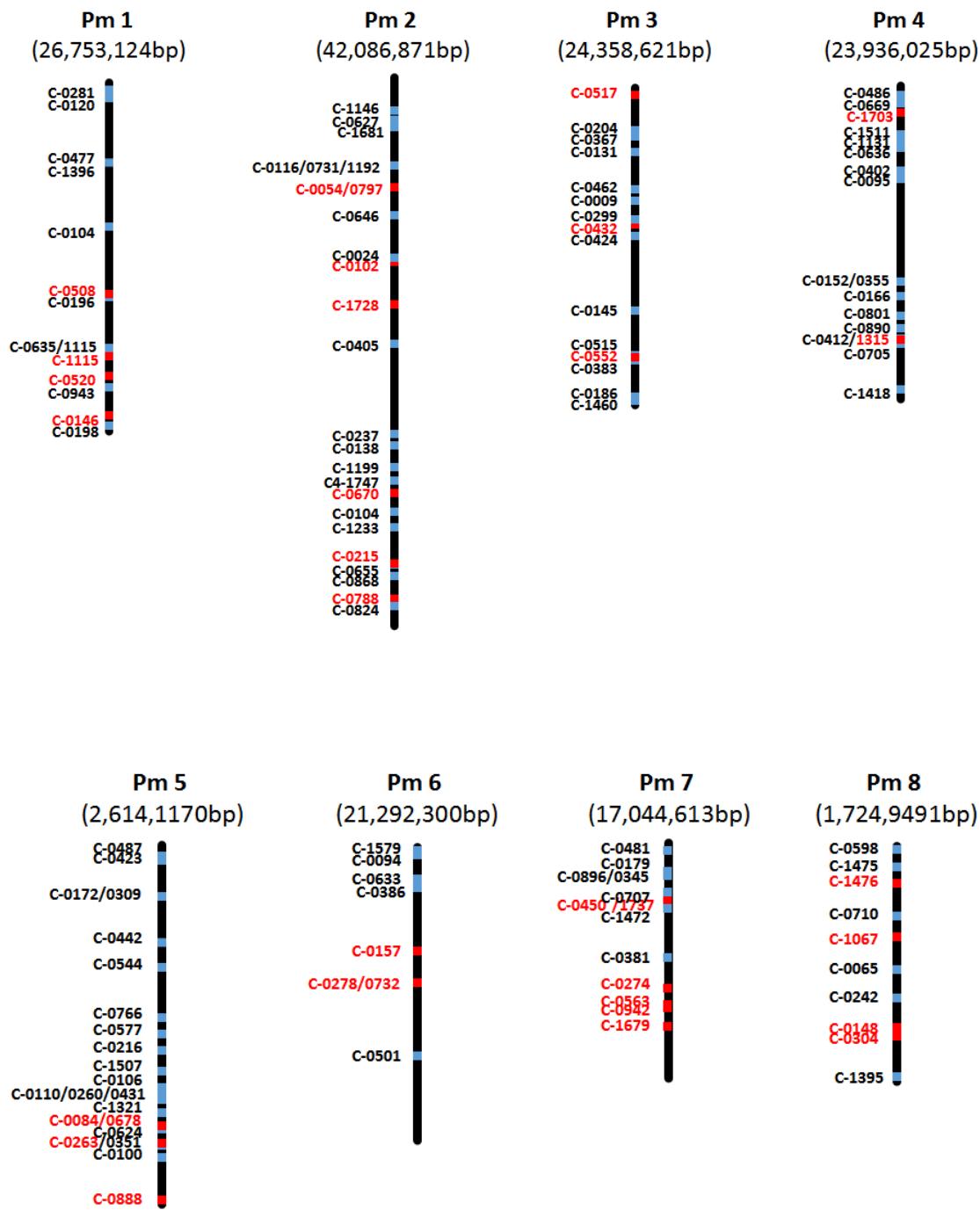


Figure 11. Locations of 156 PrCOS genes mapped to the *P. mume* chromosomes

PrCOS genes marked in red are chosen for genetic analysis of natural *Prunus* populations in Jeju island. Pm, *P. mume*

3. PrCOS 마커 검증 및 단계적 선발

제주도 자생 벚나무류의 계통분석과 유전구조 분석을 수행하기 위해서는 제주도의 주요 벚나무류 수종인 왕벚나무, 벚나무, 사옥, 산벚나무, 올벚나무 등에 PrCOS 마커가 범용적으로 적용되어야 한다. PrCOS 마커 156개에 대한 3단계의 적용성 평가를 수행하여 왕벚나무 종 구분 및 유전분석용 PrCOS 마커를 선발하였다. 선발 과정의 첫 번째 단계에서는 올벚나무와의 대립 유전자 다양성(allele diversity)을 확인하기 위하여 왕벚나무와 올벚나무에 156개의 PrCOS 프라이머 세트를 적용하여 왕벚나무와 올벚나무를 구별할 수 있는 PrCOS 마커를 선발하였다. 두 번째 단계에서는 제주도 자생 왕벚나무의 근연종 벚나무류에서 적용성을 확인하기 위하여 왕벚나무 기념물 지정 개체 일부와 왕벚나무의 모계 추정종인 올벚나무, 부계 추정종인 벚나무, 산벚나무, 사옥에 선발된 PrCOS 마커를 적용하였다. 세 번째 단계에서는 왕벚나무 종 구분 및 유전분석 분자 마커를 선발하기 위하여 왕벚나무 기념물 지정 개체 일부와 제주도 자생 근연종 벚나무류에 PrCOS 마커를 적용하였다. 최종적으로 왕벚나무 종 구분 PrCOS 분자 마커 및 왕벚나무 유전구조 분석에 사용한 왕벚나무의 유전적 이형성 PrCOS 분자 마커를 선발하였다.

156개 PrCOS 마커 프라이머 세트에 대해 왕벚나무 기념물 지정목 2개체 (Identification NO: 159-1, 159-2)와 서로 다른 집단에서 1개체씩 선발된 올벚나무 4개체(Sample ID: A04, A06, A09, A11)에 우선적으로 적용하여 적용성 평가를 수행하였다. 그 결과, 왕벚나무에서 두 개의 allele를 나타내고, 그 중 하나의 allele를 올벚나무와 공유하는 다양성 양상을 나타내는 마커 59세트를 1차 선발하였다(Figure 12). 왕벚나무 전사체 서열을 기반으로 한 PrC_v1.1 27세트에서는 4개의 프라이머 세트에서만 왕벚나무 이형성을 보였지만 올벚나무에서도 동일한 양상을 보였기 때문에 제외하였다. 반면 왕벚나무 핵 유전체 조립 서열을 기반으로 한 PrC_v1.2 프라이머 16세트, PrC_v1.3 프라이머 7세트, PrC_v1.4 프라이머 21세트를 각각 선발하였으며 왕벚나무 유전자 모델을 기반으로 한 PrC_v2.1 프라이머 세트에서 15세트를 선발하였다. 두 번째 단계 선발 과정으로 1차 선발된 마커를 왕벚나무 근연종에서 적용성을 확인하기 위해 2차 적용성 평가를 수행하였다. 2차 적용성 평가를 위해서 제주도 자생 왕벚나무

기념물 지정목 대표 7개체를 선발하였다(Sample ID: 15-Pr001, 15-Pr002, 15-Pr003, 15-Pr004, 15-Pr007, 15-Pr010, 15-Pr020). 선발된 왕벚나무 지정목 7개체와 올벚나무 (Sample ID: A04, A06), 벚나무(Sample ID: MJ_101), 산벚나무(Sample ID: MJ_111), 사옥(Sample ID: MJ_163)에 1차 선발된 프라이머 59세트를 적용하였다. 왕벚나무 지정목에서 항상 2개 allele를 나타내며 모계 후보종인 올벚나무와 하나의 대립 유전자를 공유하고, 다른 하나의 대립 유전자를 부계 후보종인 벚나무, 산벚나무, 사옥과 공유하는 마커 38세트를 2차 선발하였다. Figure 13에 나타낸 것과 같이 왕벚나무의 이형성 밴드 패턴의 경우 하나의 밴드는 올벚나무와 같은 크기의 밴드와 일치하고, 다른 하나의 밴드는 부계 추정종인 벚나무, 산벚나무, 사옥의 밴드와 일치한다. 하지만 부계 추정 종인 벚나무, 산벚나무, 사옥 내에서의 다형성은 없는 것으로 나타났다. 2차 선발 마커를 왕벚나무, 올벚나무, 벚나무, 산벚나무, 사옥, 산개벚지나무, 오시마벚나무에 적용하여 왕벚나무 종 구분 및 유전분석 마커 38세트를 선발하였다. 이상의 과정을 통해 PrCOS 마커 156개 세트로부터 제주도 자생 왕벚나무의 종 구분 및 유전 분석용 PrCOS 마커 38개를 최종 선발하였다(Figure 14 and 15, Table 6 and 7).

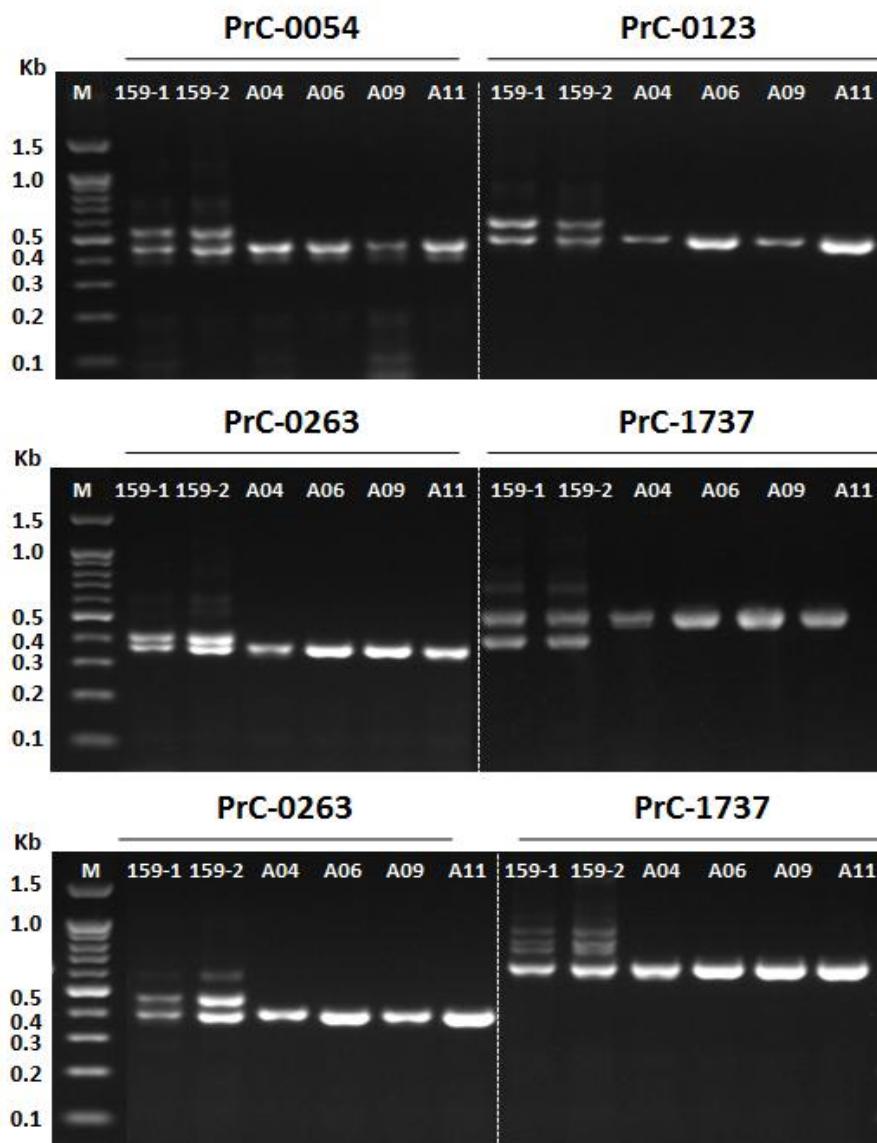


Figure 12. Examples of PCR amplification for selection of dimorphic PrCOS markers in *P. yedoensis*. Step 1: Application for *P. yedoensis* and *P. pendula* for. *ascendens*

Two *P. yedoensis* (Identification No. 159-1, 159-2) and four *P. pendula* for. *ascendens* (Sample ID: A04, A06, A09, A11) plants were tested to verify dimorphic amplification of the markers.

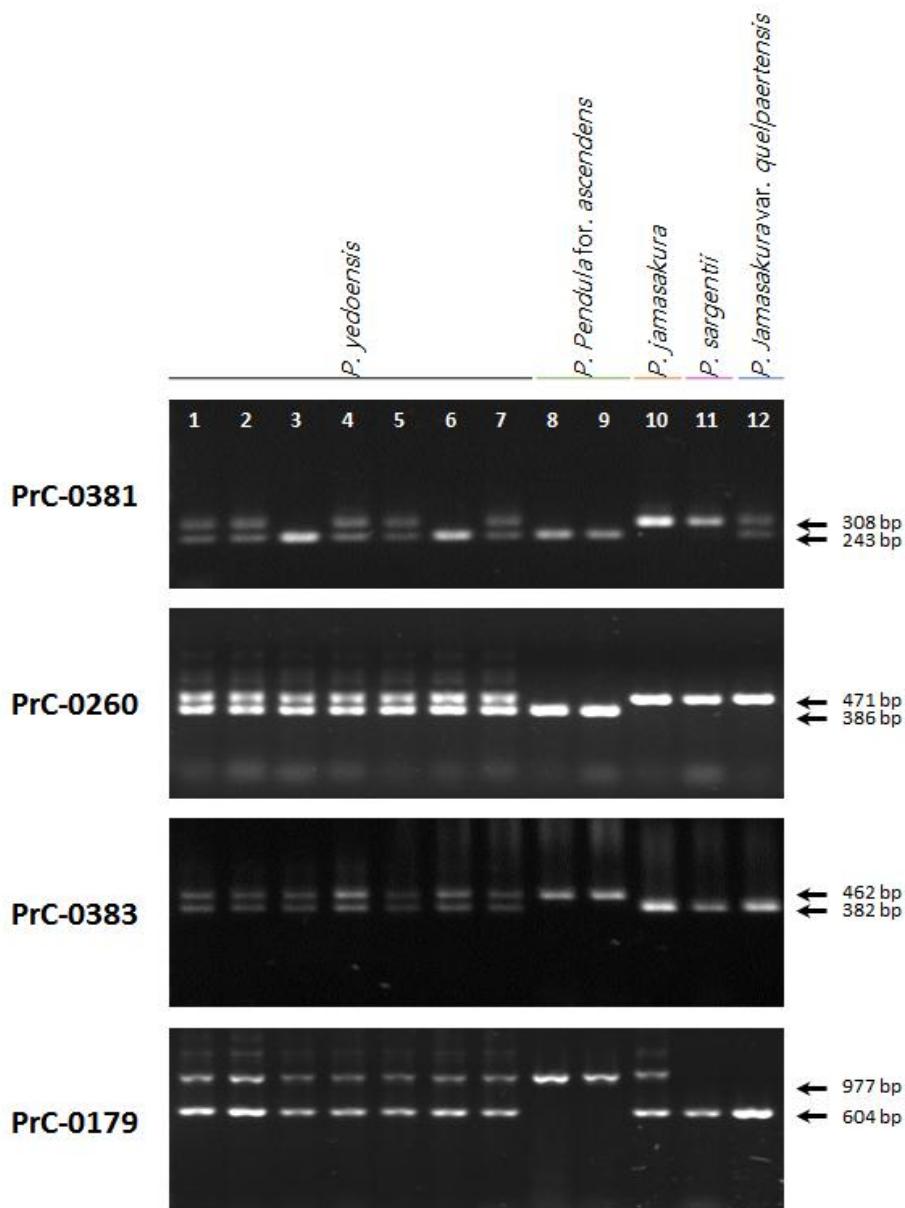


Figure 13. Examples of PCR amplification for selection of dimorphic PrCOS markers in *P. yedoensis*. Step 2: Application for *P. yedoensis* and its closely relative species

Seven *P. yedoensis* plants and five plants of closely relative *Prunus* species were tested to verify dimorphic amplification of the markers. Idendification Number is as follow: 1, 15-Pr001; 2, 15-Pr002; 3, 15-Pr-003 4, 15-Pr004; 5, 15-Pr007; 6, 15-Pr010; 7, 15-Pr020 8, A04; 9, A06; 10, MJ_101; 11, MJ_111; 12, MJ_163

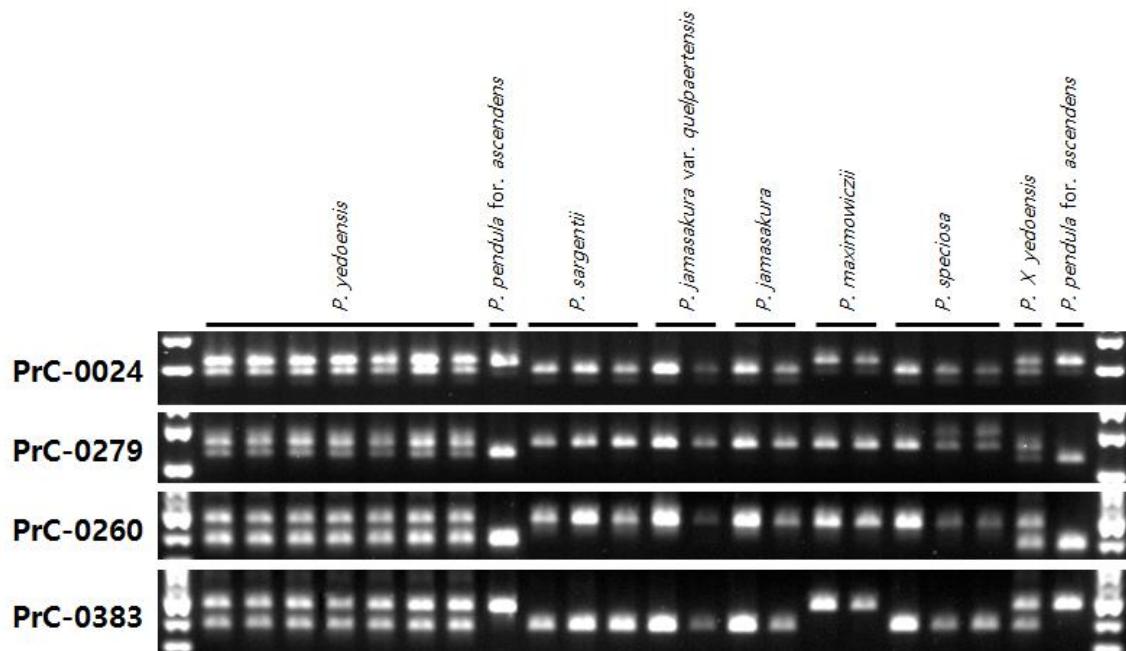


Figure 14. Examples of PCR amplification for selection of dimorphic PrCOS markers in *P. yedoensis*. Step 3: Application for *P. yedoensis* and its all available related species

