

디지털육종을 적용한 수박 흰가루병 저항성 판별



정명희¹, 이옥진², 임유진¹, 이혜은², 신윤희^{1*}

¹경기 용인시 기흥구 흥덕1로 13 흥덕IT밸리 (주)인실리코젠 R&D 센터
²전북특별자치도 완주군 이서면 농생명로 100 국립원예특작과학원 채소기반과

Abstract

박과에 속하는 수박(*Citrullus lanatus*)은 전 세계적으로 널리 소비되는 주요 작물로, 식이섬유뿐만 아니라 시트룰린(citrulline)과 라이코펜(lycopene)과 같은 기능성 물질을 풍부하게 함유하고 있다. 그러나 최근 기후변화로 인한 재배 환경 악화와 병해충 발생 증가로 수박의 생산성이 저하되고 있으며, 병해충에 의한 수확량 손실은 20~40%에 이를 것으로 추정된다. 특히 흰가루병(powdery mildew)은 박과 작물 전반에서 발생하는 대표적인 곰팡이성 병해로, 잎 표면에 흰색 가루층을 형성하여 광합성을 방해하고 생육을 크게 저하한다. 이에 따라 본 연구에서는 수박의 유전형 및 표현형 데이터를 기반으로 디지털육종 기술을 적용하여 흰가루병 저항성 품종 개발의 효율성을 높이고자 하였다. 수박 308개 자원을 대상으로 개체별 병반면적률을 조사하고, ApeKI 효소를 이용한 GBS (Genotyping by Sequencing)를 통해 유전형 데이터를 확보하였다. 확보된 고품질 SNP 22,167개를 바탕으로 GWAS (Genome Wide Association Study)를 수행하여 흰가루병 저항성과 유의하게 연관된 SNP를 도출하였으며, SNP의 중요도(feature importance)를 기준으로 선별한 최종 26 SNP를 활용하여 기계학습 기반 예측 모델을 구축하였다. 해당 모델은 흰가루병 저항성 여부를 판별하는 데 활용할 수 있으며, 예측 정확도는 87.5%로 나타났다. 이러한 결과는 병 저항성 개체의 선발 및 교배 조합 작성 시 정확하고 신속한 의사결정을 가능하게 하는 데이터 기반 육종 실용성을 입증한 것으로, 본 연구 결과는 향후 수박뿐만 아니라 다양한 원예작물의 주요 형질 예측 및 정밀 육종에도 폭넓게 활용될 수 있을 것으로 기대한다.

