

디지털육종을 적용한 무의 복합 원예 형질 예측



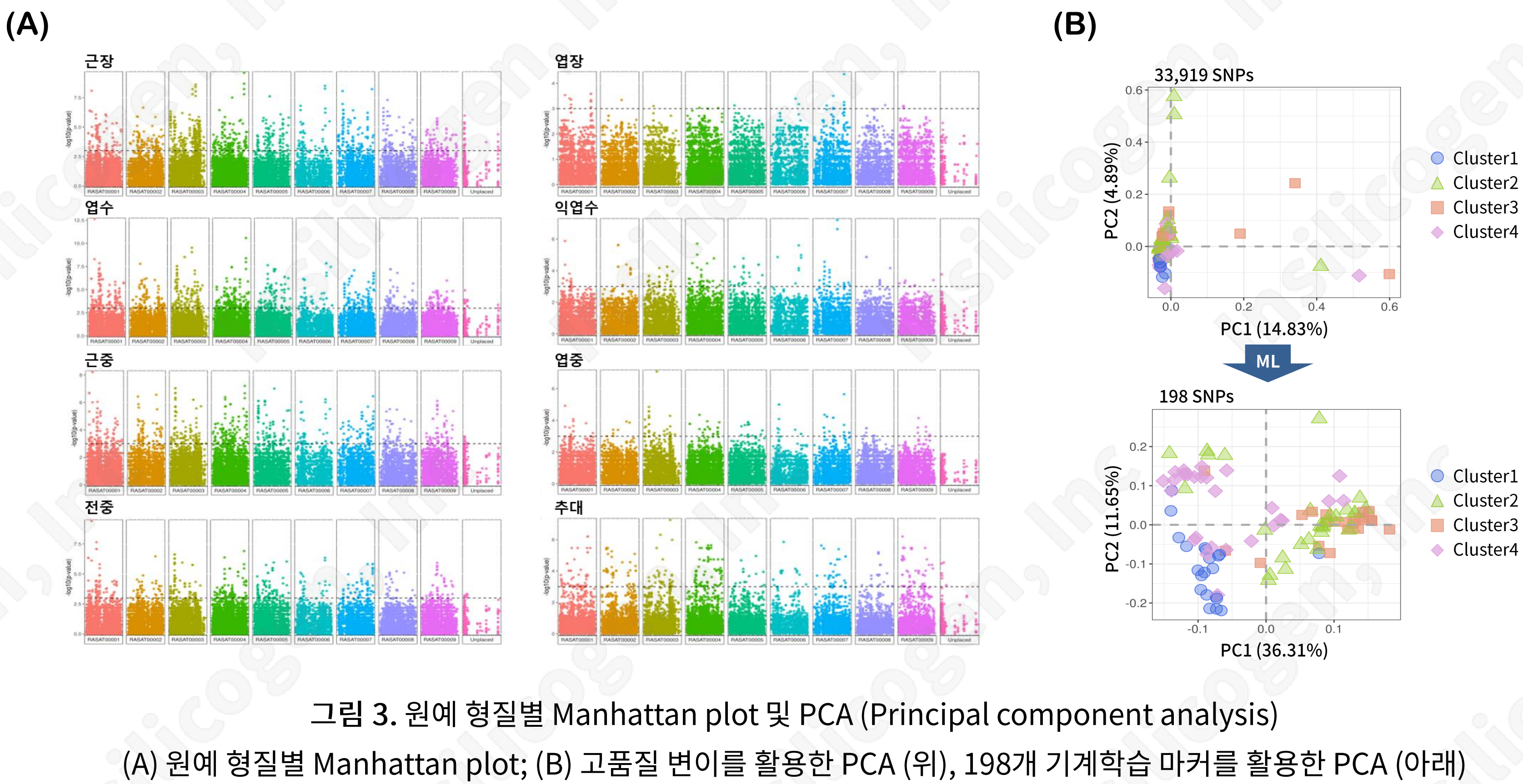
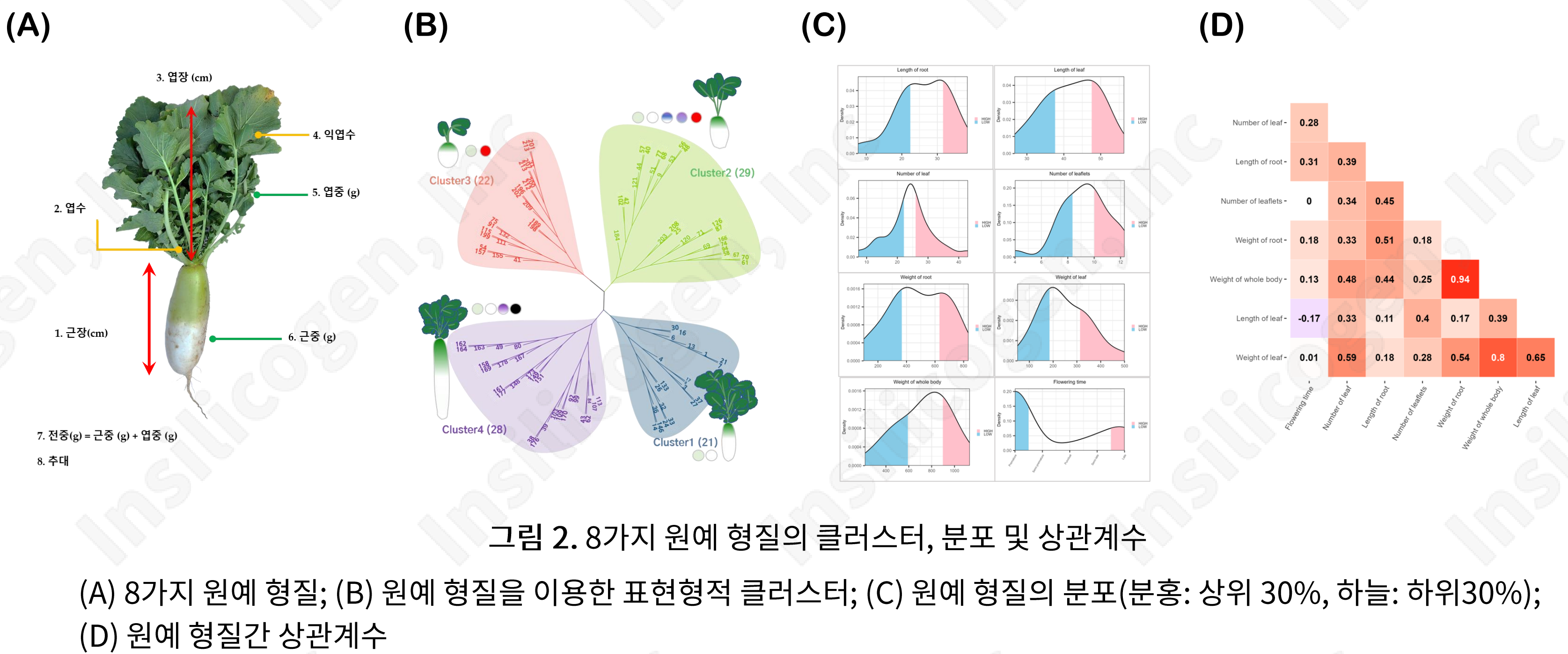
임유진¹, 박병준^{2,3}, 정명희¹, 조성현¹, Sathiyamoorthy Subramaniyam¹, 최수빈², 허성호^{2,3}, 손효정¹, 박한용², 신윤희^{1*}