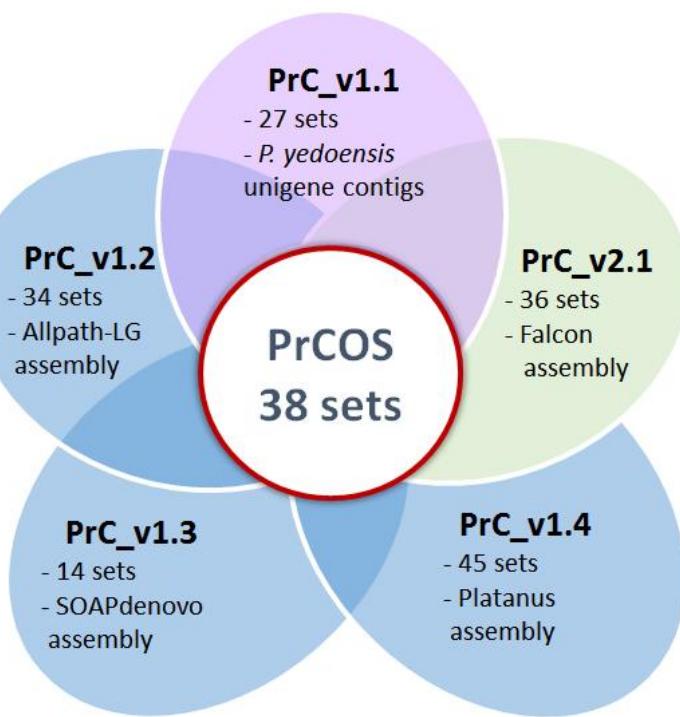


Figure 15. Summary of selected PrCOS markers

Table 6. Summarized statistics of 38 primer sets designed from dimorphic PrCOS markers

	Tm (°C)	GC (%)	Primer length (bp)	Predicted amplicon size in <i>P. yedoensis</i> (bp)	
Average	57.6	51.2	21.9	435.7	57.6
Maximum	63.5	63.2	25.0	925.0	63.5
Minimum	53.1	40.0	19.0	173.0	53.1
Median	57.3	51.1	22.0	394.5	57.3

Table 7. Summary of 38 primer sets for dimorphic PrCOS markers in *P. yedoensis*

No.	PrCOS marker ID	Primer sequence (5'→3')	Predicted amplicon size (bp)	
			A allele	B allele
1	PrC-0024	TGCTCATAGATTCTACCCTCGTTG AAGCGCAAACAATGCAAATTGCTC	194	223
2	PrC-0054	TGACGTCTAGCATGGGAACCA AGTTGTACATGAACCCGACCA	453	554
3	PrC-0110	GAGGTTGTGGATGAAGTTGTGC GCCCTCAATTGATCTAAGAGGGAG	470	534
4	PrC-0123	ACTCTCGTCTGGAAAGCA ACCCCTCCACTAAATGCCCATGA	511	612
5	PrC-0131	AGGATTCTCTGCTGTCACTGC TGCCAATGACCATCCTGTAGAC	269	579
6	PrC-0138	ACCATAGCAGCAAAGGACCTG CAAGGTTCAATGTTGGCTTCACC	636	855
7	PrC-0146	TGGAAACAGCTTACCAAGGCATGGT CCTCTTAGCAGATGACGTGACCCT	518	544
8	PrC-0148	ACTGTTGACACGTCCTGAGT ACAATACACGCCTCAAAAGG	314	339
9	PrC-0179	CGAGCATCTGAAGAAGCAGTGA AATCCGAGAACAGCTTCCCTG	552	925
10	PrC-0260	CATATGAAGGGCTTCTGATGAGGT CCCTCTGTTGGACTCCCATG	384	471
11	PrC-0263	GGTAGCTGCTCATTTCATGC TTCTGGTGCAGAGAGAAGCA	377	414
12	PrC-0279	GAAGGTCCACCCCTCCATGATT CTTCAAACATCCACCACGACC	236	264
13	PrC-0281	CTGCTGCAAATGATGTAGAGGA AGAGTTGTAGCCAAGTCTCCA	865	1,020
14	PrC-0367	AGAGTCTTGGTCAAATGGGT ACTCATAGGAGTTGGGAGAGC	453	555
15	PrC-0381	AACAGACTGGCTTGCTG TCCACAGGTACACAAGAGCAGC	243	308
16	PrC-0383	CCTGGGCAATCTGCAAAGGAGT CAGTCCCTACTGAGTGCCG	540	627
17	PrC-0386	TGGGGAACCTGTGGCTGAGAG CAGGAGCAGGTAAGTCTTAACTC	398	697

Table 7. (Continued)

No.	PrCOS marker ID	Primer sequence (5'→3')	Predicted amplicon	
			A allele	B allele
18	PrC-0402	GAGCCTGCTGAGTACACTGACA ACCACGTCTCCTTCGAGAACG	655	737
19	PrC-0454	GCTAAAACAACCCACGCCACATTG ACTCACAACAGTCCCATCAGGA	532	630
20	PrC-0465	TTGCTGCAAGCCCCATCTACG AAGGCCAAATTGCAGGATGTGGT	391	460
21	PrC-0508	ACTGTACCAGTTGAGAGGAGA GCTACCTCAACTAGGAATTCTGAG	642	821
22	PrC-0552	AAGGGAACACCCAGGACGCTG TTGCTTCAGCTCTGACCCCT	513	548
23	PrC-0563	GGCCCTTGTAGTCTCACATGGGA CGAGTATAGAGTCTCTGGCGT	222	239
24	PrC-0705	GCTGACCATTCAAGACCCAGGA GTGGTGGGCTTCACTCTTGTC	173	193
25	PrC-0732	ACAACGCCATACTGCGGAAG TCGTTGCAGTCTCCCAATGGAGA	280	314
26	PrC-0788	GCAACTCTCAAGGTGATTGGAA ACAATCAAGCGGGTCTACCA	370	397
27	PrC-0828	CGACTATACAATGATAGCCACAGC TAAGGCAGAAGCTGCCATCAGC	802	867
28	PrC-0888	TTGCAGATCCCATTGTCA TGATTAGGGCTTCGGCAACA	303	334
29	PrC-0890	GGTCAAGTTGCCAAGGATGCT ACACAACCAAGGTGAGTCAG	371	415
30	PrC-0896	TGGCTTGAAGGACAACAAGCA GTCATTCTATTCCCTTCGCCA	234	299
31	PrC-0954	ACGTTAGCAGCCAGGCTCAC CGCATTCCGTACTGGTTTC	174	194
32	PrC-1115	ACCCTAGCTCAATTCCCGCA GCAGCTTAACTCCAAATTGTCC	225	240
33	PrC-1418	TGATATCTCCTGGCTCCATCC GCTCCTGCATACAAGTCAGCAAC	500	659
34	PrC-1507	CATGGCATGTTCTCAGCTCTC ACTTGCGCATGAAGTTGGTGG	270	395
35	PrC-1535	GAGGATGAGGTTCTGGCTTCCT TGCAAGCATGACTCTCGTGGA	257	297
36	PrC-1579	CCATTGGCTATAAGACCCCTTGAG CGGTTTCGTTCTCCTTACTACC	254	275
37	PrC-1681	GGCTTCGCCAGAAGAATTAA CAGGCAAGAATAGAAAGCCATCT	607	885
38	PrC-1782	ATTGTTAAAGCCCTAGCAGGT CGTATCCCTCGATCTGCACC	380	429

제 3 절 제주도 왕벚나무 자생 집단의 유전형 및 유전구조 분석

1. 왕벚나무 자생 집단의 개체별 유전형 분석

제주도 한라산 백록담을 기준으로 구분한 자생 왕벚나무의 4집단에 속하는 52개 자연개체와 기념목 21개체를 대상으로 20개의 PrCOS 분자 마커를 적용하여 유전형 분석을 수행하였다. 또한 유전형 비교분석을 위하여 재배 왕벚나무 5개체, 올벚나무, 벚나무, 산벚나무, 사옥을 각각 1개체씩 포함하였다. 재배 왕벚나무와 근연종 벚나무류는 종을 기준으로 각각 하나의 집단으로 구성하였다.

제주도 왕벚나무 자생 집단의 유전구조 비교를 위하여 총 82개체 벚나무류의 유전형을 20개의 PrCOS 분자 마커(Table 8)로 분석한 결과, 왕벚나무에서 서로 다른 두 개의 대립 유전자에 대한 밴드를 육안으로 뚜렷하게 구분할 수 있었다. 왕벚나무 각 개체의 PrCOS 마커에 대한 유전형은 올벚나무와 동일한 대립 유전자를 가지는 밴드를 A 밴드, 부계 후보종인 벚나무, 산벚나무, 사옥과 같은 대립 유전자를 가지는 밴드를 B 밴드라고 지정했다. 각 PrCOS 마커를 적용한 개체별 PCR 반응산물의 아가로스 젤 전기영동 결과를 육안으로 확인하여 A 밴드와 B 밴드에 대한 유전형을 결정하였다(Figure 16). 유전형 결정 결과는 마이크로소프트사의 엑셀 프로그램을 이용하여 매트릭스로 작성한 후 codominant data template로 구성하여 계통수 작성과 유전형 분석에 사용하였다(Table 9).

Table 8. Summary of 20 primer sets for dimorphic PrCOS markers used in phylogenetic and population structure analysis of *P. yedoensis*

No.	PrCOS marker ID	Primer sequence (5'→3')	Predicted amplicon size (bp)	
			A allele	B allele
1	PrC-0024*	TGCTCATAGATTCTACCCTCGTT AAGCGCAAACAAC TGCAAATTGCTC	194	223
2	PrC-0054*	TGACGTCTAGCATGGGAACCA AGTTGTACATGAACCCGACCA	453	554
3	PrC-0131	AGGATTCTCTGCTGTCAGTGC TGCCAATGACCATCCTGTAGAC	269	579
4	PrC-0138	ACCATAGCAGCAAAGGACCTG CAAGGTTCAATGTTGGCTTCACC	636	855
5	PrC-0179	CGAGCATCTGAAGAACAGCAGTGA AATCCGAGAACAGCTTCCCTG	552	925
6	PrC-0237	GCCTGACAGTGATTTCACAGC CCTGGAAACTCCTACAGCT	475	726
7	PrC-0260	CATATGAAGGGCTTCTTGATGAGGT CCCTCTTGTGGACTCCATG	384	471
8	PrC-0263*	GGTAGCTGCTCATTTCATGC TTCTGGTGCAGAGAACAGCA	377	414
9	PrC-0281	CTGCTGCAAATGATGTAGAGGA AGAGTTGTAGCCAAGTCTCCA	865	1,020
10	PrC-0304*	AGAGCCATGCAGGTCTCTTC TAAGAAGGGGCCAATCTCGCA	229	245
11	PrC-0383*	CCTGGCAATCTGCAAAGGAGT CAGTCCCTACTGAGTGCCG	540	627
12	PrC-0386	TGGGAACTTGTGGCTGAGAG CAGGAGCAGGTAAGTCTTAAC	398	697
13	PrC-0402	GAGCCTGCTGAGTACACTGACA ACCACGTCTCCTCGAGAAC	655	737
14	PrC-0508*	ACTGTACCACTGAGAGGAGA GCTACCTCAACTAGGAATTCTGAG	642	821
15	PrC-0517*	TGCCTATTGTGATGGGGTCCA TGAGCACTTGTGCTTCCGGT	439	459
16	PrC-0563*	GGCCCTTGTAGTCTCACATGGGA CGAGTATAGAGTCTCTGGCGT	222	239
17	PrC-0705*	GCTGACCATTACAGACCCAGGA GTGGTGGGCTTCACTCTTGTGTC	173	193

Table 8. (Continued)

No.	PrCOS marker ID	Primer sequence (5'→3')	Predicted amplicon size (bp)	
			A allele	B allele
18	PrC-0890*	GGTCAAGTTGCCAAGGATGCT ACACAACCAAGGTGAGTCAG	371	415
19	PrC-1418	TGATATCTCCTTGGCTCCATCC GCTCCTGCATACAAGTCAGCAAC	500	659
20	PrC-1681*	GGCTTCGCCCAGAAAGAATT CAGGCAAGAATAGAAAGCCATCT	607	885

*Target sequence of PrCOS markers were defined to *P. yedoensis* gene model.

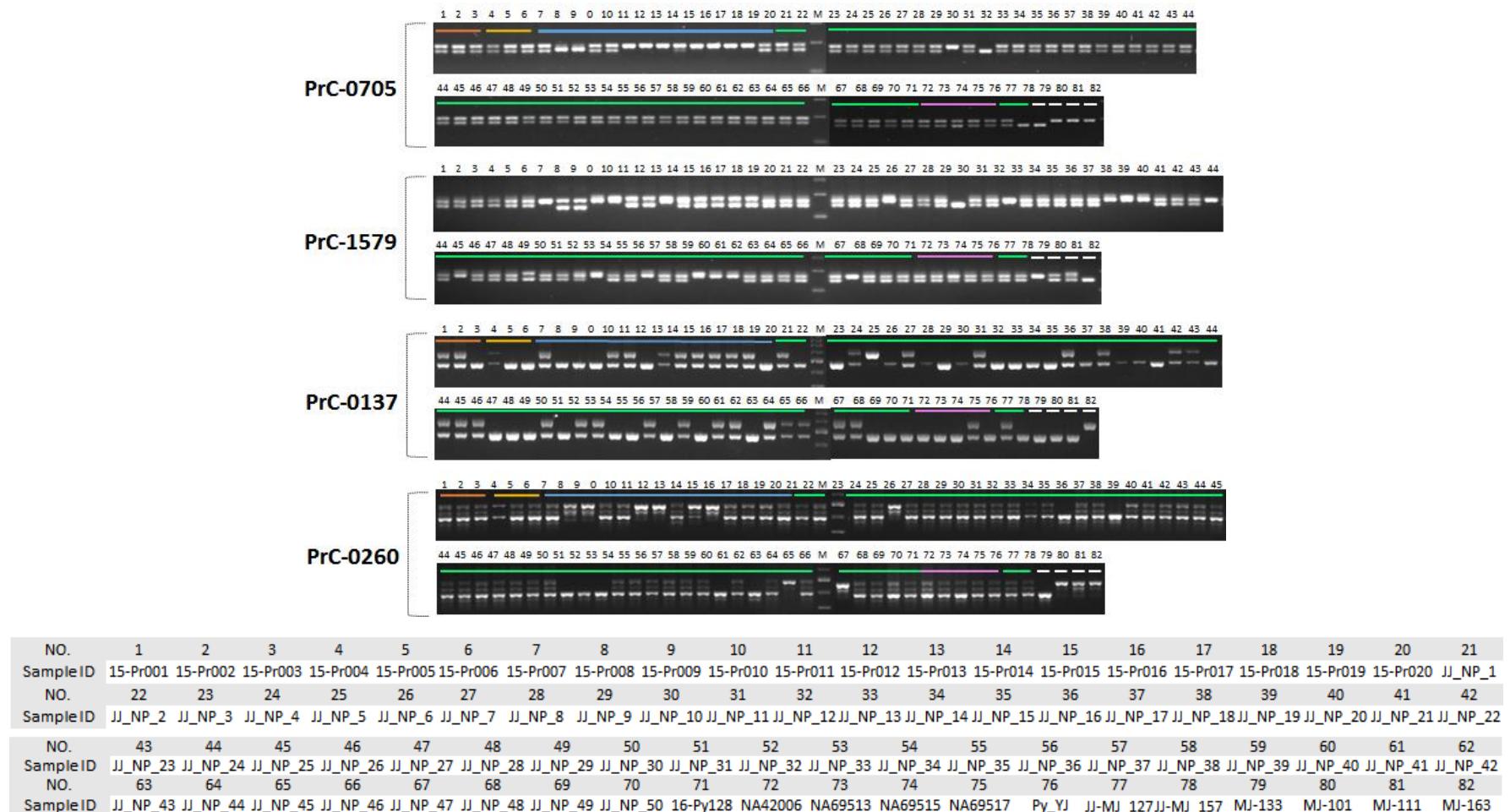


Figure 16. Examples of application of the dimorphic PrCOS markers for genotyping of wild *Prunus* populations

Table 9. Genotyping matrix of wild *P. yedoensis* plants using 20 PrCOS markers

Table 9 (Continued)

Table 9 (Continued)

	1681-B	1418-B	1418-A	1418-B	1681-A
JJ_NP_35					
JJ_NP_36					
JJ_NP_37					
JJ_NP_38					
JJ_NP_39					
JJ_NP_40					
JJ_NP_41					
JJ_NP_42					
JJ_NP_43					
JJ_NP_44					
JJ_NP_45					
JJ_NP_46					
JJ_NP_47					
JJ_NP_48					
JJ_NP_49					
JJ_NP_50					
16-Py128					
<i>P. Xy-1¹⁾</i>					
<i>P. Xy-2¹⁾</i>					
<i>P. Xy-3¹⁾</i>					
<i>P. Xy-4¹⁾</i>					
<i>P. Xy-5¹⁾</i>					
JJ-MJ_127					
JJ-MJ_157					
MJ_133					
MJ_101					
MJ_111					
0024-B					
0024-A					
0054-B					
0131-B					
0131-A					
0138-B					
0138-A					

¹⁾, data not available
P. X yedoensis

2. 제주도 왕벚나무 자생 집단의 유전구조 및 계통 비교 분석

제주도 자생 왕벚나무 각 개체간의 유전적 관계(genetic relationship), 집단 내의 유전적 유사성(genetic similarity), 집단별 계통 등을 파악하기 위하여 계통 분석과 유전구조 분석을 수행하였다.

가. 계통 분석

올벚나무, 벚나무, 산벚나무, 사옥을 외집단(outgroup)으로 하여 왕벚나무 자생 집단과 지정목, 재배 왕벚나무의 계통적 유연관계를 MEGA7 프로그램의 Maximum Likelihood 방법으로 부트스트랩 1,000회로 지정하여 분석하였다. 그 결과, 제주도 벚나무류는 2개의 클레이드(clade)를 형성하는 것으로 나타났다(Figure 17). 첫 번째 클레이드는 대부분의 왕벚나무와 올벚나무가 속하는 집단으로서 전체 왕벚나무의 90% (전체 왕벚나무 78개체 중 70개체)를 포함하고 있었다. 특히 제주도 자생 왕벚나무와 재배 왕벚나무는 계통적으로 구분되지 않았으며, 관음사에서 채집된 6개체들은 올벚나무와 유연관계가 가까운 것으로 나타났다. 한편, 두 번째 클레이드는 부계 추정종인 벚나무, 산벚나무, 사옥과 함께 자생 왕벚나무 중 신예리에서 채집된 7개체들과 제주도 북동 지역(JJ-NE)에서 채집된 1개체로 구성되었는데, 이 결과로 볼 때 이들 왕벚나무 개체들은 부계 추정종들과 유연관계가 가까운 것으로 판단되었다.