Materials & Methods

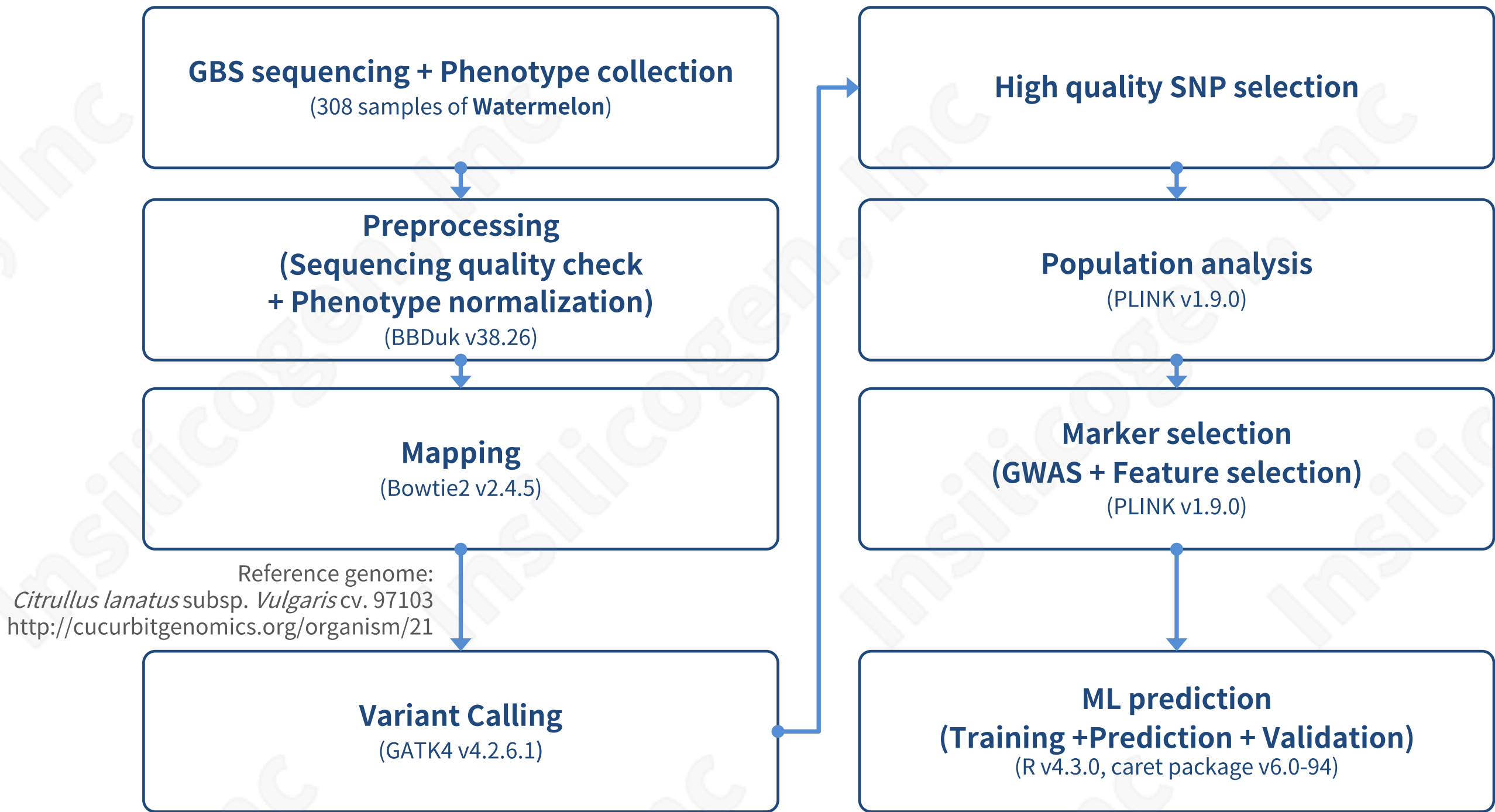


그림 1. 수박 308 계통의 디지털육종 모식도

Results

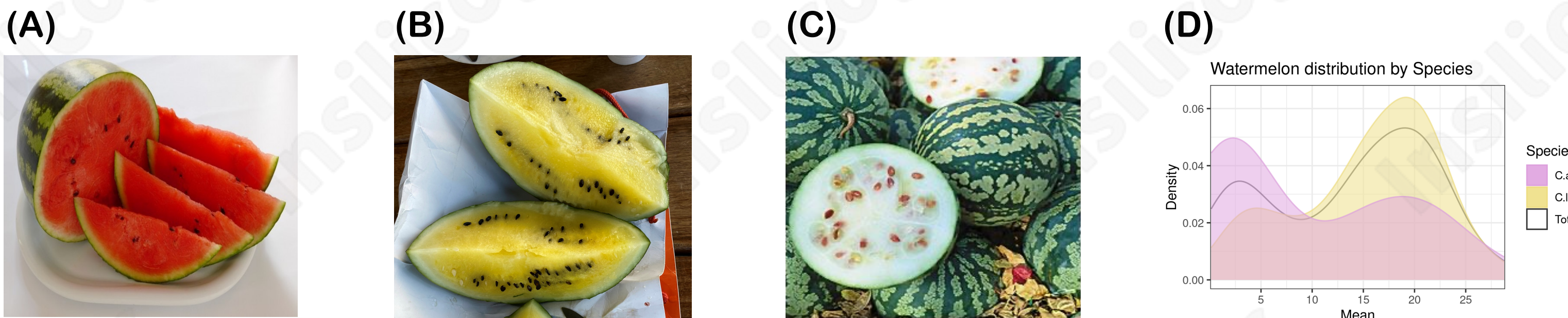


그림 2. 수박(*Citrullus* spp.) 병반면적률 분포

(A) *C. lanatus* (수박; watermelon); (B) *C. mucospermus* (에구시 멜론; Egusi melon); (C) *C. amarus* (유자멜론; Citron melon); (D) 분석에 사용된 수박 품종별 병반면적률 분포