¹경기 용인시 기흥구 흥덕1로 13 흥덕IT밸리 (주)인실리코젠 R&D 센터 ²서울특별시 광진구 능동로 209 세종대학교 바이오자원과학학과
³경기 평택시 이화로 251-29 농업회사법인 다산바이오㈜

Abstract

무는 배추과의 뿌리 채소로 김치, 깍두기, 무말랭이, 쌈무, 단무지 등으로 많이 소비되고 있는 원예 작물로 특히, 김장 문화로 인해 국내 종자 시장에서 큰 비중을 차지하고 있습니다. 이러한 무는 봄 무, 여름 무, 가을 무, 월동 무, 알타리 및 열무 등 쓰임새에 따라 여러 가지 계통으로 나뉘지므로 목적에 맞는 우수한 품종 개발이 요구됩니다. 또한 무는 뿌리, 줄기, 잎이 모두 소비되므로 하나의 원예 형질이 아닌 다수의 원예 형질이 동시에 육종되어야 하며, 자가 불화합성, 기후 변화, 병 저항성 등 여러 요인이 복합적으로 작용하고, 고토리당 종자 수도 적어 육종이 어려운 작물입니다. 이러한 난제 해결을 위해 우리는 100개의 무 엘리트 계통을 대상으로 8가지 원예 형질(근장, 엽장, 엽중, 엽수, 익엽수, 근중, 전중, 추대)을 고려한 디지털육종을 수행하였습니다. 원예 형질 정보를 수집하고, 다변량 통계 기법을 사용하여 클러스터링을 수행한 결과, 4개의 클러스터로 구분되었고, 이들은 계통간 표현형적 차이를 명확히 구분하였습니다. 또한, 박달무 유전체(*Raphanus sativus* L. Bakdal)를 참조서열로 한 33,919개의 고품질 변이 정보를 산출하고, 이를 바탕으로 형질별 연관 마커를 선별하여 유전적 다양성을 모두 반영한 기계학습 예측 모델을 구축하였습니다. 이들 8개의 원예 형질 예측 정확도는 평균 89.82%로 하나의 원예 형질이 아닌 다수의 양적 원예 형질을 통합적으로 고려하여 개체 선별과 교배조합 작성에 대한 의사결정에 도움을 줄 것입니다. 본 연구는 무의 우수 육종 계통에 대한 유전형 정보를 바탕으로 원예 형질과 연관된 SNP 마커를 제시하고, 정량적 원예 형질을 정확하게 식별하는 데 있어 연관 마커와 더불어 기계학습 예측 모델의 효율성을 입증함으로써, 빅데이터 기반의 고품질 마커 개발 및 디지털육종의 기반을 마련하는 데 기여할 것으로 기대합니다.

Results



Conclusion

본 연구는 무의 우수 육종 계통에 대한 유전형 정보를 바탕으로 원예 형질과 연관된 SNP 마커를 제시하며, 정량적 원예 형질을 정확하게 식별하는 데 있어 분자 마커와 더불어 기계학습 예측 모델의 효율성을 입증하였습니다. 즉, 형질별 21 ~ 40개의 총 198개 SNP 마커를 사용하면 엽중, 추대, 근장 등 8가지 모든 원예 형질의 우수 또는 불량 여부를 판별할 수 있다는 것을 의미하며, 이는 빅데이터 기반의 고품질 분자 마커 개발 및 디지털육종의 기반을 마련하는데 기여할 것입니다.

Acknowledgement

본 결과물은 농림축산식품부의 재원으로 농림식품기술기획평가원의 디지털육종전환기술개발사업의 지원을 받아 연구되었습니다(RS-2022-IP322069).