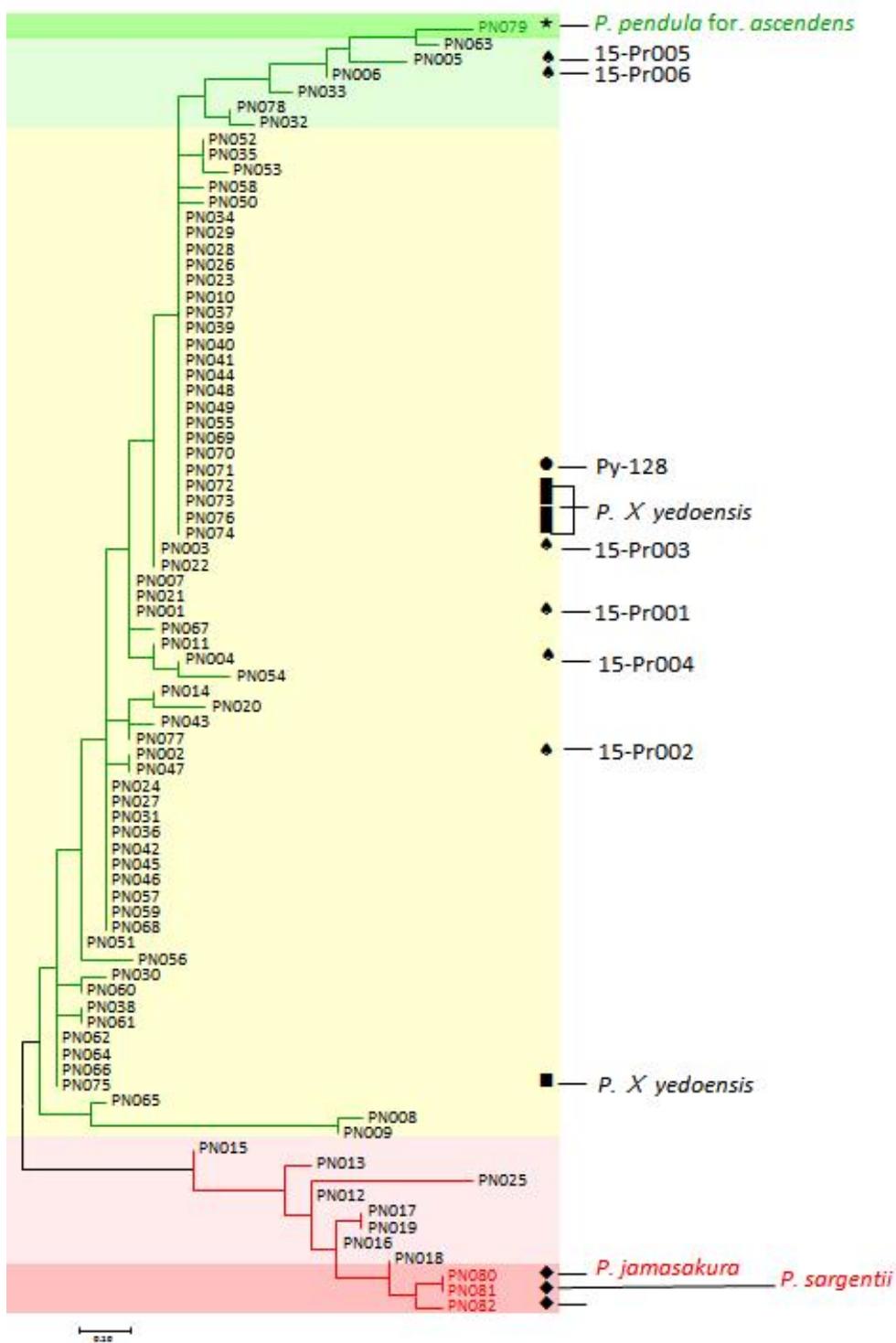


Figure 17. A maximum Likelihood tree of *P. yedoensis* and its closely related species using 20 PrCOS markers

나. 유전구조 분석

왕벚나무 각 개체에서 모계 및 부계 유전형의 유전적 조성을 STRUCTURE 2.3.4 프로그램으로 분석하였다. Evanno method를 이용한 ad hoc criterion (ΔK) 분석 결과, 왕벚나무 유전 집단의 수(K)는 확률적으로 볼 때 $K=2$ 로 결정하였다(Table 10). 대립 유전자의 빈도(allele frequency)를 기준으로 유전 집단을 2개로 구분한 결과, 왕벚나무의 대부분은 모계 추정종인 올벚나무와 하나의 집단을 형성했으며, 8개체는 부계 추정종인 벚나무, 산벚나무, 사옥과 하나의 집단을 형성하여 계통 분석에 의한 유연관계와 동일한 양상을 보였다(Figure 18A). 특히 제주도 자생 왕벚나무의 약 80%에 해당하는 59개체들은 모계와 부계 유전형을 절반씩 나타내고 있어 이형접합체(heterozygote)로 판단되었으며, 약 8%인 6개체들은 모계 유전형을 75%, 약 11%인 8개체들은 부계 유전형을 75%씩 가지고 있는 것으로 분석되었다. 지역적으로 볼 때, 모계와 부계의 유전형에 치우친 유전적 조성을 보이는 개체들은 각각 관음사와 신예리 지역에서 채집된 개체들로서 이들은 계통 분석 결과에서도 각각 모계와 부계에 유연관계가 가까운 것으로 나타난 것들이다.

지역적으로 왕벚나무 개체들의 유전형을 재분석한 결과(Figure 18B), 제주 북동 집단(JJ-NE)은 올벚나무와 부계 추정종의 교잡형(37개체, 84%), 올벚나무의 유전형 75% 개체(6개체, 14%), 부계 추정종의 유전형 75% 개체(1개체, 2%)로 구성되어 가장 복잡한 양상을 보였다. 제주 남동 집단(JJ-SE)은 올벚나무와 부계 추정종의 교잡형(6개체, 46%), 부계 추정종의 유전형 75% 개체(7개체, 54%)로 구성되어 있다. 한편, 제주도 서쪽 지역 집단(JJ-NW, JJ-SW)은 모두 올벚나무와 부계 추정종의 교잡형으로만 구성되어 있어 가장 단순한 양상을 보였다(Figure 19).

Table 10. Determination of optimum K value for *P. yedoensis* populations using Evanno method

K	Replications	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	Ln''(K)	Delta K
1	10	-2200.66	1.121705	—	—	—
2	10	-2037.33	9.949322	163.33	155.23	15.60207
3	10	-2029.23	3.875292	8.1	7.97	2.056619
4	10	-2029.1	3.824773	0.13	0.93	0.243152
5	10	-2028.04	1.402537	1.06	4.73	3.372459
6	10	-2031.71	8.842882	-3.67	4.05	0.457996
7	10	-2031.33	5.922096	0.38	8.72	1.472452
8	10	-2039.67	30.619349	-8.34	0.46	0.015023
9	10	-2047.55	32.128674	-7.88	0.56	0.01743
10	10	-2054.87	51.664603	-7.32	6.41	0.124069
11	10	-2055.78	50.871992	-0.91	13.24	0.260261
12	10	-2043.45	41.990852	12.33	19.42	0.462482
13	10	-2050.54	43.291292	-7.09	0.03	0.000693
14	10	-2057.6	54.440855	-7.06	4.14	0.076046
15	10	-2068.8	58.687099	-11.2	—	—

A

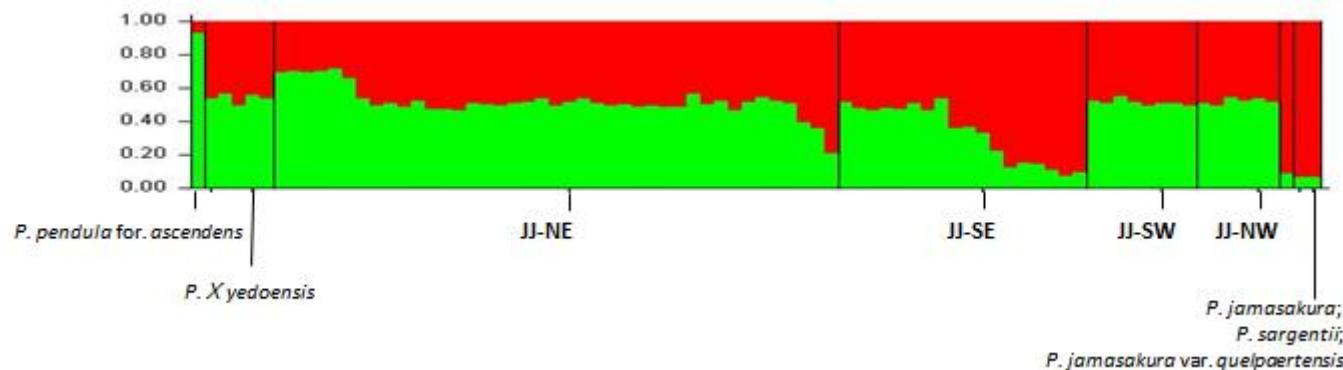
B


Figure 18. Genetic structure analysis of wild *P. yedoensis* population

A, Population structure representing two tentative groups of wild *P. yedoensis* genotypes based on 20 PrCOS markers, **B**, genetic composition of individual plants in each wild *P. yedoensis* population.

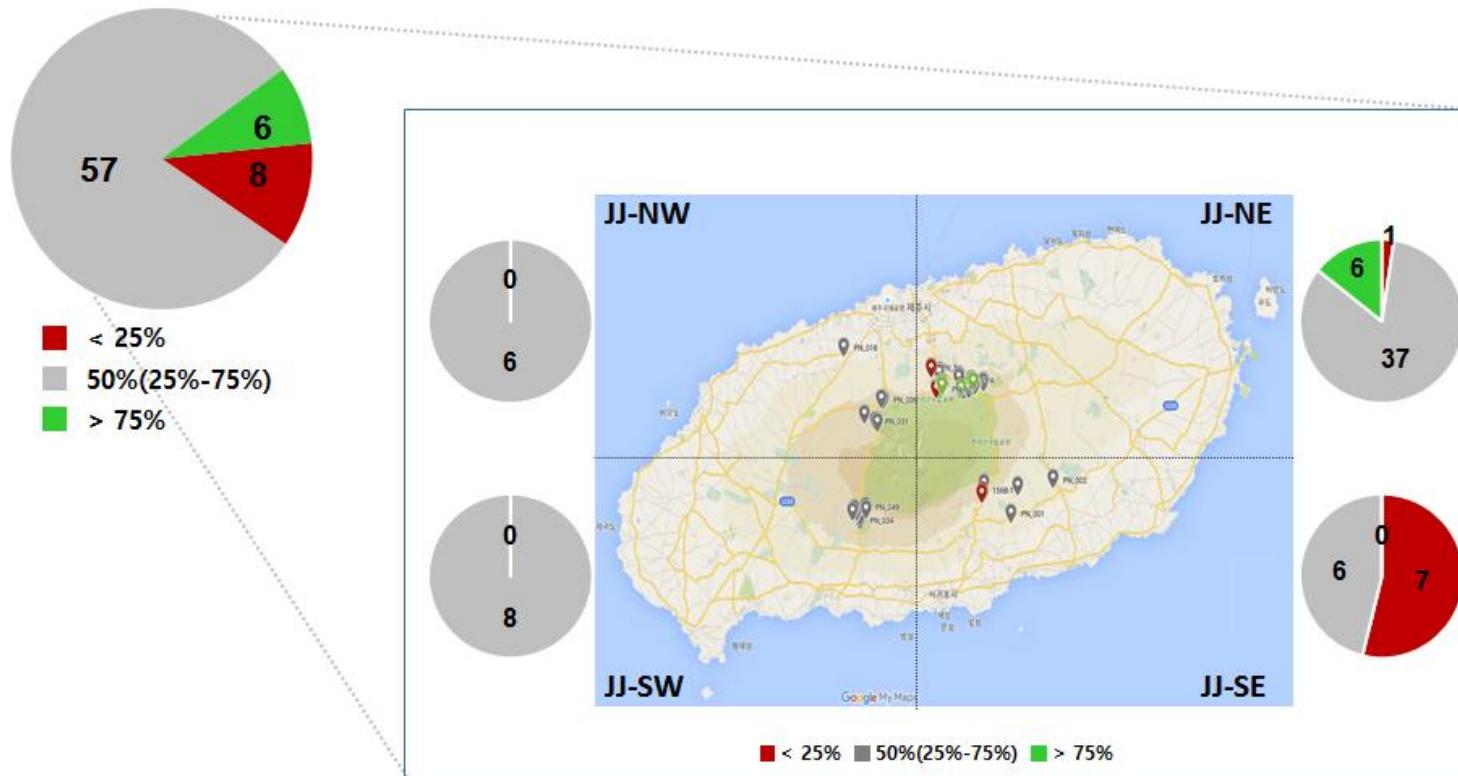


Figure 19. Haplotype analysis of each wild *P. yedoensis* population. Pie charts were presented based on percentages of maternal genotype

다. 자생 왕벚나무 유전형의 재구성

제주도 자생 왕벚나무의 유전형 분석 결과, 대부분의 왕벚나무는 모계와 부계 유전자를 절반씩 가지고 있어 이형교잡종인 것으로 추정된다. 한편, 일부 왕벚나무의 경우에는 모계 혹은 부계 유전자를 75% 가지고 있어 이형교잡종의 개체가 모계나 부계 유전형의 개체와 역교배되어 이들의 유전형이 추가 유입된 것으로 판단할 수 있다(Figure 20).

지역적으로 볼 때, 제주 북동 지역(JJ-NE)에서는 자생 왕벚나무에 모계 추정종인 올벚나무의 유전자가 추가 유입된 개체가, 제주 남동 지역(JJ-SE)에서는 자생 왕벚나무에 부계 추정종인 벚나무, 산벚나무 또는 사옥의 유전자가 추가 유입된 개체가 주로 나타난다. 이러한 현상은 제주도 북동 지역과 남동 지역의 환경적 차이, 예를 들어 지역의 일조량, 기온, 풍속 등의 차이에 의해 벚나무류의 개화 조건이 달라져 지역적으로 서로 다른 교잡종이 형성될 수 있었을 것으로 추정하며 향후 이에 대한 추가 연구가 필요할 것으로 사료된다.

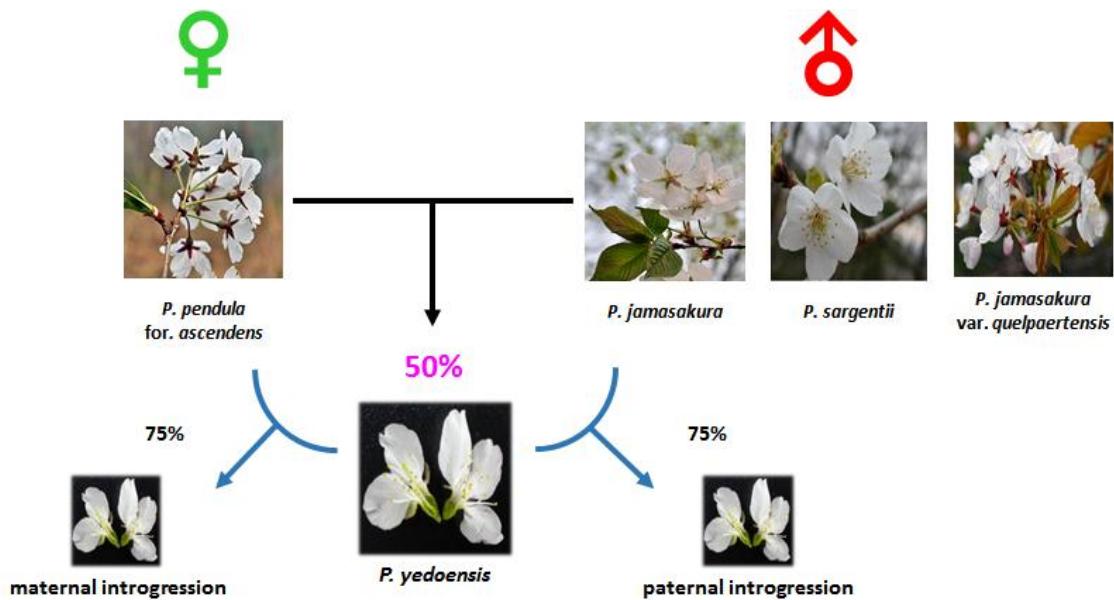


Figure 20. Reconstruction of genetic structure of wild *P. yedoensis* in Jeju island

Most of wild *P. yedoensis* plants in Jeju island are hybrid offspring of a cross between maternal *P. pendula* for. *ascendens* and paternal *P. jamasakura*, *P. sargentii* or *P. jamasakura* var. *quelpaertensis*. However approximately 8% and 11% of wild *P. yedoensis* in Jeju island are supposed to be originated by introgression of maternal or paternal genotype, respectively.

제 4 절 관음사 및 봉개동 왕벚나무 자생 집단의 개체별 유전형 분석

1. 벚나무류 근연종의 유전구조 분석 및 왕벚나무와의 계통 비교 분석

가. 왕벚나무와 근연 벚나무류 자연 집단의 구성

왕벚나무와 근연 벚나무류가 혼재하는 자연 집단의 유전구조를 분석하기 위하여 왕벚나무 및 근연종 벚나무류의 자생지를 고려하여 봉개동 집단과 관음사 집단을 구성하였다. 관음사 집단은 왕벚나무 5개체, 올벚나무 5개체, 벚나무 7개체, 산벚나무 4개체, 사옥 1개체로(Table 11), 봉개동 집단은 왕벚나무 4개체, 올벚나무 6개체, 벚나무 4개체, 산벚나무 1개체, 사옥 1개체로 구성되었다(Table 12).

나. 관음사 집단과 봉개동 집단의 유전형 결정

제주도 자생 왕벚나무와 근연종 벚나무류의 계통 분석을 위하여 관음사 집단과 봉개동 집단의 유전형을 앞서 선발한 PrCOS 분자 마커 20개를 이용하여 분석하였다. 그 결과 제주도 자생 왕벚나무와 근연 벚나무류의 종내 유전적 다형성은 매우 적었다. 또한 부계 후보로 추정되는 벚나무와 산벚나무, 사옥간에는 마커의 길이 다형성이 전혀 없어 왕벚나무의 부계를 직접 추정할 수 없었다. 그러나 PrCOS 마커의 PCR 빈드 증폭 양상을 통하여 자생 왕벚나무와 근연종 벚나무류는 명확하게 구분할 수 있었다. 지리적 위치를 고려하여 나눈 제주도 북동(JJ-NE) 집단 전체에서는 왕벚나무로의 모계 및 부계 유전자 유입이 확인되었으나 북동 집단(JJ-NE)내의 관음사 집단과 봉개동 집단에서는 PrCOS 분자 마커가 모계나 부계에 특이하게 증폭되는 경우가 많지 않아 유전자 유입을 확인하기 어려웠다(Figure 21).

Table 11. List of populations and individual wild *Prunus* species collected from Kwaneumsa temple region

Scientific name	Sample ID ¹⁾	Voucher No. ²⁾
<i>P. yedoensis</i>	15-Pr004	
	15-Pr005	Registered as Natural
	15-Pr006	Monument
	JJ_MJ_127	63401
	JJ_MJ_128	63402
<i>P. pendula</i> for. <i>ascendens</i>	JJ_MJ_103	63377
	JJ_MJ_107	63381
	JJ_MJ_116	63390
	JJ_MJ_121	63395
	JJ_MJ_126	63400
<i>P. jamasakura</i>	JJ_MJ_101	63375
	JJ_MJ_102	63376
	JJ_MJ_106	63380
	JJ_MJ_110	63384
	JJ_MJ_117	63391
	JJ_MJ_118	63392
	JJ_MJ_125	63399
<i>P. sargentii</i>	JJ_MJ_105	63379
	JJ_MJ_111	63385
	JJ_MJ_114	63388
	JJ_MJ_115	63389
<i>P. jamasakura</i> var. <i>quelpaertensis</i>	JJ_MJ_163	63437

¹⁾Sample IDs were given according to an order of collection in this study.