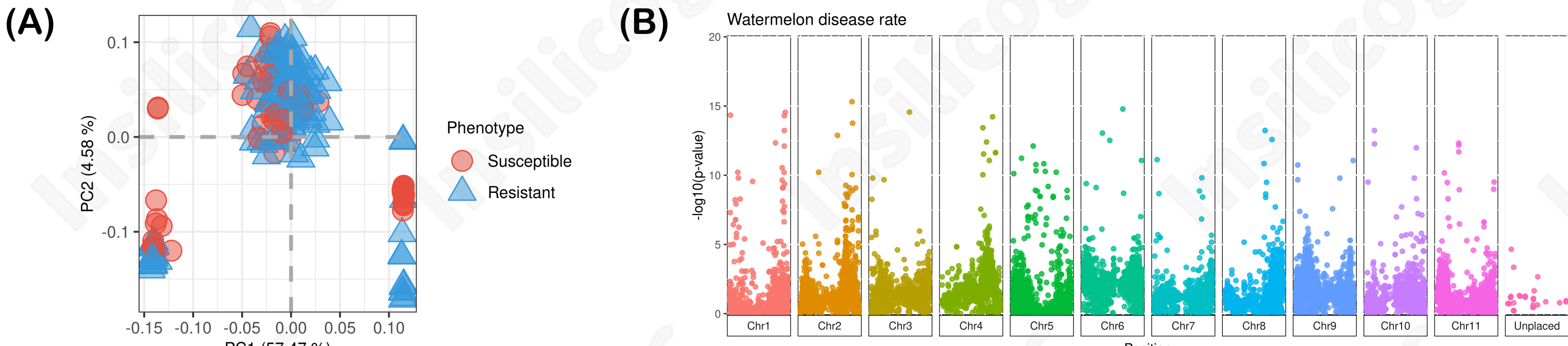


그림 3. PCA(Principal component analysis) 및 Manhattan plot
(A) 고품질 변이를 활용한 GBS 데이터 별 PCA; (B) 고품질 변이를 활용한 Manhattan plot

표 1. 흰가루병 저항성 기계학습 예측 모델 구축 결과

26SNPs로 구축된 흰가루병 저항성 기계학습 예측 모델					
ML Model	Accuracy	Specificity	Sensitivity	#of train set	#of validation set
SVM	0.89	0.90	0.88	128	56
KNN	0.84	0.80	0.89	128	56
RF	0.89	0.83	1.00	128	56
C5	0.89	0.89	0.78	128	56
PLS	0.84	0.84	0.78	128	56
GBM	0.96	0.96	0.67	128	56