Materials & Methods

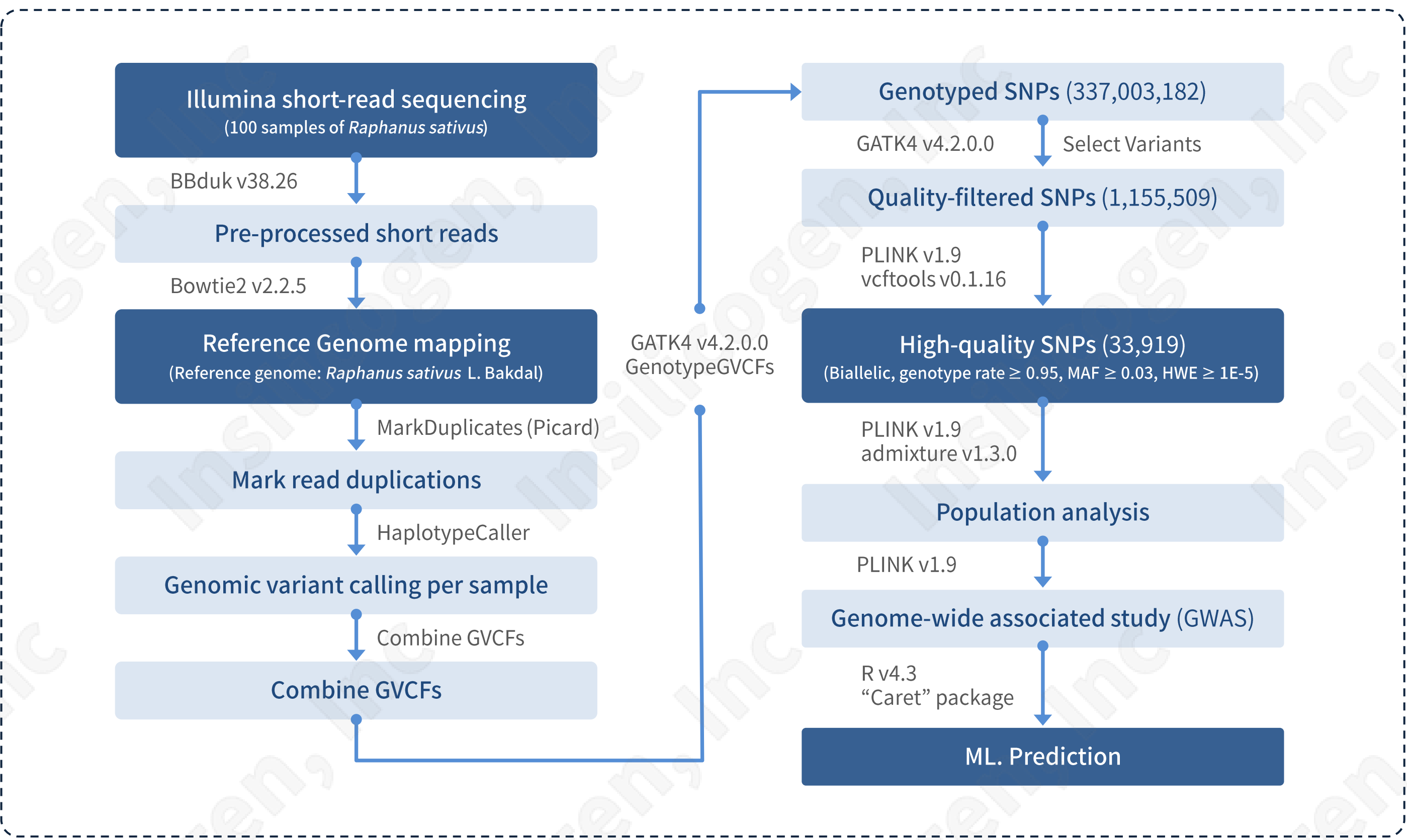


그림 1. 박달무 유전체를 이용한 무 엘리트 100계통의 디지털육종 모식도

표 1. 원예 형질별 기계학습 예측 모델 정확도

원예 형질	개별 원예 형질별 연관 마커로 구축된 모델						8개 다중 형질 통합 마커로 구축된 모델 (198 SNPs)			
	# of SNP	ML model	Acc.	Sp	Sn		ML model	Acc.	Sp	Sn
근장	34	KNN	0.89	0.90	0.88		C5	0.89	0.90	0.88
엽장	31	SVM	0.84	0.80	0.89		C5	0.79	0.60	1.00
엽수	33	C5	0.89	0.83	1.00		KNN	0.95	0.92	1.00
익엽수	25	SVM	0.89	0.89	0.78		RF	0.95	1.00	0.89
근중	36	RF	0.84	0.84	0.78		SVM	0.84	0.90	0.78
엽중	21	SVM	0.90	0.90	0.88		C5	0.90	1.00	0.75
전중	40	RF	0.83	0.83	0.89		SVM	0.94	0.89	1.00
추대	23	KNN	0.96	0.96	0.67		SVM	0.92	0.67	0.95