²⁾Vouchers were deposited at the herbarium of Warm-Temperature and Subtropical Forest Research Center, NIFS.

Table 12. List of populations and individual wild *Prunus* species collected from Bonggae-dong region

Scientific name	Sample ID ¹⁾	Voucher No. ²⁾
<i>P. yedoensis</i>	15-Pr001	Registered as Natural Monument
	15-Pr002	
	15-Pr003	
<i>P. pendula</i> for. <i>ascendens</i>	JJ_MJ_157	63431
	JJ_MJ_133	63407
	JJ_MJ_152	63426
	JJ_MJ_154	63428
	JJ_MJ_158	63432
	JJ_MJ_160	63434
	JJ_MJ_162	63436
<i>P. jamasakura</i>	JJ_MJ_150	63424
	JJ_MJ_151	63425
	JJ_MJ_155	63429
	JJ_MJ_161	63435
<i>P. sargentii</i>	JJ_MJ_153	63427
<i>P. jamasakura</i> var. <i>quelpaertensis</i>	JJ_MJ_159	63433

¹⁾Sample IDs were given according to an order of collection in this study.

²⁾Vouchers were deposited at the herbarium of Warm-Temperature and Subtropical Forest Research Center, NIFS.

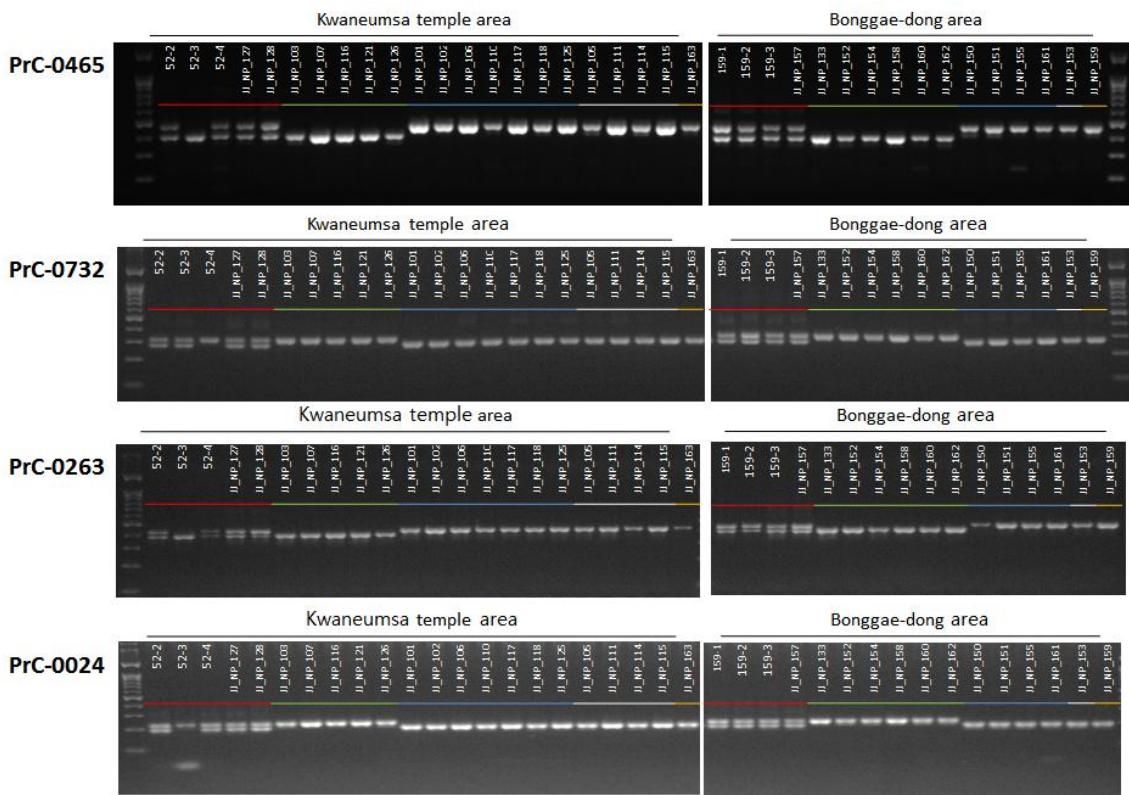


Figure 21. Examples of PCR amplification of PrCOS markers using *Prunus* populations in Kwaneumsa temple area and Bonggae-dong area

다. 관음사 집단과 봉개동 집단의 계통 및 유전구조 분석

관음사 집단과 봉개동 집단의 근연종간의 집단 내 유전적 유연관계를 파악하기 위하여 상기 기술한 방법에 따라 Maximum Likelihood 계통 분석을 수행하였다. 그 결과, 두 집단 모두 자생 왕벚나무와 올벚나무는 하나의 클레이드를 형성했으나 각각 독립된 집단으로 묶였다. 이에 비하여 벚나무와 산벚나무, 사옥은 함께 하나의 클레이드를 형성하여 왕벚나무/올벚나무 클레이드와 분리되었다. 이러한 결과를 통하여 관음사와 봉개동의 왕벚나무는 모계인 올벚나무와 부계인 벚나무, 산벚나무, 사옥 등에서 유래했으며 집단 내 근연종의 유연관계는 유사한 것으로 판단된다(Figure 22).

한편, 집단의 유전구조 분석 결과도 계통 분석과 동일한 결과를 나타냈는데, 관음사 집단과 봉개동 집단 모두 모계와 부계 유전형을 각각 50%씩 가지고 있는 이형접합체의 특성을 보였으며, 두 집단 간에 다형성은 나타나지 않았다(Figure 23). 이상의 결과를 종합할 때, 제주도의 자생 왕벚나무는 각 지역에서 모계 추정종인 올벚나무와 부계 추정종인 벚나무, 산벚나무, 사옥 사이에서 자연 교접에 의해 형성된 것으로 추정되었다.

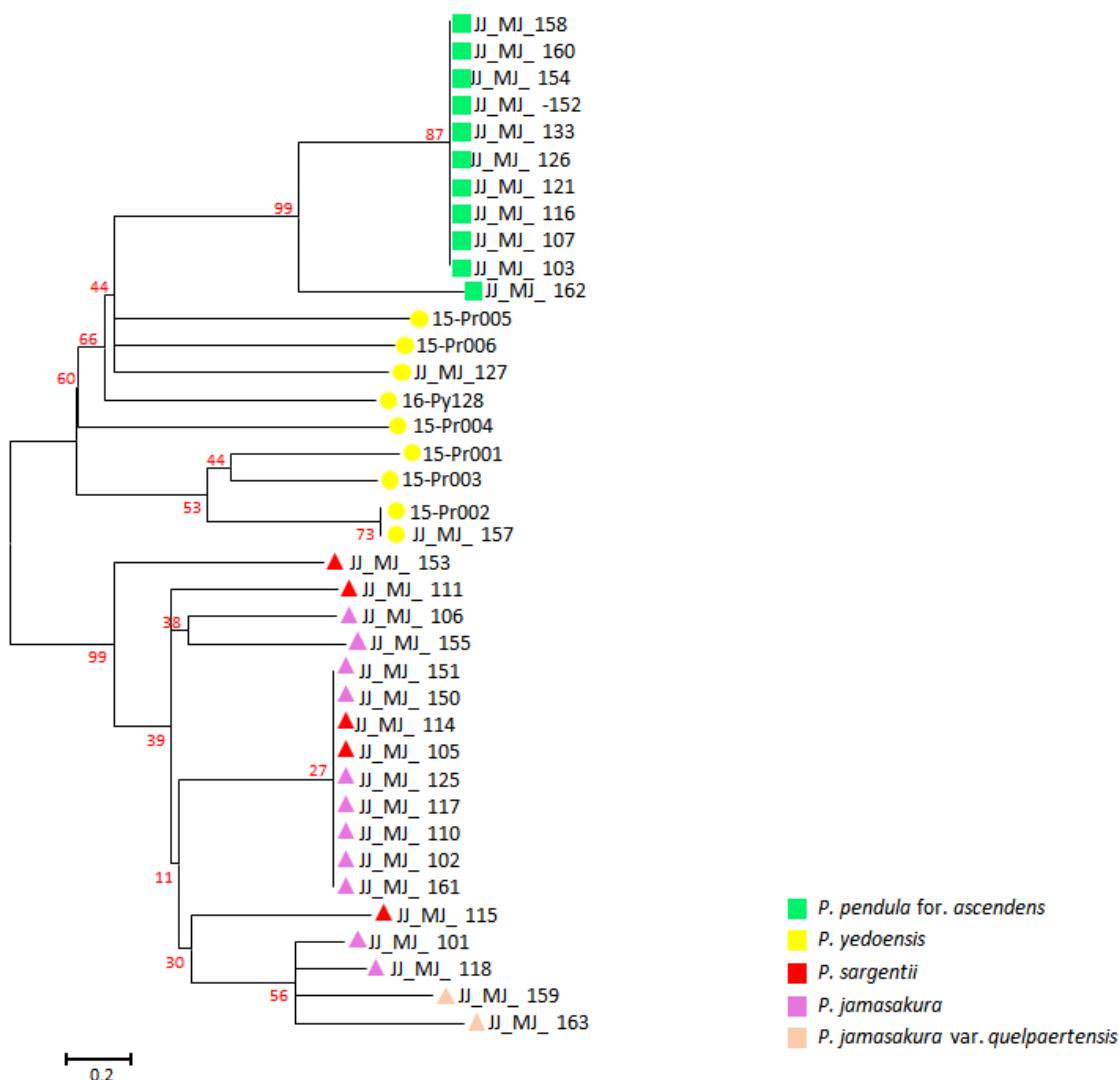


Figure 22. A Maximum Likelihood tree of *Prunus* species in Kwaneumsa temple and Bonggae-dong populations based on 20 PrCOS markers

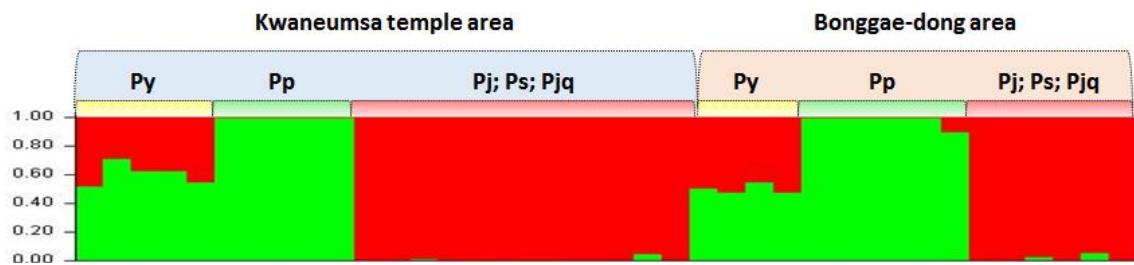


Figure 23. Characterization of genetic structure of wild *Prunus* species in Kwaneumsa temple area and Bonggae-dong area based on population structure analysis using 20 PrCOS markers

Py, *P. yedoensis*; Pp, *P. pendula* for. *ascendens*; Pj, *P. jamasakura*; Ps, *P. sargentii*; Pjq, *P. jamasakura* var. *quelpaertensis*

제 5 절 제주도 자생 왕벚나무의 부계 추정

1. 왕벚나무 이형성 PrCOS 분자 마커의 목표 영역 서열 분석

가. 왕벚나무 이형성 PrCOS 분자 마커의 목표 영역 서열 추출

왕벚나무와 근연종 벚나무류의 서열 유사도를 비교하고 왕벚나무의 부계를 추정하기 위하여 우선 왕벚나무 이형성 PrCOS 분자 마커의 목표 영역을 탐색하였다. 왕벚나무 유전구조 분석에 사용한 왕벚나무 이형성 PrCOS 분자 마커 20개의 핵 유전체 조립 서열과 왕벚나무 유전자 모델에 대하여 E-value cutoff $1E^{-10}$ 의 조건에서 BLASTN을 수행하였다. 그 결과, 왕벚나무 이형성 PrCOS 분자 마커 11개의 목표 영역 서열을 유전자 모델에서 정의할 수 있었다. 하지만 왕벚나무 이형성 PrCOS 분자 마커 9개의 목표 영역 서열은 왕벚나무 유전자 모델과 대응되지 않아 핵 유전체 조립 서열을 그대로 사용하였다. 왕벚나무 이형성 PrCOS 분자 마커 11개의 목표 영역 서열을 왕벚나무 유전자 모델에서 추출하였으며, 나머지 PrCOS 분자 마커 9개의 핵 유전체 조립 서열을 추출하였다. 왕벚나무 이형성 PrCOS 분자 마커의 목표 영역 서열을 왕벚나무 유전자 모델에서 정의한 11개는 Table 8에서 별표(*)로 표시하였다.

나. 왕벚나무 및 근연종 벚나무류의 공통서열 추출

왕벚나무의 부계를 추정하기 위하여 왕벚나무와 근연종 벚나무류의 개체별 핵 유전체 Illumina short read 서열을 기반으로 한 공통서열을 추출하였다. 제주도 자생 왕벚나무 기념목 지정 4개체와 왕벚나무 근연종 벚나무류 4종에 대하여 각각 1개체씩 핵 유전체 단편 서열(read)을 왕벚나무 유전자 모델에 매핑하였고 각 개체별로 핵 유전체 단편 서열을 기반으로 한 공통서열을 만들었다. 왕벚나무와 근연종 벚나무류의 공통서열에서 PrCOS 분자 마커 11개의 목표 영역 유전자를 확인하고 해당 영역의 서열을 추출하였다. 한편 왕벚나무 유전자 모델에 대응하지 않은 PrCOS 분자 마커 9개에 대하여 왕벚나무 핵 유전체 조립 서열에 근연종 벚나무류 및 왕벚나무 핵 유전체 단편 서열을 매핑하고 핵 유전체 단편 서열을 기반으로 공통서열을 만들어 추출하였다.

다. 부계 대립 유전자 서열 선발

왕벚나무 부계를 추정하기 위하여 우선, PrCOS 분자 마커의 목표 영역 서열 20세트의 길이와 왕벚나무 PCR 반응산물의 아가로스 젤 전기영동 결과 크기를 비교하여 부계 추정 서열과 모계 추정 서열을 구분하였다. 아가로스 젤 전기영동 결과 PCR 밴드의 크기를 완벽하게 추정할 수 없으므로 부계 후보종의 핵 유전체 단편 서열과 모계 후보종의 핵 유전체 단편 서열의 매핑 개수(read coverage)를 확인하여 최종적으로 부계 대립 유전자 서열과 모계 대립 유전자 서열을 구분하였다.

2. 서열 일치도 분석

가. 서열 일치도에 근거한 부계 추정

왕벚나무 모계 대립 유전자 서열과 부계 대립 유전자 서열을 각각 모계인 올벚나무의 공통 서열과 부계 후보종인 벚나무와 산벚나무 그리고 사옥의 공통서열을 정렬하였다. 정렬된 서열의 Sequence Identify Matrix를 제작하여 각 개체별 서열 일치도를 비교하였다. Table 13에서 보는 것과 같이 왕벚나무의 모계 대립 유전자 서열과 올벚나무의 공통서열은 평균 0.99의 서열 일치도를 보였다. 올벚나무는 왕벚나무 지정목 4개체에 대하여 모두 동일한 서열 일치도를 나타내어 엽록체 DNA 뿐만 아니라 핵 유전체 서열 결과에서도 왕벚나무의 모계가 올벚나무라는 사실을 다시 한 번 확인 할 수 있었다. 반면 왕벚나무의 부계 대립 유전자 서열과 부계 후보종의 공통서열을 비교한 결과, 왕벚나무와 벚나무는 평균 0.99, 사옥은 0.98 그리고 산벚나무는 0.93의 서열 일치도를 보였다. 왕벚나무와 벚나무에서는 개체별로 약 0.98에서 약 0.99의 유사도를 보였으며 왕벚나무와 사옥도 비슷한 양상을 보였다. 그러나 왕벚나무와 산벚나무의 경우 약 0.92에서 약 0.93의 서열 일치도를 보여 왕벚나무의 부계 후보종 가장 낮았다. 따라서 왕벚나무의 부계 서열 유사도에 근거하여 왕벚나무의 부계는 벚나무가 가장 유력하다고 사료된다.

Table 13. Sequence identity of the paternal and maternal allele between *P. yedoensis* and its closely relative species

Maternal allele	<i>P. yedoensis</i>				Average
	159-1	159-2	159-3	PY128	
<i>P. pendula</i> for. <i>ascendens</i>	0.99	1.00	0.99	0.99	0.99
Paternal allele	<i>P. yedoensis</i>				Average
	159-1	159-2	159-3	PY128	
<i>P. jamasakura</i>	0.98	0.99	0.99	0.98	0.99
<i>P. jamasakura</i> var. <i>quelpaertensis</i>	0.98	0.99	0.98	0.98	0.98
<i>P. sargentii</i>	0.92	0.93	0.93	0.92	0.93

제 4 장 결 론

본 연구에서는 중국 매실과 복숭아 그리고 왕벚나무의 유전체 정보를 이용하여 벚나무속 COS 유전자 1,303개를 발굴하였다, 발굴한 COS 유전자를 이용하여 벚나무속에 공통적으로 활용 가능한 유전자 기반 범용 분자 마커 156개를 설계하였으며 그 중 왕벚나무 종판별이 가능한 분자 마커 38개를 개발하였다. 개발한 COS 분자 마커는 왕벚나무의 종판별과 제주도 자생 왕벚나무 및 근연종 벚나무류의 유전구조의 파악과 집단 간 계통 분석에 이용하였다.

제주도 자생 왕벚나무의 유전형을 분석한 결과, 왕벚나무의 대부분은 모계 추정종인 올벚나무와 부계 후보종인 벚나무, 산벚나무, 사옥의 유전형을 50%씩 가지고 있는 이형접합체로 판별되었다. 일부 개체들은 모계 혹은 부계 후보종의 유전형을 약 75%씩 가지고 있었으며 이는 이형접합체에 모계나 부계의 유전형의 개체가 역교배되어 이들의 유전형이 추가로 유입된 것으로 판단할 수 있었다. 특히 한라산 백록담을 기준으로 제주도 북동 지역(JJ-NE)에서는 왕벚나무에 모계 추정종인 올벚나무의 유전자가 추가 유입된 개체가, 제주 남동 지역(JJ-SE)에서는 왕벚나무에 부계 추정종인 벚나무, 산벚나무 또는 사옥의 유전자가 추가 유입된 개체가 구분되어 나타났다. 이는 지역의 일조량, 기온 등의 차이에 의해 벚나무류의 개화 조건이 달라져 지역적으로 서로 다른 왕벚나무 교잡종이 형성되었기 때문으로 추정되며 이에 대한 추가적인 연구가 필요하다고 사료된다.