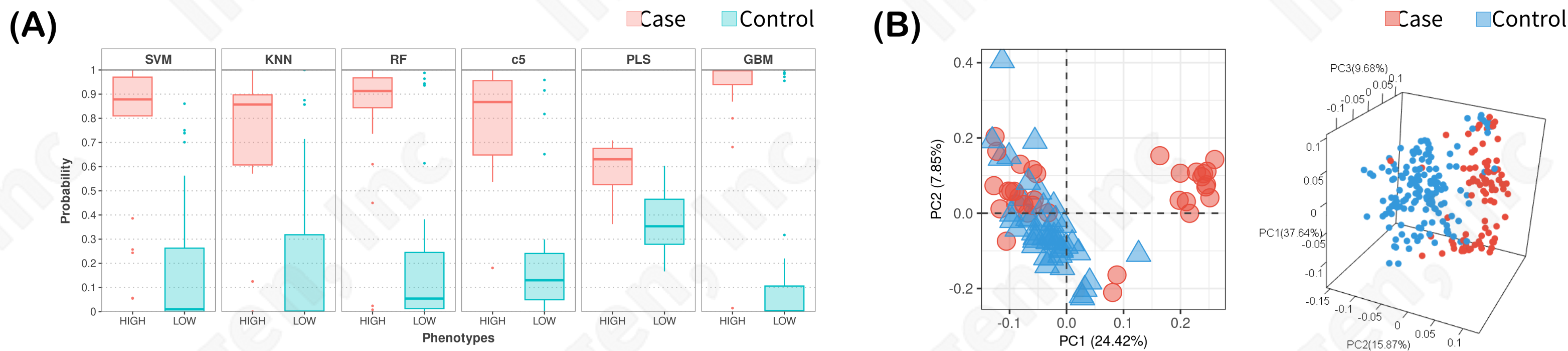


그림 4. 선별된 SNP 마커를 이용한 기계학습 모델의 prediction box plot 및 PCA

(A) 26개 SNP 마커를 이용한 기계학습 모델의 예측 확률 분포(왼쪽부터 SVM, KNN, RF, C5, PLS, GBM; 빨강: 감수성 30%(Case); 초록: 저항성 30%(Control)); (B) 모델 구축에 사용되지 않은 개체들을 이용한 PCA(원), 전체 개체를 이용한 PCA(오)

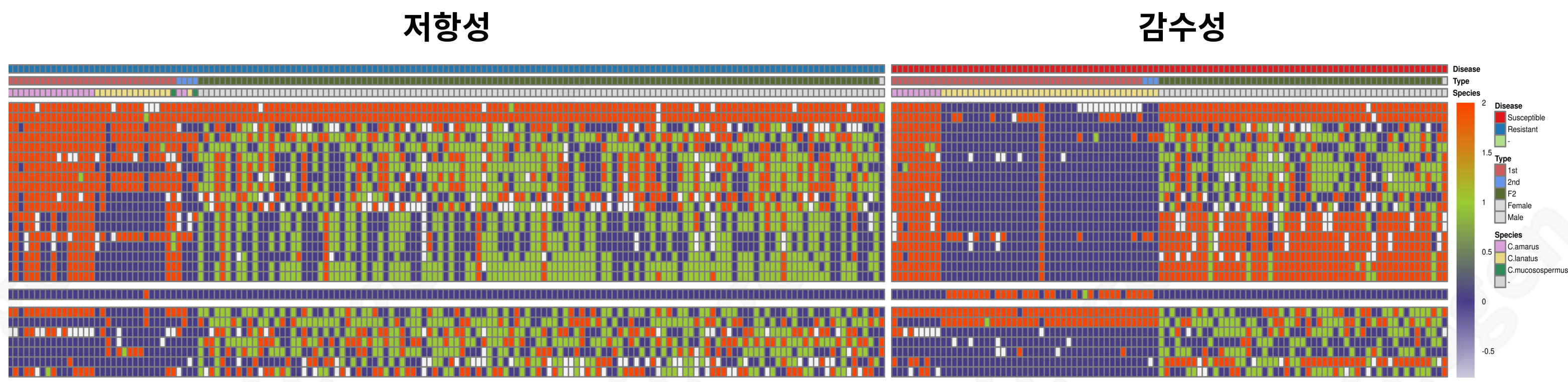


그림 5. 모델 구축에 사용되지 않은 개체들의 26개 SNP에 대한 유전형



그림 6. 디지털육종 플랫폼 내 데이터 탑재 및 형질 예측

Conclusion

본 연구는 수박 계통에 대한 유전형 정보를 바탕으로 흰가루병 저항성과 연관된 SNP 마커를 제시하며, 흰가루병 저항성 개체를 정확하게 식별하고 교배 조합 작성시 정확하고 신속한 의사결정을 가능하게 하는 데이터 기반 육종 실용성을 입증하였다. 즉, 총 26개 SNP 마커를 사용하면 흰가루병 저항성 개체 및 교배조합을 선별 할 수 있다는 것을 의미하며, 향후 수박 뿐만 아니라 다양한 원예작물의 주요 형질 예측 및 정밀 육종에도 폭넓게 활용될 수 있을 것으로 기대한다.

Acknowledgement

본 결과물은 농림축산식품부의 재원으로 농림식품기술기획평가원의 디지털육종전환기술개발사업의 지원을 받아 연구되었음(RS-2022-IP322074).