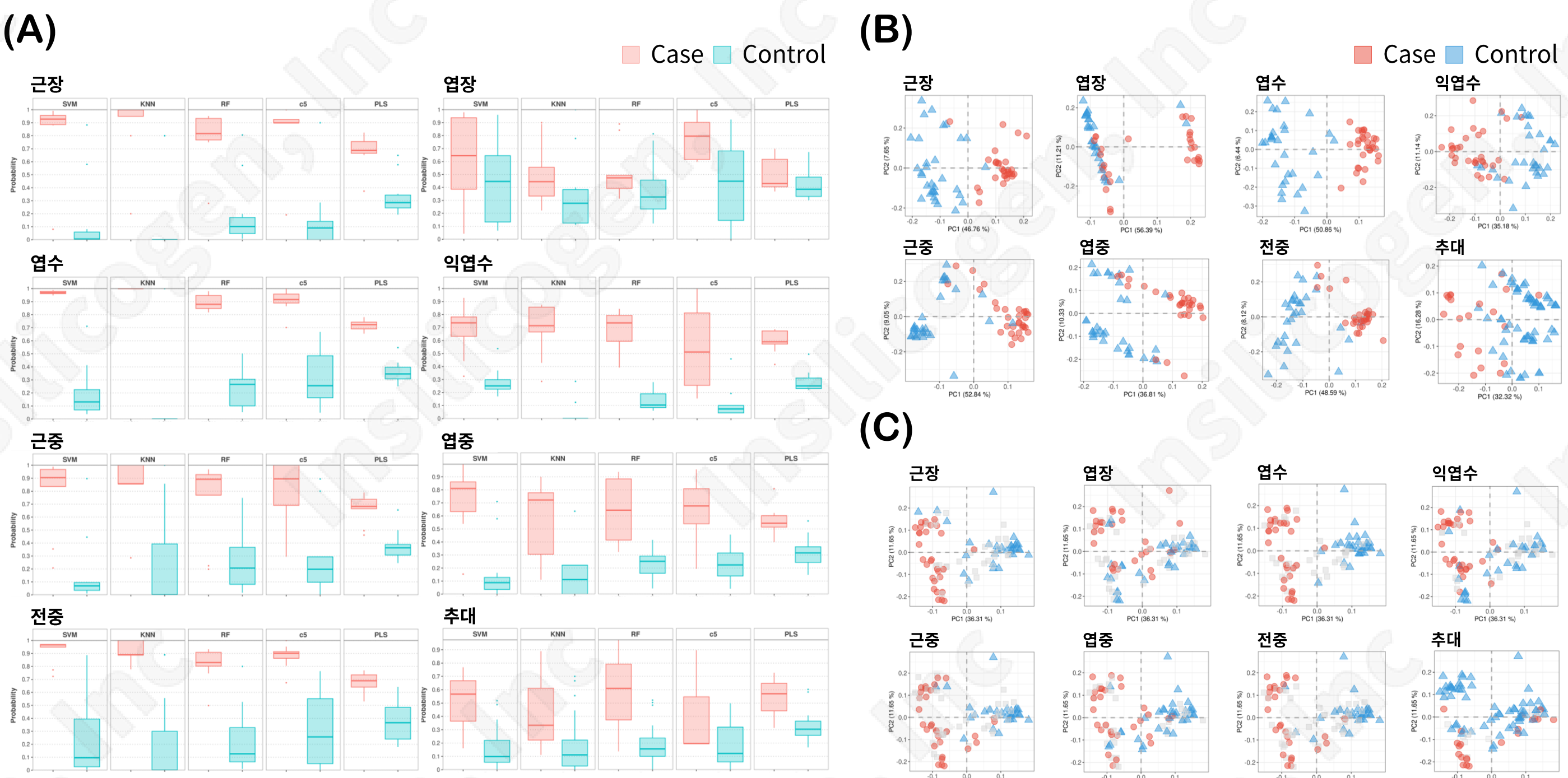


표 2. 원예 형질별 기계학습 예측 모델 정확도

계통명	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	...	S98
형질 예측 모델														
근장	0.99	0.80	0.11	0.67	0.03	1.00	0.01	0.09	0.77	0.79	0.00	1.00	...	0.91
엽장	0.65	0.29	0.04	1.00	0.04	0.03	0.47	0.63	0.15	0.51	0.23	0.98	...	0.65
엽수	0.98	0.98	0.64	1.00	0.03	1.00	0.08	0.26	1.00	0.87	0.01	1.00	...	1.00
익엽수	0.65	0.09	0.34	0.98	0.11	0.13	0.00	0.15	0.48	0.59	0.07	0.92	...	0.90
근중	0.98	0.98	0.07	0.95	0.00	0.99	0.31	0.77	0.94	0.45	0.12	1.00	...	0.89
엽중	0.99	0.99	0.09	0.99	0.01	1.00	0.11	0.58	0.05	0.90	0.18	0.98	...	0.79
전중	0.94	0.94	0.39	0.99	0.10	0.99	0.47	0.70	0.88	0.61	0.27	0.95	...	0.89
추대	0.97	0.97	0.02	0.00	0.04	1.00	0.09	0.79	0.06	0.14	0.01	0.86	...	0.42