제주도 자생 왕벚나무와 근연종 벚나무 수종 중에서 모계 추정종인 올벚나무와 부계 추정종인 벚나무, 산벚나무, 사옥과의 계통을 비교 분석하였다. 그 결과 왕벚나무와 올벚나무는 각각 독립된 집단으로 묶여 다른 벚나무류와 구분되었다. 그러나 벚나무, 산벚나무, 사옥은 혼합되어 있는 하나의 집단으로 묶여 왕벚나무와 올벚나무와 구분되었다. 한편 왕벚나무 종 내에서 자생 왕벚나무와 재배 왕벚나무는 구분되지 않았다.

제주도 자생 왕벚나무 기념목 지정 4개체와 근연종 벚나무류인 올벚나무, 벚나무, 산벚나무 그리고 사옥의 핵 유전체 단편 서열을 이용하여 왕벚나무 두 개의 대립 유전자에 대한 모계 추정 공통 서열과 부계 추정 공통 서열을 만들었다. 이후 왕벚나무 모계 서열과 올벚나무의 서열 그리고 왕벚나무 부계 서열과 벚나무, 산벚나무 그리고

사옥의 서열을 비교하여 서열 일치도 분석을 수행하였다. 왕벚나무의 모계인 올벚나무의 서열과 왕벚나무의 모계 대립 유전자 서열의 일치도는 평균 0.99로 매우 높은 반면, 왕벚나무의 부계 후보종인 벚나무와 사옥 그리고 산벚나무의 대립 유전자 서열의 일치도는 각각 평균 0.99, 0.98 그리고 0.93으로 벚나무의 서열이 가장 유사한 것으로 판별되었다. 이상의 결과로 볼 때 왕벚나무의 모계는 올벚나무로 확인하였으며 부계는 벚나무가 가장 유력한 것으로 판단하였다.

참 고 문 헌

- Brunner AM, Busov VB, Strauss SH (2004) Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends Plant Sci* 9: 49-56
- Cho M-S, Kim C-S, Kim S-H, Kim S-C (2016) Comparative phylogenetic relationship between wild and cultivated *Prunus yedoensis* Matsum. (Rosaceae) with regard to Taquet's collection. *Korean J Pl Taxon* 46: 247-255
- Cho M-S, Kim C-S, Kim S-H, Kim TO, Heo K-I, Jun J, Kim S-C (2014) Molecular and morphological data reveal hybrid origin of wild *Prunus yedoensis* (Rosaceae) from Jeju Island, Korea: Implications for the origin of the flowering cherry. *Am J Bot* 101: 1976-1986
- Choi H-K, Mun J-H, Kim D-J, Zhu H, Baek J-M, Mudge J, Roe B, Ellis N, Doyle J, Kiss GB, Young ND, Cook DR (2004) Estimating genome conservation between crop and model legume species. *Proc Natl Acad Sci USA* 101: 15289-15294
- Cronquist A (1981) An integrated system of classification of flowering plants. Columbia University Press 34: 268
- Doganlar S, Frary A, Daunay M-C, Lester RN, Tanksley SD (2002) A comparative genetic linkage map of eggplant (*Solanum melongena*) and its implications for genome evolution in the Solanaceae. *Genet* 161: 1697-1711
- Echt CS, Nelson CD (1997) Linkage mapping and genome length in eastern white pine (*Pinus strobus* L.). *Theor Appl Genet* 94: 1031-1037
- Falush D, Stephens M, Pritchard JK (2003) Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies. *Genet* 164: 1567-1587
- Gray A (1867) Manual of the botany of the northern United States. Ivison & Company
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 95-98
- Harn C, Kim Y, Yang S, Chung H (1977) Studies on the origin of *Prunus yedoensis* Matsumura. 1. A comparative electrophoretic study on wild *Prunus subhirtella* in Mt. hanla, cultivated *Prunus yedoensis* and *Prunus donarium*. *Korean J Bot* 101: 1976-1986
- Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Resour* 9: 1322-1332

- Iketani H, Ohta S, Kawahara T, Katsuki T, Mase N, Sato Y, Yamamoto T (2007) Analyses of clonal Status in lsquo; Somei-yoshino and confirmation of genealogical record in other cultivars of *Prunus yedoensis* by microsatellite Markers. Breed Sci 57: 1-6
- Jeong Y-H, Go M-H, Kim G-O, Jeong Y-C, Kim M-H (1998) Phylogenetic relationship among selected taxa of *Prunus* in Mt. Halla and cultivated *Prunus yedoensis* by the RAPD analysis. Korean J Pl Taxony 27: 415-415
- Jung Y, Ko M, Oh Y, Han S, Kim M, Koh S, Oh M (1998) Genetic relationship of *Prunus yedoensis*, native and cultivar, based on internal transcribed spacer sequences of ribosomal DNA. Korean J Genet 20: 109-116
- Jung YH, Han S-H, Oh M-Y (2002) Phylogenetic analysis of Korean *Prunus* (Rosaceae) based on ITS sequences of nuclear ribosomal DNA. Korean J Genet 24: 247-258
- Kalendar R, Lee D, Schulman AH (2009) FastPCR software for PCR primer and probe design and repeat search. G3 (Bethesda) 3: 1-14
- Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28: 1647-1649
- Khanuja SP, Shasany AK, Darokar MP, Kumar S (1999) Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. Plant Mol Biol Rep 17: 74-74
- Kim M-H, Kim C-S, Lee G-Y, Mun M-O, Hyeon J-H, Im B-S (1998) Natural habitat of *Prunus yedoensis* Matsumura and its morphological variation. Korean J Pl Taxony 28: 117-117
- Kim M (1985) Flora of vascular plant in Cheju-do. Seoul: Report of the academic survey of Halasan (Mountain) natural preserve: 246-298
- Koehne E (1912) *Prunus yedoensis* var. *nudiflora*, nov. var. Feddes Repert. 10: 507-507
- Koidzumi G (1932) *Prunus yedoensis* Matsum. is a native of Quelpaert. APG 1: 177-179
- Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol doi: 10.1093/molbev/msw054

- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth* 9: 357-359
- Larkin MA, Blackshields G, Brown N, Chenna R, McGgettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948
- Lee S, Wen J (2001) A phylogenetic analysis of *Prunus* and the Amygdaloideae (Rosaceae) using ITS sequences of nuclear ribosomal DNA. *Am J Bot* 88: 150-160
- Liewlaksaneeyanawin C, Zhuang J, Tang M, Farzaneh N, Lueng G, Cullis C, Findlay S, Ritland CE, Bohlmann J, Ritland K (2009) Identification of COS markers in the *Pinaceae*. *Tree Genet Genomes* 5: 247-255
- Matsumura J (1901) *Cerasi japonicae duae species novae*. *J Plant Res* 15: 99-101
- Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8: 4321-4326
- Park S-H, Kim M-H, Kim K-S, Lee S-T (1984) 漢拏山에 自生하는 벚나무數種의 花粉學的 研究(A palynological study of some *Prunus* species in Mt. Halla.) *Korean J Pl Taxon* 14: 153-159
- Potter D, Eriksson T, Evans RC, Oh S, Smedmark J, Morgan DR, Kerr M, Robertson KR, Arsenault M, Dickinson TA (2007) Phylogeny and classification of Rosaceae. *Plant Syst Evol* 266: 5-43
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genet.* 155: 945-959
- Quraishi UM, Abrouk M, Bolot S, Pont C, Throude M, Guilhot N, Confolent C, Bortolini F, Praud S, Murigneux A, Charmet G, Salse J (2009) Genomics in cereals: from genome-wide conserved orthologous set (COS) sequences to candidate genes for trait dissection. *Funct Integr Genomics* 9: 473
- Roh MS, Cheong EJ, Choi I-Y, Joung YH (2007) Characterization of wild *Prunus yedoensis* analyzed by inter-simple sequence repeat and chloroplast DNA. *Sci Hortic* 114: 121-128
- Takenaka Y (1963) The origin of the Yoshino cherry tree. *J Hered* 54: 207-211
- Verde I, Abbott AG, Scalabrin S, Jung S, Shu S, Marroni F, Zhebentyayeva T, Dettori MT, Grimwood J, Cattonaro F, Zuccolo A, Rossini L, Jenkins J, Vendramin E, Meisel LA, Decroocq V, Sosinski B, Prochnik S, Mitros T,

Policriti A, Cipriani G, Dondini L, Ficklin S, Goodstein DM, Xuan P, Fabbro CD, Aramini V, Copetti D, Gonzalez S, Horner DS, Falchi R, Lucas S, Mica E, Maldonado J, Lazzari B, Bielenberg D, Pirona R, Miculan M, Barakat A, Testolin R, Stella A, Tartarini S, Tonutti P, Arus P, Orellana A, Wells C, Main D, Vizzotto G, Silva H, Salamini F, Schmutz J, Morgante M, Rokhsar DS (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet* 45: 487-494

Wheelan SJ, Church DM, Ostell JM (2001) Spidey: A tool for mRNA-to-genomic alignments. *Genome Res* 11: 1952-1957

Wilson EH (1916) The cherries of Japan. Printed at the University Press

Zhang Q, Chen W, Sun L, Zhao F, Huang B, Yang W, Tao Y, Wang J, Yuan Z, Fan G, Xing Z, Han C, Pan H, Zhong X, Shi W, Liang X, Du D, Sun F, Xu Z, Hao R, Lv T, Lv Y, Zheng Z, Sun M, Luo L, Cai M, Gao Y, Wang J, Yin Y, Xu X, Cheng T, Wang J (2012) The genome of *Prunus mume*. *Nat Commun* 3: 1318

이창복 (1996) 왕벚나무의 自生地. 서울大學校 樹木園 研究報告 16: 50-54

한창열 (1964) 속보(速報): 한라산 자생의 왕벚 및 추정양친에(推定兩親) 관하여. *J. Plant Biol* 7: 14-16

한창열 (1965) 보문(報文): 한라산 자생 왕벚 및 추정양친에(推定兩親) 관한 연구. *J. Plant Biol* 8: 11-18

한창열, 김영진, 양서영, 정희준 (1977) 보문 (報文): 왕벚의 기원에 관한 연구. 한라산 자생 올벚, 재배 왕벚 및 산벚의 전기영동형 (電氣泳動型) 비교. *J. Plant Biol* 20: 1-5

부 록

제주도 자생 왕벚나무의 양친을 추정하기 위하여 PrCOS 마커의 목표 구간에 대한 서열 일치도를 분석하였다. 왕벚나무 159-1번 개체, 159-2번 개체, 159-3번 개체, 그리고 올벚나무, 벚나무, 산벚나무 그리고 사옥을 각각 1개체씩 sequencing한 Illumina read를 왕벚나무 핵 유전체 초안에 매핑하여 공통서열을 만들어 정렬하였다. 이후 정렬한 서열은 자생 왕벚나무의 유전형 분석에 사용한 20개 마커의 목표 영역별로 추출하고, 모계 서열(M)과 부계 서열(P)로 나누었다. 핵 유전체 해독 개체인 159-2번 개체를 기준으로 서열 일치도를 표기하였으며 기준 개체와 같은 염기서열은 점(.)으로, 서로 다른 염기서열은 염기서열로 표기하였다. 프라이머 영역은 노란색으로 표시하였다.

PrC-0024



PrC-0054

159-2_M	T	G	A	C	G	T	C	A	G	C	A	T	G	T	T	G	T	C	A	A	T	C	T	G	A	G	C
159-1_M	
159-3_M		
Pp_M		
10 20 30 40 50 60 70 80																											
159-2_M	A	T	T	G	A	G	A	G	T	G	C	T	A	G	T	C	A	T	T	C	A	T	T	G	A	C	G
159-1_M		
159-3_M			
Pp_M			
90 100 110 120 130 140 150 160																											
159-2_M	G	G	A	A	C	T	C	G	G	C	C	A	C	T	C	G	T	A	T	T	C	T	T	G	C	T	T
159-1_M		
159-3_M			
Pp_M			
170 180 190 200 210 220 230 240																											
159-2_M	C	A	T	A	A	A	A	G	T	A	T	T	T	T	C	T	G	A	A	T	A	A	T	T	T	T	T
159-1_M		
159-3_M			
Pp_M			
250 260 270 280 290 300 310 320																											
159-2_M	T	A	G	A	G	G	G	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	G	G	A	A
159-1_M		
159-3_M			
Pp_M			
330 340 350 360 370 380 390 400																											
159-2_M	T	A	G	A	G	G	G	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	G	G	A	A
159-1_M		
159-3_M			
Pp_M			
410 420 430 440 450																											
159-2_P	G	T	T	G	G	A	C	A	G	T	T	C	A	A	T	T	C	A	T	G	T	T	G	T	A	A	C
159-1_P		
159-3_P			
Pj_P			
Ps_P			
Pjq_P			
10 20 30 40 50 60 70 80																											
159-2_P	T	G	A	C	G	T	C	A	G	T	T	G	C	A	A	T	T	C	A	T	T	G	T	A	C	G	
159-1_P			
159-3_P				
Pj_P			
Ps_P			
Pjq_P			
90 100 110 120 130 140 150 160																											
159-2_P	G	G	A	A	C	T	C	G	G	C	C	A	C	T	C	G	T	A	A	T	A	A	T	T	G	C	T
159-1_P			
159-3_P				
Pj_P			
Ps_P			
Pjq_P			
170 180 190 200 210 220 230 240																											
159-2_P	C	A	T	A	A	A	A	G	T	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
159-1_P			
159-3_P				
Pj_P			
Ps_P			
Pjq_P			
250 260 270 280 290 300 310 320																											

PrC-0131

159-2_P AGGATTCTCTGCTGTCAGTGCTTGTCAAAATTCTGACACACATTCAAGTAACATATGGGATTGGGATTCTATT
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

159-2_P AAGGAGGAAAGATCCACTTCTAGGTGTTTTAGTACTTTTTAAATAATGTTAACTATTCTGAAAAGTTACAAAGA
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

159-2_P AAAGAGGAACCTTACTCCCTTTTATGCTGCTGCATCATTAAATAAGTGTACAACATGGCATTATTCAATCATACTTA
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

159-2_P AATCTTAGTCTACAGGATGGTCATTGGCA
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

159-2_M AGGATTCTCTGCTGTCAGTGCTTGTCAAAATTCTGACACACCTTCAAGTAACATATGGGATTGGGATTCTCTT
159-1_M
159-3_M
Pp_M

159-2_M ACAGTGCCTTGGATGAGGAAATTAAAGTACAAAGGATTCAATAATGACGGAATTAAATGACGGAATTAAATGTT
159-1_M
159-3_M
Pp_M

159-2_M CTAGAAATTGAAAGTTCTTGTGTTGGATTCTCTCTGGATTAGAAATGACGCACTTCAAGATAATTAAAGTT
159-1_M
159-3_M
Pp_M

159-2_M GGGATTTATACTCTAAATTCCATGTTTTCCACTGGAAATTCTCAAATTCTATGTTTTTAAATCAAAACACGGG
159-1_M
159-3_M
Pp_M

159-2_M AATTGGTTCATGTCATTTAGAAATTCTGAGTTTGTCAAAATCCAGTTATTTCCCTCATCCAAATGCACCAAGAAA
159-1_M
159-3_M
Pp_M

159-2_M AGATCCACTCTCGGTGTTTTAGTACTTTTTAAATAATGTTAACTATTCTGAAAAGTTACAAAGAAAAAGGGAA
159-1_M
159-3_M
Pp_M

159-2_M CCTTACTCCCTTTTATGCTTATCATCATTAAATAAGTGTACAACACATTGCAATTATTCAATCATACTTAAATCTAAGT
159-1_M
159-3_M
Pp_M

159-2_M CTACAGGATGGTCATTGGCA
159-1_M
159-3_M
Pp_M

PrC-0138

159-2_M	ACCATAGCAGCAAAGGACCTG	CAGAAGAAA	TTCTG	TGAAA	ATCCTG	TTGAA	ATTGAA	ATTA	ATCCAGCAGCTAATT	TTA	10	20	30	40	50	60	70	80
159-1_M
Pp_M
159-2_M	CGTAAAAGGGGATG	CTGGT	GATTATGAGAAGGG	TATATTGAAA	ACTGCA	ATATA	TTGGGAT	TTCTGG	TTATCATG	A	90	100	110	120	130	140	150	160
159-1_M
159-3_M
Pp_M
159-2_M	AACTGAAA	TTGAAGT	TTAGCT	GGGACG	ACACAG	CAAGGA	ACTCTA	CAGAGA	ATTTTTG	C	170	180	190	200	210	220	230	240
159-1_M
159-3_M
Pp_M
159-2_M	ATTTGGA	ATCTG	CTAGT	CGTCT	CTGAGCT	CCAAAGT	TCGAT	CCAGT	ATTCAG	C	250	260	270	280	290	300	310	320
159-1_M
159-3_M
Pp_M
159-2_M	TATGAGAAGG	ATAAAACTG	AAAGGTTTG	ATATAA	TTGGGATG	ATGGT	ATCATG	AGAAAGCAGG	TAAGTATG	T	330	340	350	360	370	380	390	400
159-1_M
159-3_M
Pp_M
159-2_M	TGGGATG	ACTACAGG	AAAGCAT	CTCACAGA	AAAAATG	TTGTTG	C	410	420	430	440	450	460	470	480	490	500	510
159-1_M
159-3_M
Pp_M
159-2_M	TCTCTGAG	CTCCAAG	TTGATCC	CAGTAT	CCAGCT	CGCATG	AAAGTTGG	GTATT	CAGGTG	A	520	530	540	550	560	570	580	590
159-1_M
159-3_M
Pp_M
159-2_M	AGGTTTG	ATATAATT	GGGGATG	ATGGTG	ATCATG	AGAAAAT	CAGATA	AACTG	TTTTA	ACTGGG	50	510	520	530	540	550	560	570
159-1_M
159-3_M
Pp_M
159-2_M	ATCTACAG	AGAAAAAT	TTGTTG	CAGAGAA	ATCCT	TATAACT	AGTTT	GGAA	ATCTGCC	AGTAG	580	590	600	610	620	630	640	650
159-1_M
159-3_M
Pp_M
159-2_M	TAGCAAAA	TTTAAATG	TGATG	ACCC	TGGCGAT	CATGAT	GGTGG	TGAA	ATTG	AAAAGAG	730	740	750	760	770	780	790	800
159-1_M
159-3_M
Pp_M
159-2_M	AGAAGGG	AAAGGAC	CTGTTG	AAAATCT	TTTTG	G	810	820	830	840	850	860	870	880	890	900	910	920
159-1_M
159-3_M
Pp_M

PrC-0138

159-2_P	ACCATAGCAGCAAAGGACCTGCAGAAGAAAATTCTGTGGAAAATCCTGCTGAAATCGAAATTAAATCCAGCAGCTAATTAA	10 20 30 40 50 60 70 80
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	CATGAAAATGGGGTGCCTGGTGATTATGAGAAGGGTATATTGAAACCTGCAATATAATTGGAGATTCTGGTTATCATGA	90 100 110 120 130 140 150 160
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	AACTGAAAATTTAAGTTAGCTGGGATGACAACAGCAAGGAATCTACAGAGAAATTTTTTGCTAAAGACCTGTGAGTA	170 180 190 200 210 220 230 240
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	ATTTGGAATCTGCTAATTGTCCTCTGAGCTCCAAAGTCATCCAGTATCCAGCTCACATGAAGTTGGTGATGCAGGTGAT	250 260 270 280 290 300 310 320
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	TATGAGAAGGATAAAACTGAAAGGTTTGATATAATTGGGATGACGGTGATCATGAGAAAGCTGAAAAGGATGTTTAACTG	330 340 350 360 370 380 390 400
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	TGGGATGACTACAGGAAAGTATCTACAGAGAAAATTGTTGCAGAGAAATACTATAACTAGTTGGAAATCTGCCAGTAGTC	410 420 430 440 450 460 470 480
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	TCTCCAAACTCAAGCTAATTAGCAACAAATTAAATGATGATGACCCCTGGCGATCATGATGGTGGTGAAATTGAAAAG	490 500 510 520 530 540 550 560
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		
159-2_P	AGTATTATATCTGGAAATGAGAGCAGGGAGGACCTGTTGAAAAAAATCTTTGGTGAAGCCAACTTGAACCTTG	570 580 590 600 610 620 630 640
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
.....		

PrC-0179

159-2_M	CGAGC ATCTGAAGAACAGTG ACCATAATTATAAGAATTGCAGGTTATCTTTCTCC TATTTATTTCGTAATA	10	20	30	40	50	60	70	80
159-1_M
Pp_M
159-2_M	TACTATGGTCAGGCAGGTTTTCTTCTGACCCATGCCTAACCAAGGAGTCATTAGATGTCTTCTACTGCCTTGCA	90	100	110	120	130	140	150	160
159-1_M
Pp_M
159-2_M	TTTTCTTGTGGGTTCTGTTAAGTTCTGCAGTTAGTGCCTACTATGTTGGATGCTGTCTTAGTTATGTGTT	170	180	190	200	210	220	230	240
159-1_M
Pp_M
159-2_M	TATGTGGAAGTAGTACCATCAC-TATTGAAATTCTTAATTATGCAGTATGTTAGGTTATGCCTATATCTCAGA	250	260	270	280	290	300	310	320
159-1_M
159-3_M
Pp_M
159-2_M	TGAGGATTAGCTTTAGGTTGCTATAATGTCGACATTCTATACTACTATGTTAGTAGTCGGTATTACCCTTGTGGT	330	340	350	360	370	380	390	400
159-1_M
159-3_M
Pp_M
159-2_M	GATAATTTCTTTAGTGAAGGGTTCTGGTGGAAATGTTGGTAGCATTTGCTTTGGATGAATAACATGTTCAAGTAA	410	420	430	440	450	460	470	480
159-1_M
159-3_M
Pp_M
159-2_M	G TGATGCATGAGTCGACCTAAAAAGCTCTCAAGTGCATGGGCTGGTGAGTCTTGATGAGATGTTCTGCTGG	490	500	510	520	530	540	550	560
159-1_M
159-3_M
Pp_M
159-2_M	T CCTATAAGACAATAAGTAATAGTTGAATCAAGGGACAATGTTTCACTATCTGTCAAAGTTATTACTATTATCTGT	570	580	590	600	610	620	630	640
159-1_M
159-3_M
Pp_M
159-2_M	TACTTGTGTTCTGGGACTATTACAGTACTATCTTAAGCATACTTACAGAAGAAATCCGTTTCACCTCTGATCTGATGTGT	650	660	670	680	690	700	710	720
159-1_M
159-3_M
Pp_M
159-2_M	GACTTTATCAGCATCAATGTAAGTGCAGGAAAAATTAGAAATTTTTATCTTGATCTGATGCCACCAAGATTTTAGAA	730	740	750	760	770	780	790	800
159-1_M
159-3_M
Pp_M
159-2_M	AAGCCAAGTGTAAAAATCACATTCTAATGAATAACCTACTGACTGTAACCTTATCTAGGTTTTGGAAATGTAT	810	820	830	840	850	860	870	880
159-1_M
159-3_M
Pp_M
159-2_M	CTATGATTTCTATTTGCTTCAAGGAAAGCTGTCGGATT	890	900	910	920
159-1_M
159-3_M
Pp_M

PrC-0179



PrC-0237



PrC-0237

PrC-0260

159-2_M	CATATGAAGGGCTTCTGATGAGGTAAGTTGTCGTAAAAACGTGAAGGAATAATTATTCA	TGGCATGACACAACAAAAAG	10	20	30	40	50	60	70	80
159-1_M
159-3_M
Pp_M
159-2_M	CTCCGTGAGAGTAGGAAGCTAGGTGAGCAAGTACATGTGAAATTGGCTCATGAGGTATTTCT	CATCTGGGTG	90	100	110	120	130	140	150	160
159-1_M
159-3_M
Pp_M
159-2_M	TTGCATACCAAGTTCAATTGGTATGGTAAAAAGAAAAAAAGTAAAGGTTTTTTCCAAAGAACTTTAATTATTGGGG	TTT	170	180	190	200	210	220	230	240
159-1_M
159-3_M
Pp_M
159-2_M	GTTTTTGAACCTGAATGGAAACAATTCA	TGTGTTGGTAACTGTAAGCTATCCCTAAATAGTTTCAGTACAGT	250	260	270	280	290	300	310	320
159-1_M
159-3_M
Pp_M
159-2_M	TCTCGCATATCAATAATGGCTGTCGAGCTTGTGATGCATCTGTGATGGAGTCCAAACAAAGAGGG	TTT	330	340	350	360	370	380	390	400
159-1_M
159-3_M
Pp_M
159-2_P	CATATGAAGGGCTTCTGATGAGGTAATCGTCGTAAAAACGTGAAGGAATAATTATTCA	TGGCATGACACAACAAAAAG	10	20	30	40	50	60	70	80
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
159-2_P	CTCTGTGAGAGTAGGTAGCTAGGTGAGCAAGTACATGTGAAATTGGCTCATGAGGTATTTCTCATCTGGATT	TTT	90	100	110	120	130	140	150	160
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
159-2_P	GTTGCATACCAAGTTCAATTGGCATAGTGAAAAAGAAAAAAAGTAAAGGTTTTTTCCAAAGAACTTTAATTATTGGGG	TTT	170	180	190	200	210	220	230	240
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
159-2_P	GGTTTTGAATTTTTCTTAATAATCATCTGTGTTCTATTGGTTAACTAATACCCATTTATAAGGCATG	TTT	250	260	270	280	290	300	310	320
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
159-2_P	AGACATTTTTCTGATGGAAACAATTCA	TGTGTTGGTAATCTGAATTAGTTGTAAGCTATCCCTAAATAGTTCA	330	340	350	360	370	380	390	400
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
159-2_P	GTCACAGTTCTGCATATCAATAATGGTGCTGTCGAGCTTGTGATGCATCTGTCA	TGGAGTCCAAACAAAGAGGG	410	420	430	440	450	460	470	480
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

PrC-0263

159_2_M	GGTGTAGCTGCTCATTTCATGCTGTTACATCTTTGTCACCGTAAATGGTGCCTTCCCATACTCAAACCTTTAGTGTG	10	20	30	40	50	60	70	80
159_1_M
159_3_M
Pp_M
159_2_M	GATCCACATGATGGTACCTCACATTATTTTATTCATCTACATCATACTCCATCTGTTCCCTCCCTCTTCACTGTACT	90	100	110	120	130	140	150	160
159_1_M
159_3_M
Pp_M
159_2_M	TATGAGGTATATGATTTATTATGGTTTGACAGGTATATTCTAGAGTTGAAACTGTGTTGCTGTCGGGAAGCTTCACTGAA	170	180	190	200	210	220	230	240
159_1_M
159_3_M
Pp_M
159_2_M	TCTCACTTGTTTAGCAAAAGAGTGACACCTGTTGGCTCTCCGTTCGCCTTGCTATGAAGAACTTTTACCATTC	250	260	270	280	290	300	310	320
159_1_M
159_3_M
Pp_M
159_2_M	CACAATGTTATATCCCTAACAGTTGATTATCTCAACGCTGCTTCTCGCACCAAGAA	330	340	350	360	370			
159_1_M
159_3_M
Pp_M
159_2_P	GGTGTAGCTGCTCATTTCATGCTTTACATCTCTTGTCCACCGTAAATGGTGCCTTCCCATACTCAAACCTTTAGTGTG	10	20	30	40	50	60	70	80
159_1_P
159_3_P
Pj_P
Ps_P
Pjq_P
159_2_P	GATCCACATGATGGTACCTCACATTATTTGTTCATCTACATCCTCTGAGGTCATTTCTCACTCTCACTCTCAAA	90	100	110	120	130	140	150	160
159_1_P
159_3_P
Pj_P
Ps_P
Pjq_P	NNN
159_2_P	TTACCTCTGCCCTTCCCTCTTATCTCCCTGACTTATGAGGTATATGATTATATTGGTTTGACAGGTATATCTA	170	180	190	200	210	220	230	240
159_1_P
159_3_P
Pj_P
Ps_P	S
Pjq_P	S
159_2_P	GAGTTGAAACTGTTGCTGTCGGGAAGCTTCACTGAATCTCACTTGTTAGCAAAGAGTGACACCTGTTGGTTCT	250	260	270	280	290	300	310	320
159_1_P
159_3_P
Pj_P
Ps_P
Pjq_P
159_2_P	CCGCTTCGCCCTGCTATGAAGAATCTTACCAATTACACACAACTGATATATCCCTAACAGTTGATTATCTCAACACTGCTTC	330	340	350	360	370	380	390	400
159_1_P
159_3_P
Pj_P
Ps_P
Pjq_P
159_2_P	TCTCGCACCAAGAA	410							
159_1_P
159_3_P
Pj_P
Ps_P
Pjq_P

PrC-0281

Sequence alignment of 159-2_M, 159-1_M, 159-3_M, and Pp_M across 1000 positions. The alignment shows high conservation with minor variations. A yellow box highlights a region from position 970 to 1020.

Position	159-2_M	159-1_M	159-3_M	Pp_M
10	CTGCTGCAAA	TGATG	TAGAGGATA	AGGAGGTTGAGGCTACT
20			CAAGCTTAAAGGG	CAAGCTTAAAGGG
30			TAATATTTC	TAATATTTC
40			CTAGCTGCMIT	CTAGCTGCMIT
50			TCAAACT	TCAAACT
60				
70				
80	G TACTGTT	CCCTACAT	TTAGTGTT	TTAATTTAAT
90	TTAGTGTT	ATTTAAT	TTAAT	ATGATRTAT
100	TTAAT	TTAAT	TTAAT	GGAAATTAAAG
110			AGCATTTG	TTGACTAAAT
120			TTGACTAAAT	TTGACTAAAT
130				
140				
150				
160				
170	CATTGCTTA	ATGATTTC	TTACCTT	GGTTTCCTT
180	TTACCTT	GGTTTCCTT	GGTTTCCTT	GGAAATTAAAG
190	GGTTTCCTT	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG
200	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG
210	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG
220	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG
230	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG
240	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG	GGAAATTAAAG
250	TGAAATGTAAT	GATTGCTT	TTAGTCATT	CAAACTTC
260	GATTGCTT	CAAACTTC	CAAACTTC	CAACCTAACT
270	CAAACTTC	CAACCTAACT	CAACCTAACT	AGTTGTTG
280	CAACCTAACT	AGTTGTTG	AGTTGTTG	GAAGTCGAAGA
290	AGTTGTTG	GAAGTCGAAGA	GAAGTCGAAGA	TCATTTCTAACAGTG
300	GAAGTCGAAGA	TCATTTCTAACAGTG	TCATTTCTAACAGTG	TCATTTCTAACAGTG
310	TCATTTCTAACAGTG	TCATTTCTAACAGTG	TCATTTCTAACAGTG	TCATTTCTAACAGTG
320	TCATTTCTAACAGTG	TCATTTCTAACAGTG	TCATTTCTAACAGTG	TCATTTCTAACAGTG
330	TACATTAAT	TRGCCCTT	TTCTT	GTATTAAT
340	TRGCCCTT	GTATTAAT	GTATTAAT	GAATCATTGTT
350	GTATTAAT	GAATCATTGTT	GAATCATTGTT	CTCTATCTGGG
360	GAATCATTGTT	CTCTATCTGGG	CTCTATCTGGG	CTTAAGCTGAGAGAGAGG
370	CTCTATCTGGG	CTTAAGCTGAGAGAGG	CTTAAGCTGAGAGAGG	ATGGTGTG
380	CTTAAGCTGAGAGAGG	ATGGTGTG	ATGGTGTG	TCAGAGGATTGGTGTG
390	ATGGTGTG	TCAGAGGATTGGTGTG	TCAGAGGATTGGTGTG	TCAGAGGATTGGTGTG
400	TCAGAGGATTGGTGTG	TCAGAGGATTGGTGTG	TCAGAGGATTGGTGTG	TCAGAGGATTGGTGTG
410	TGAAAATATCGT	ATGAAAGG	TAAGG	TAAGGAACTTGT
420	ATGAAAGG	TAAGG	TAAGG	TTCACTGTT
430	TAAGG	TTCACTGTT	TTCACTGTT	TTCACTGTT
440	TTCACTGTT	TTCACTGTT	TTCACTGTT	TTCACTGTT
450	TTCACTGTT	TTCACTGTT	TTCACTGTT	TTCACTGTT
460	TTCACTGTT	TTCACTGTT	TTCACTGTT	TTCACTGTT
470	TTCACTGTT	TTCACTGTT	TTCACTGTT	TTCACTGTT
480	TTCACTGTT	TTCACTGTT	TTCACTGTT	TTCACTGTT
490	AGGTTGGTAGG	CCCTTG	TGCCAA	TCTTGACT
500	GGTAGG	TGCCAA	TCTTGACT	TTGTTCTAG
510	TGCCAA	TCTTGACT	TTGTTCTAG	TTGTTCTAG
520	TCTTGACT	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG
530	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG
540	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG
550	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG
560	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG	TTGTTCTAG
570	TTATATGGT	ATTTAAT	CTTTT	CTTCTT
580	ATTTAAT	CTTTT	CTTCTT	CTTCTT
590	CTTTT	CTTCTT	CTTCTT	CTTCTT
600	CTTCTT	CTTCTT	CTTCTT	CTTCTT
610	CTTCTT	CTTCTT	CTTCTT	CTTCTT
620	CTTCTT	CTTCTT	CTTCTT	CTTCTT
630	CTTCTT	CTTCTT	CTTCTT	CTTCTT
640	CTTCTT	CTTCTT	CTTCTT	CTTCTT
650	CTTTGGTTGAG	ATTCT	CTTG	CTTCCAG
660	ATTCT	CTTG	CTTCCAG	CCACAG
670	CTTG	CTTCCAG	CCACAG	TTGACT
680	CTTCCAG	CCACAG	TTGACT	CTTGT
690	CCACAG	TTGACT	CTTGT	TTGAAATG
700	TTGACT	CTTGT	TTGAAATG	GAACAG
710	CTTGT	TTGAAATG	GAACAG	TTGACT
720	TTGAAATG	GAACAG	TTGACT	CTCTATCTAC
730	ACATTTCTAG	AAAGGAA	TCCTT	GATGCTT
740	AAAGGAA	TCCTT	GATGCTT	TTCATG
750	TCCTT	GATGCTT	TTCATG	AAATGTTT
760	GATGCTT	TTCATG	AAATGTTT	CTTCTT
770	TTCATG	AAATGTTT	CTTCTT	CTTCTT
780	AAATGTTT	CTTCTT	CTTCTT	CTTCTT
790	CTTCTT	CTTCTT	CTTCTT	CTTCTT
800	CTTCTT	CTTCTT	CTTCTT	CTTCTT
810	TTTCTTTT	GGCCCTT	GGGTTA	AAATTCTT
820	GGCCCTT	GGGTTA	AAATTCTT	CTTCTT
830	GGGTTA	AAATTCTT	CTTCTT	CTTCTT
840	AAATTCTT	CTTCTT	CTTCTT	CTTCTT
850	CTTCTT	CTTCTT	CTTCTT	CTTCTT
860	CTTCTT	CTTCTT	CTTCTT	CTTCTT
870	CTTCTT	CTTCTT	CTTCTT	CTTCTT
880	CTTCTT	CTTCTT	CTTCTT	CTTCTT
890	TGATAGAATG	TGACT	TTTGT	TCATAAA
900	TGACT	TTTGT	TCATAAA	TCAAATT
910	TTTGT	TCATAAA	TCAAATT	TGGTTT
920	TCATAAA	TCAAATT	TGGTTT	ATGGTT
930	TCAAATT	TGGTTT	ATGGTT	CATCAT
940	TGGTTT	ATGGTT	CATCAT	TGGTTG
950	ATGGTT	CATCAT	TGGTTG	TTGTTG
960	CATCAT	TGGTTG	TTGTTG	TTGTTG
970	TGGCTGCAGA	TAAAAACCGA	TACTGCA	ATTGGGTAAGAATG
980	TAAAAACCGA	TACTGCA	ATTGGGTAAGAATG	TGGAGAC
990	TACTGCA	ATTGGGTAAGAATG	TGGAGAC	CTGGCTACAACTCT
1000	ATTGGGTAAGAATG	TGGAGAC	CTGGCTACAACTCT	
1010	TGGAGAC	CTGGCTACAACTCT		
1020	CTGGCTACAACTCT			

PrC-0281

PrC-0304

159-2_M AGAGCCATGCAGGTCTCTTCAACCTTTGATAGTAGTGCTTGTGCTGAAACAGCAGGCTATCATCGAGAATTAAATG
159-1_M
159-3_M
Pp_M

159-2_M AAGGTTATCCTGTTGATAAATTATGCCTGTTAGCGGAATTTTTATAGATGCCTATGGTTAAATTATTGTTG
159-1_M
159-3_M
Pp_M

159-2_M AGTATGGTTGGTTAAATTAAACTGGTTTTGTTAGTTCAAGATCATTCGAGATTGGCCCCCTTCTTA
159-1_M
159-3_M
Pp_M

159-2_P AGAGCCATGCAGGTCTCTTCAACCTGTGATAGTAGTGCTTGTGCTGAAACAGCAGGCTATCATCGAGAATTAAATG
159-1_P
159-3_P
P_s_P
P_q_P
P_l_P

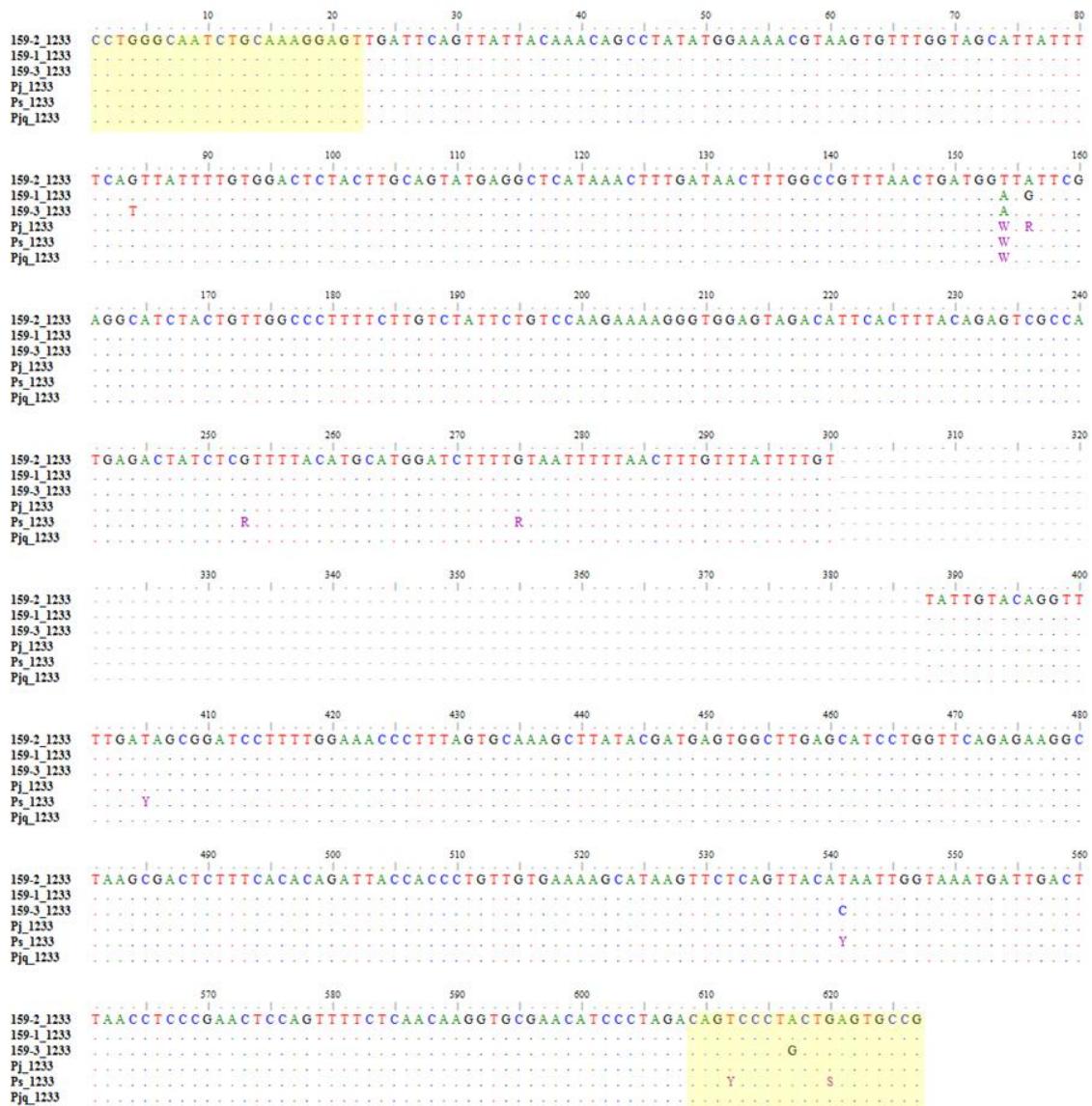
159-2_P AAGGTTTGCATATGCCTGATTCCTTGTGTTAGTTAAACTTCATGCCGTTAGCGGAATTTTTATAGATGCCTATG
159-1_P
159-3_P
P_s_P
P_q_P
P_l_P

159-2_P GTTTAATTATGTTGCAGTATGGTTGGTTAAATTAAACTGGTTTTGGTTAGTTCAAGATCATTCGAGATTGGCCCCCT
159-1_P
159-3_P
P_s_P
P_q_P
P_l_P

159-2_P TCTTA
159-1_P C
159-3_P
P_s_P
P_q_P
P_l_P

PrC-0383

159-2_M	C C T G G G C A A T C T G C A A A G G A G T T G A T T C A G T T A T T A C A A A C A G C C T A T A T G G A A A A C G T A A G T G T T G G T A G C A T T A T T	10	20	30	40	50	60	70	80
159-1_M
159-3_M
Pp_M
.....									
159-2_M	T C A G T T A T T T G T G G A C T C T A C T T G C T G T A T G A G G C T C A T A A A C T T G A T A A C T T T G G C G T T A A C T G A T G G A T A T T C G	90	100	110	120	130	140	150	160
159-1_M
159-3_M
Pp_M
.....									
159-2_M	A G G C A T C T A C T G T T G G C C C T T T C T G T C T A T T C T G T C C A A G A A A A G G G G G A G T A G A C A T T C A C T T A C A G A G T T G C C A	170	180	190	200	210	220	230	240
159-1_M
159-3_M
Pp_M
.....									
159-2_M	T G A G A C T A T C T C G T T T A C A T G C A T G G A T C T T T G T A A T T T T T A A C T T T G T T T A T T T G T T A T T A G C T T G C T G T G T G	250	260	270	280	290	300	310	320
159-1_M
159-3_M
Pp_M
.....									
159-2_M	T A T T G A A C C C T A T C T T T G G G C T T G G G A A T A A T G A T G A T G T T T G T A G G C T T T G A A T G A C A A T C C T T A A T A T T G T A C A G G T	330	340	350	360	370	380	390	400
159-1_M
159-3_M
Pp_M
.....									
159-2_M	T T G A T A G C G G A T C C T T T G G A A A C C C T T A G T G C A A A G C T T A T A C G A T G A G T G G C T T G A G G C A T C C T G G T T C A G A G A A G G C	410	420	430	440	450	460	470	480
159-1_M
159-3_M
Pp_M
.....									
159-2_M	T A A G C G A C T C T T C A C A C A G A T T A C C A C C C T G T T G T G A A A A G C A T A A G T T C T C A G T T A C A C A A T T G G T A A A T G A T T G A C T	490	500	510	520	530	540	550	560
159-1_M
159-3_M
Pp_M
.....									
159-2_M	T A A C C T C C C G A A C T C C A G T T T T C T C A A C A A G G T G C G A G C A T C C C T A G A C A G T C C C T A C T G A G G T G C C G	570	580	590	600	610	620		
159-1_M
159-3_M
Pp_M

PrC-0383


PrC-0386


PrC-0386

159-2_P TGGGGAACTTGTGGCTGAGAGT GACAAGCTAGCTTCTCAAGAAGTCGGAGATGGAGACGACAATGCGAGGAGGCGACGAC
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

90 100 110 120 130 140 150 160

159-2_P GTGAAAAAAATTGAAAGACTTGTCTCAAAAGATGGAGGTTGTTAGTTATTTCCTTCTTCATGTTTCTTGTAATAAC
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

170 180 190 200 210 220 230 240

159-2_P TATCTTAAAGAATTATTGTTCCAAATTGTTCTGAGTGCGATAATGTTGTTCTATGACATAATTGTCATAATG
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

250 260 270 280 290 300 310 320

159-2_P AAAATGCAAGGAAACATTAAACAAAGTTAACATCTGAAATTCAAAAGAAATTATTTCCTCCGGAGGGTTGTTGCAAGGCATT
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

330 340 350 360 370 380 390 400

159-2_P GTCCCCATCGTTTGCATGGCTTTATGATTCTCTGCTCACCTAAATAGATGTGTTAGTTAACATGATCCGCTTA
159-1_P
159-3_P
Pj_P T
Ps_P Y
Pjq_P

410 420 430 440 450 460 470 480

159-2_P GACAAAGTCATTGATTCGACCTGTAACACTCAAGGATGACGGTTGCTGGCTTCACTGCTGGCAATCACTGATTTTACTAATT
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

490 500 510 520 530 540 550 560

159-2_P TTCTATGGTTATGCTAGCATACCATCTCTTGTGTTCTGTTCTGCTTCATTCGAGTTGATGTCATACTAA
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

570 580 590 600 610 620 630 640

159-2_P ATTGTTGCTTGTACCATACCTCATGCATATTCTAATGTTGTCATTCTTTATGAGCTGATATGGTAAATGAAATGCAAGG
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

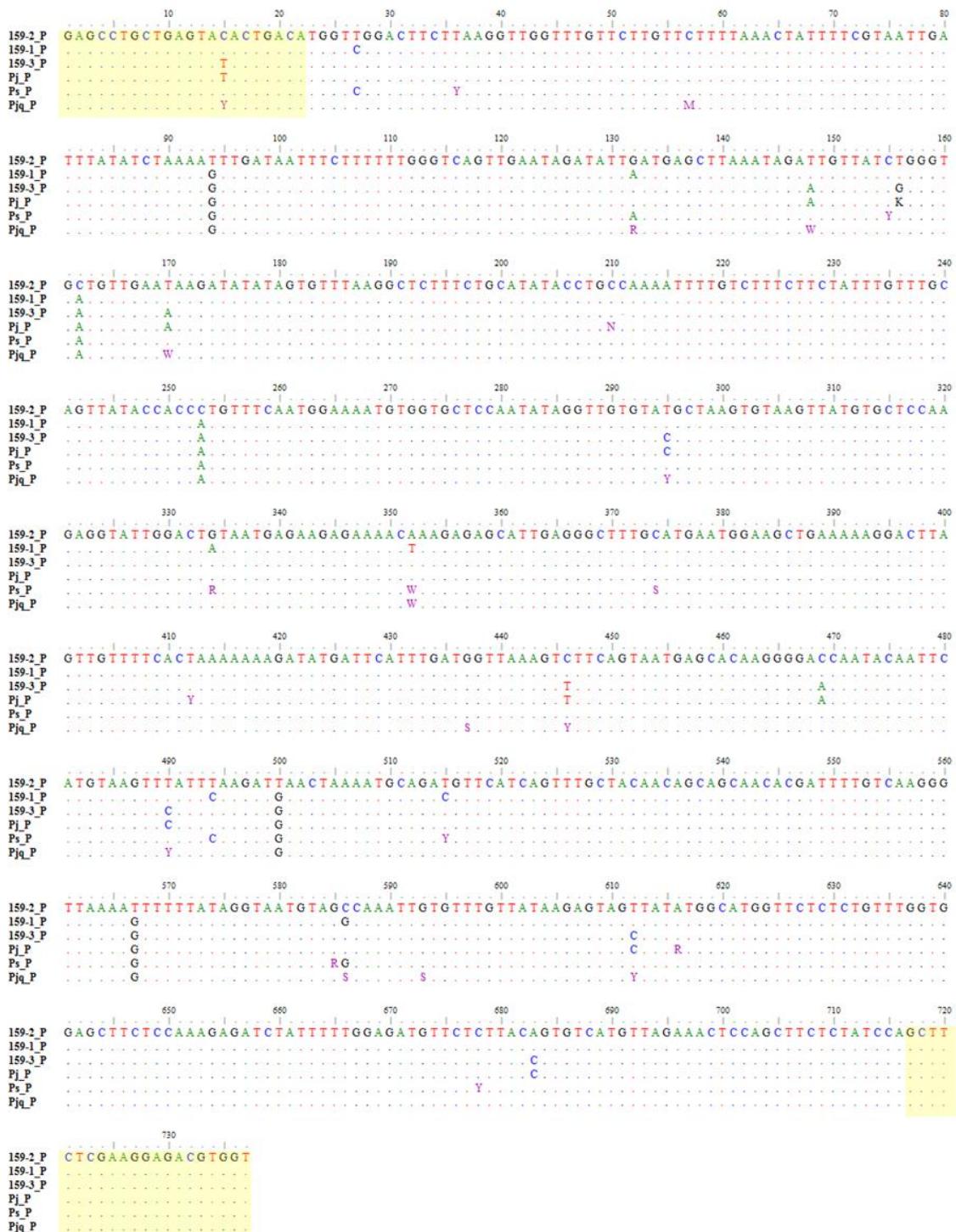
650 660 670 680 690 700

159-2_P CACCAACTAAAGCTATATAGATCAACTTAATAGAGTTAAAGACTTACCTGCTCCCTG
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P

PrC-0402

159-2_M	GAG CCT GCT GAG TAC ACT GAC ATGG T TGG ACT TTCT TAAGG TTGG T TTG TTCT G TCT TTT TAAACT ATTT CGTA ATT GA	10 20 30 40 50 60 70 80
159-1_M	Y	
159-3_M	Y	
Pp_M		
159-2_M	TTT AT AT C T A A A A G T T G A T A A T T C T T T T T G G G T C A G T T G A A T A G A T A T T G A T G A G C T T A A A T A G A T T G T T A T C T G G G T	90 100 110 120 130 140 150 160
159-1_M		
159-3_M		
Pp_M		
159-2_M	GAT G TT G A A T A A G A T A T A G T G T T A A G G C T C T T C T G C A T A T A C C T G C C A A A A T T T G T C T T C T T C T A T T T C C A A G A	170 180 190 200 210 220 230 240
159-1_M		
159-3_M		
Pp_M		
159-2_M	GG T A T T G G A C T G T A A T G A G A G A G A A A C A A A G A G A G T A T T G A G G G C T T T G C A T G A A T G G A A G C T G A A A A A G G A C T T A G T	250 260 270 280 290 300 310 320
159-1_M		
159-3_M		
Pp_M		
159-2_M	T G T T T T C A T T A A A A A G A T A T G A T T C A T T G A T G G T T A A A G T C T T C A G T A A T G A G C A A A A G G G A C T A A T A C A A T T C A T	330 340 350 360 370 380 390 400
159-1_M		
159-3_M		
Pp_M		
159-2_M	G T A A G T T T A T T A A G A T G A A C T A A A A T G C A G A T G T C A T C A G T T G C T A C A A C A G C A C A C G A T T T G T C A A G G G T T	410 420 430 440 450 460 470 480
159-1_M	Y	
159-3_M	Y	
Pp_M		
159-2_M	A A A A G T T T T T A T A G G T A A T G T A G C C A A A T T G T G T T G T T A A G A G T A G T T A T A T G G C A T G G T T C T C T C T G T T T G G T G G A	490 500 510 520 530 540 550 560
159-1_M		
159-3_M	S	
Pp_M		
159-2_M	G C T T C T C C A A A G A G A T C T A A T T T G G A G A T G T T C T T A C A G T G T C A T G T T A G A A A C T C C A G C T T C T C T A T C C A G C T T C T	570 580 590 600 610 620 630 640
159-1_M		
159-3_M		
Pp_M		
159-2_M	CGA A G G A G A C G T G G T	650
159-1_M		
159-3_M		
Pp_M		

PrC-0402



PrC-0508

159-2_M	AGGATGGATTGGTGTCCCTGGTTGTCAGCGTACAATTAACCTTGATACAAGTATGCACAAAAAGAAAATGCAGTTCTT	10 20 30 40 50 60 70 80
159-1_M
159-3_M
Pp_M
159-2_M	TGTTTTGAACGGACCTTAGCATGTTTACACATTCTACTCAGAGACATCCATAGTAGCAGGTAGTAACTAACGGAATT	90 100 110 120 130 140 150 160
159-1_M
159-3_M
Pp_M
159-2_M	G GTGATTGGAGCAAAACAGAAAATAAACCATATGTTAAGGAACAGCAGCCTTAATCTATAAGCATTAACTCCCTGTAGACT	170 180 190 200 210 220 230 240
159-1_M
159-3_M
Pp_M
159-2_M	GAGATAACGATACATCAGTTGATACAGTGGTCCCTTGAGACAATATTATTGCTGTGTTGCAATGAAATACAAGCTTTA	250 260 270 280 290 300 310 320
159-1_M
159-3_M
Pp_M
159-2_M	TGTGTGTTAGATATTAGAAAATATGATAGCTTCCACATATATATTGTTGGTATATATCTGTGTCATGATGTTAAATTCTAGA	330 340 350 360 370 380 390 400
159-1_M
159-3_M
Pp_M
159-2_M	ACAGTGGATTTCATCAATTCGACTGTTTTTTTTTCATCAGGTTGCTTTGGATAAAATATGTAATGACTCTAATAA	410 420 430 440 450 460 470 480
159-1_M
159-3_M
Pp_M
159-2_M	GTAATGCCACCCATTTCACCACTTCAGGCAAAATACATAGCTGAGGCTGAAGAGAAGTATAACAAATTACTGAAAAATG	490 500 510 520 530 540 550 560
159-1_M
159-3_M
Pp_M
159-2_M	CTGGATCCTCAGCTGCTTGAATTACAGAGAAGATGCGTAGAAGCA	570 580 590 600
159-1_M
159-3_M
Pp_M

PrC-0508

159-2_P	AGGATAGATTGGTGTCCCCTGGTTGTCAGCGTACAATTA	CTTACTACAGTACAAAGTATGCAC	TAAGAATGCAAGTTCTTT	10	20	30	40	50	60	70	80			
159-1_P			
159-3_P			
Pj_P			
Ps_P			
Pjq_P	M	.	.	.			
159-2_P	TGTTTTGAAACGGACCTT	AGCATGTTTACACAT	CCAATGTAAGCAGTAG	TGAGTGA	TAGGATTTAAATTAA	T	90	100	110	120	130	140	150	160
159-1_P	
159-3_P	
Pj_P	
Ps_P	R	
159-2_P	CACTTAACAATTGAACATAGAAGAA	TATAAAGAGATCGAGGAGA	AGAAAATATAGATCGAA	AGACAAAGGGCAGAA	AGACAAAGGGCAGAA	W	170	180	190	200	210	220	230	240
159-1_P	.	AG	.	K	
159-3_P	T..C	
Pj_P	
Ps_P	
Pjq_P	
159-2_P	ACAGTAGAACAAAGAAGAA	TATTGGAGGTCAGAGGAGA	AGAAAATATAGATCGA	AGGGGACAAATCTGAGA	GGTTCTAGGAA	A	250	260	270	280	290	300	310	320
159-1_P	.	.	YR	
159-3_P	.	.	Y	
Pj_P	
Ps_P	
Pjq_P	
159-2_P	AAGGTTTACTCAACATT	CAATCTGAATTAAAGAA	CTACATGGGGTATT	TATTA	AGCAAACTAAAC	CTAGACATAATA	330	340	350	360	370	380	390	400
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
159-2_P	GGATAAAAAATCCCCTT	GCTTAAACATAAACTT	CTAACCCAAATAAAAA	TTAGGAAACTTAA	AACTAGAAATCTG	GAATCTGGAA	410	420	430	440	450	460	470	480
159-1_P	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
159-2_P	AAAGGTAGAACATCCTA	CTAGAACATTGGAA	AACTAAAGAATCTA	GGAAACTAACAA	AACTAACAAATA	CATTAGGAAACT	490	500	510	520	530	540	550	560
159-1_P	G.C	W	
159-3_P	G.N	
Pj_P	
Ps_P	
Pjq_P	G.	
159-2_P	AAATAAATTAGGTCTT	ACACGTCCTGCA	CAAGCAGGTAGTA	ACTAACCGGA	ATTTGGTGA	TTGGAGCAACAGAA	610	620	630	640	650	660	670	680
159-1_P	A	
159-3_P	A	
Pj_P	.	.	NN	
Ps_P	
Pjq_P	
159-2_P	CCATATGTTAAGGA	ACAGCAGCCTTAAAT	CTAAAGTTCAACT	CTCTGTAGACTG	AGATAACGATA	ACATCGGTTGATA	690	700	710	720	730	740	750	760
159-1_P	T	
159-3_P	T	
Pj_P	Y	
Ps_P	
Pjq_P	
159-2_P	AGTGGTTCTTGAGACA	ATATTATTGCTGTG	TTGTCATGAA	TACAAAGCTT	ATGTTGTTAGA	TATAAGAAATATGA	770	780	790	800	810	820	830	840
159-1_P	Y	
159-3_P	
Pj_P	
Ps_P	
Pjq_P	
159-2_P	TAGCTTCCACATATA	TATTGTTGGTATTAT	CTATCATGATGTT	AAATTCCTACAA	ACACAGTGAA	TTTCATCAATTG	850	860	870	880	890	900	910	920
159-1_P	G	
159-3_P	G	
Pj_P	G	
Ps_P	GK	
Pjq_P	G	

PrC-0508

PrC-0517

Sequence alignment of 159-2, 159-1, and 159-3 isolates across various samples (M, P, Pj, Ps, Pq) and positions (10-400). The alignment shows high conservation of the sequence across all isolates and samples. A yellow box highlights a region from position 10 to 150, and a green box highlights a region from position 170 to 220.

159-2_M TG CCT ATT GTG ATGGGGTTCCATTGCCAAAGACCCTTT GAGGTGACCCAGAAAAGGTTTTGGTCTCACTATCAAT
159-1_M
159-3_M
Pp_P

159-2_M CCTCTAGTGTGCAAAACAATTAGAAGGGCAAGAGCTAAGACTCTGGGATTTAGTGAAGGAATGAGGAGTAACATTCCGA
159-1_M
159-3_M
Pp_P

159-2_M GATGGATTATGTGAGACAGGAGCTAGAATTTCGCTCGAAGGATCTTGCAACAGAACTCCTGCTGCCAGTAATTGGTGAGA
159-1_M
159-3_M
Pp_P

159-2_M ATATTGCACCTTCACATCTGTGCTCATATCAGTTGGTTTTAGTATGATCTTGTGGCTGTTTCGTTAAGCTTCAT
159-1_M
159-3_M
Pp_P

159-2_M TTGGTTGAGTTAACATTGTCGGCTGAGTTAACATCCTGCTCGTATTGGTTGATTGAGAGGTAACAGGCAAAGCAA
159-1_M
159-3_M
Pp_P

159-2_M TTGAAGAAACTGCTGCTTGTACTGAGGCTTTATCATGACCGGAAGCACAAGTGCTCA
159-1_M
159-3_M
Pp_P

159-2_P TG CCT ATT GTG ATGGGGTTCCATTGCCAAACACCCTTT GAGGTGACCCAGAAAAGGTTTTGGTCTCACTATCAAT
159-1_P
159-3_P
Pj_P
Ps_P
Pq_P

159-2_P CCTCTAGTGTGCAAAACAATTAGAAGGGCAAGAGCTAAGACTCTGGGTTTTAGTGAAGGAATGAGGAGTAACATTCCGA
159-1_P
159-3_P
Pj_P
Ps_P
Pq_P

159-2_P GATGGATTATGTGAGACAGGAGCTGGATTTCGCTCGAAGGATACTTGCACAGAACTCCTATCTGCCAGTAATTGGTGAGA
159-1_P
159-3_P
Pj_P
Ps_P
Pq_P

159-2_P ATATTGCACCTTCACATCTGTGCTCATATCATAATTGGTTGATTGAGAGGTAACAGGCAAAGCAATTGAAGAAACTGCTGCTGTT
159-1_P
159-3_P
Pj_P
Ps_P
Pq_P

159-2_P TTGGTTGAGTTAACATTCCCTGCTCATATTGGTTGATTGAGAGGTAACAGGCAAAGCAATTGAAGAAACTGCTGCTGTT
159-1_P
159-3_P
Pj_P
Ps_P
Pq_P

159-2_P G TACTGAGGCTTTATCATGACCGGAAGCACAAGTGCTCA

PrC-0563

159-2_M
159-1_M
159-3_M
Pp_M

159-2_M
159-1_M
159-3_M
Pp_M

159-2_M
159-1_M
159-3_M
Pp_M

159-2_P
159-1_P
159-3_P
Pj_P
Ps_P
PjQ_P

159-2_P
159-1_P
159-3_P
Pj_P
Ps_P
PjQ_P

159-2_P
159-1_P
159-3_P
Pj_P
Ps_P
PjQ_P

PrC-0705



PrC-0890

PrC-1418

159-2_M	T G A T A T C T C C T T G G C T C C A T C C A G T T T G G C C A A G C A A A C A C A C G G G T A T T T T A A T C T T A T T T C C A A C T A T T G A A A T T T	10	20	30	40	50	60	70	80		
159-1_M		
159-3_M		
Pp_M		
.....											
159-2_M	G A C T G A A A T G C A T A A T T T A T T T C A T T A A C A T C A A C T T C C A T C A C A T T G A A C A T A G A A A G T A A A G C C A T A T A A C C A	90	100	110	120	130	140	150	160		
159-1_M		
159-3_M		
Pp_M		
.....											
159-2_M	C T T T A A A T G T A A T G G T A G A A G A A G A A A A A C A A A G T C C C T T A G A A T G C A G C A C A T A T A T T G A T A T A G T T G T	170	180	190	200	210	220	230	240		
159-1_M	S		
159-3_M	S		
Pp_M	W		
.....											
159-2_M	T A T T C A G G G T A T A A T T G G A T C T G A A G G A A T C A G A A T T T G A A A C A T G T A G A A G A A G G A A T C T T C A G T T T A A C A T C T A A C T	250	260	270	280	290	300	310	320		
159-1_M		
159-3_M		
Pp_M	R		
.....											
159-2_M	A T C C T G A T T G G A A C C T G T T T T C T G A A A C T T T A T T T G G A A A T C A T A C C T T C T T A A T T T A C A T T T C T T G A A A A G T T T	330	340	350	360	370	380	390	400		
159-1_M	Y		
159-3_M	K		
Pp_M	G		
.....											
159-2_M	C T A C T C T A T A T C A C C T G G T C T A A T G C T A C T T C A C C C A G G A A T C T T C G G T T T A A C A T C T A A C C A T C C T G A T T G G A A C C T G	410	420	430	440	450	460	470	480		
159-1_M	Y		
159-3_M	H		
Pp_M	D		
.....											
159-2_M	T T T T C T G G A A A T T T G A T T T T G G A A A A T C A T A C C T T C T T A A T T T A C A T T T C T T G A A A A G T T T C T G C T C T A T A T C A C C T G	490	500	510	520	530	540	550	560		
159-1_M	Y		
159-3_M	Y		
Pp_M	R		
.....											
159-2_M	G T C T A A T G C T A C T T C A T C C C A G G C T T T G A T A C C T T G C T T C G G A A G T T A C A A A G G T A T G T C C C T C A A G G A T C A T C T G T T G	570	580	590	600	610	620	630	640		
159-1_M	Y		
159-3_M	Y		
Pp_M	R		
.....											
159-2_M	C T G A C T T G T A T G C A G G A G C	590									
159-1_M										
159-3_M										
Pp_M										

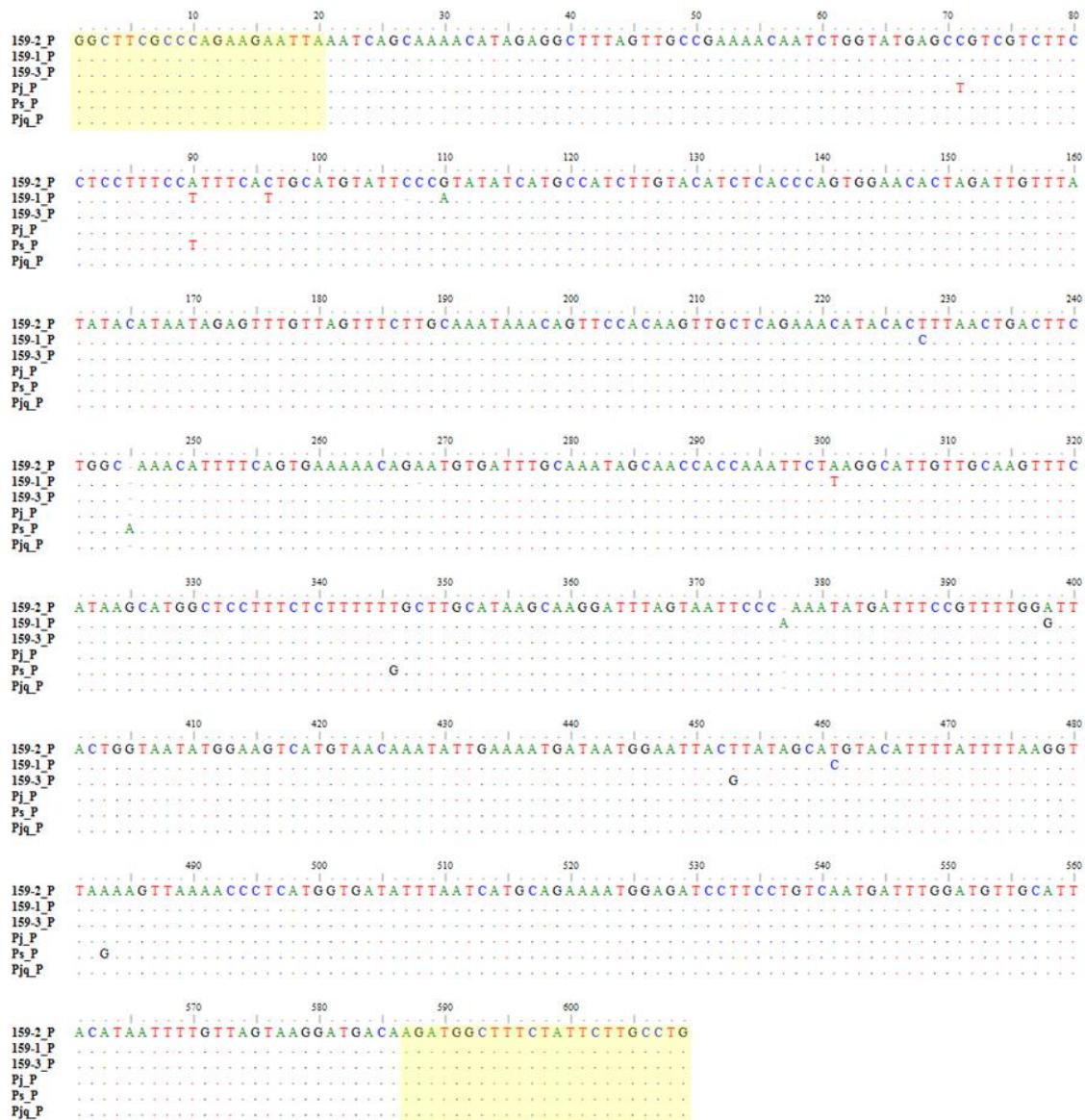
PrC-1418

159-2_P	TGATATCTCCCTGGCTCCATCCAGTTTGCCAAAGCAACACACGGGTATTTAATCTTAT	TTTCTAACT	ACTGA
159-1_P	CCAAC	GA. T TG
159-3_P	NNAANCN ANN	SWA HTNTG
Pj_P	CCAAC	GA. T TG
Ps_P	S	NNAANNN ANNNNNRA	WINTG
Pjq_P	CCAAC	GA. T TG
10 20 30 40 50 60 70 80			
159-2_P	AATGCATAATTATTTCTTAACATCAACCTTCCATCACACTGTGAGCGTAGAAATTAAAGCCATATAAACCACTGTAA	W. M. T. A. A	G
159-1_P	T
159-3_P	A. M. T. A. A	G. M.
Pj_P	A. A. T. A. A	G
Ps_P	A. A. T. A. A	G
Pjq_P	A. M. T. A. A	G
90 100 110 120 130 140 150 160			
159-2_P	ATGTAAGGGTAGAAGAAGAAAAAGAAAAACAAAAA TCCCT	TAGAATGCA	CATATATTATGATATAGTTGT
159-1_P	T	S. G	GCA
159-3_P	T	S. G	NNNGCA
Pj_P	T	G	GCA
Ps_P	T	G	GCA
Pjq_P	T	NNNNNGTCCCT	GCA
170 180 190 200 210 220 230 240			
159-2_P	TATTCAAGGGTATAATTGGATCCTGAAGGAATCAGAAATTGAAACATGTAAGGAAATCTTGGTTAACATCTAAC	R. Y
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
250 260 270 280 290 300 310 320			
159-2_P	ATCCTGACTGGAACCTGTTTCTGAAACTTGATTATGGAAAATCATA	CCTTCTTAATTTCATTTACATTCTTGTAAAAAG
159-1_P	Y	N. T	NT
159-3_P	Y	N. T	NT
Pj_P	Y	Y	NN. TA
Ps_P	N. T	Y. NT
Pjq_P	N. T	NT	NN. TA
330 340 350 360 370 380 390 400			
159-2_P	TTCTACTCTATATCACCTGGTCTAATGCTACTTCATCCCAGGCTTTGATACCTTGTCTCCGAAGTTACAAAGGTATGT	R. Y
159-1_P
159-3_P
Pj_P
Ps_P	NNNNNNNNNNNN	R. Y
Pjq_P	NNNNNNNNNNNN	Y
410 420 430 440 450 460 470 480			
159-2_P	CCCTCAAGGATCATCTGTGGCTGACTTGTATGCAGGAGC
159-1_P
159-3_P
Pj_P
Ps_P
Pjq_P
490 500 510			

PrC-1681

159-2_M	GGCTTCGCCAGAAGAATTAAATCAGCAAAACATAGAGGCTTTAGTTGCTGAAACAACTGGTATGAGCCGTCGTCTTC	10 20 30 40 50 60 70 80
159-1_M	
159-3_M	
Pp_M	
159-2_M	CTCCTTTCTTTCACTGCATGTTACATGCCATCTGTACATCTCACCCAGTGGACAACTAGACTGTTA	90 100 110 120 130 140 150 160
159-1_M	
159-3_M	
Pp_M	
159-2_M	TATAACATAATAGAGTTTGTAGTTCTGCAAAATAACAGTTCCACAAGTTGCTCAGAAACATAACTTTAACTGTGAGA	170 180 190 200 210 220 230 240
159-1_M	
159-3_M	
Pp_M	
159-2_M	TCCCCACATCGACCAACGGAGGGGGGTGATGTCCTTATATGTCACACCCCATCCATCTAGCACGAGGCCCTTGGG	250 260 270 280 290 300 310 320
159-1_M	
159-3_M	
Pp_M	
159-2_M	AGCTCACTGGCTTCGGAGTCGTAGGAACCTCGAAAGTTAACCGAGCTGGGGCTAGAGCAATCCAGGATGGGTGACCCAC	330 340 350 360 370 380 390 400
159-1_M	
159-3_M	
Pp_M	
159-2_M	TGGGAAGTTGGTCGTGAGCTCCAGAAACAAACCCATAAGGGCCAGAGAGGGGGGCCAAAGCGGACAATATCGTGCTA	410 420 430 440 450 460 470 480
159-1_M	
159-3_M	
Pp_M	
159-2_M	CGGCGGAGCGATCCGGGATGTGACATAACTGACTCTGGCAAAACATTTCAGTGGAAAAACAGAAATGTGATTGCAAAAT	490 500 510 520 530 540 550 560
159-1_M	
159-3_M	
Pp_M	
159-2_M	AGCAACCCACAAATTCTAAGGCATTGTTGCAAGTTTCAAGCATGGCTCTTCTCTTTGCTTGCATAAGCAAGGA	570 580 590 600 610 620 630 640
159-1_M	
159-3_M	
Pp_M	
159-2_M	TTTAGTAATTCCAAATATGATTTCGTTGGATTACTGGGAATATGGAACTCATGTAACAAATATTGAAAAATGATAAT	650 660 670 680 690 700 710 720
159-1_M	
159-3_M	
Pp_M	
159-2_M	GGAATTACTTATAGCATGTACATTTCAGTTAAGGTAAAGTTAAACCCCTCATGGTGATTTAATCATGCAAGAAAT	730 740 750 760 770 780 790 800
159-1_M	
159-3_M	
Pp_M	
159-2_M	GGAGATCCTTCTGTCAATGATTGGATGTTGCAATTACATAATTGTTAGTAAGGATGACAAGATGGCTTCTATTCTT	810 820 830 840 850 860 870 880
159-1_M	
159-3_M	
Pp_M	
159-2_M	GCCTG	
159-1_M	
159-3_M	
Pp_M	

PrC-1681



Phylogenetic and Population Structure Analysis of Wild *Prunus yedoensis* (Rosaceae) in Jeju Island, Korea

Cho Ara

Department of Biological Sciences
Graduate School, Myongji University
Directed by Professor Mun Jeong-hwan

Prunus yedoensis Mat. is one of the most popular ornamental trees in northeastern Asia and its native wild populations have been found in Jeju island, Korea. Previous studies suggested that wild *P. yedoensis* is a hybrid species. However, no solid evidences have been made on the parental origin and phylogenetic relationship with its closely related *Prunus* species. This study aims to estimate the genetic structure of wild *P. yedoensis* in Jeju island based on molecular markers developed from *Prunus* Conserved Ortholog Set (COS) and to deduce the putative parental species of *P. yedoensis*. Using the draft genome sequence of *P. yedoensis* as well as two reported genomes of *P. mume* and *P. persica*, a total of 1,303 COS genes were identified by reciprocal BLASTP searches. Among them, 38 dimorphic insertion-deletion Polymerase Chain Reaction (InDel PCR) markers were designed. For plant materials, a total of 73 wild *P. yedoensis* accessions were collected, of which 21 trees are designated as the Natural Monuments by the government or Jeju province and 52 trees are wild plants distributed in the forest of Mt. Halla. In addition, 45 samples included in *P. pendula* for. *ascendens* (19), *P. jamasakura* (16), *P. sargentii* (7) and *P. jamasakura* var. *quelpaertensis* (3) were collected. Genotyping and population structure analysis of wild *P. yedoensis* using 20 PrCOS markers showed that approximately 90% (70 of 78 accessions) of wild *P. yedoensis* are heterozygote

with 50% compartment of maternal genotype of *P. pendula* for. *ascendens* and 50% compartment of paternal genotype of *P. jamasakura*, *P. sargentii* or *P. jamasakura* var. *quelpaertensis*. The remaining 8 accessions showed 75% compartment of maternal or paternal genotype, suggesting a backcrossing with its parents. Moreover, phylogenetic analysis presented that wild *P. yedoensis* and *P. pendula* for. *ascendens* were grouped into one cluster and they were separated from other *Prunus* species, whereas *P. jamasakura*, *P. sargentii* and *P. jamasakura* var. *quelpaertensis* were clustered together. In order to identify the parental origin of wild *P. yedoensis*, sequence comparison of target regions for PrCOS markers was performed. The maternal allelic sequence of *P. yedoensis* was most similar to that of *P. pendula* for. *ascendens* at sequence identity 0.99. On the other hand, the paternal allelic sequence of *P. yedoensis* showed the highest sequence identity is *P. jamasakura* (0.99) followed by *P. jamasakura* var. *quelpaertensis* (0.98) and *P. sargentii* (0.93). These findings suggest that wild *P. yedoensis* in Jeju island, Korea is a homoploid hybrid species originated from a cross between maternal *P. pendula* for. *ascendens* and paternal *P. jamasakura*. Furthermore, additional introgression of parent genotypes by backcrossing is ongoing process.

Keyword

Prunus yedoensis, COS, InDel, Molecular marker, Genetic structure, Phylogenetic analysis, Genome sequence identity, Homoploid hybrid