한국형 밀 핵심집단의 유전적 다양성과 집단 구조 분석

민경도¹ · 강유나² · 김창수² · 최창현³ · 김재윤^{1*}

¹국립공주대학교 산업과학대학 식물자원학과, ²국립충남대학교 농업과학대학 식물자원학과, ³국립식량과학원 밀연구팀

Genetic Diversity and Population Structure of Korean Common Wheat (*Triticum Aestivum*)

Kyeong Do Min¹, Yu Na Kang², Chang Soo Kim², Chang Hyun Choi³, and Jae Yoon Kim^{1*}

¹Department of Plant Resources, College of Industrial Science, Kongju National University, Yesan 32439, Republic of Korea

²Department of Plant Computational Genomics, Chungnam National University, Daejeon 34134, Republic of Korea

³National Institute of Crop Science, RDA, Wanju 55365, Republic of Korea

Abstract Wheat (*Triticum aestivum*) is one of the three major food crops, along with rice and corn, and is the second most consumed crop after rice in Korea. However, the domestic production of wheat is insufficient, and the self-sufficiency rate is recorded in single digits. As wheat has a large genome size of 17 Gbp, and contains many repeated nucleotide sequences, it is difficult to conduct breeding studies and genome-based breeding lags behind that of other crops. To overcome the above challenges, we constructed a wheat core collection using simple sequence repeat markers that are suitable for the domestic cultivation environment with excellent reproducibility. Genetic diversity and population structure were analyzed using a core collection. Agricultural traits were evaluated in the Korean wheat core collection. Single marker analysis was correlated with 21 agricultural traits to identify potential molecular markers. These results may be useful for wheat breeding programs in the precision breeding era.

Keywords core collection, genetic diversity, population structure, SSR marker, wheat

Received on July 23, 2021. Revised on July 23, 2021. Accepted on August 4, 2021.

* Corresponding Author (E-mail: jaeyoonkim@kongju.ac.kr, Tel: +82-41-330-1203, Fax: +82-41-330-1209)

서론

밀은 미네랄, 비타민 등 유용성분이 많이 함유되어 있는 식량작물이며, 쌀, 옥수수과 더불어 3대 주요 작물 중 하나로 많은 노동력을 필요로 하지 않고 척박하고 건조한 환경에서도 생육이 원활하여 세계 인구의 약 30%가 주식으로 활용하고 있다(Kang et al. 2016, Oh et al. 2019). 또한, 밀은 쌀, 보리, 옥수수와 달리 종실 자체를 식용으로 활용하는 것 보다는 2차 가공(제분)과정을 거쳐 밀가루로 사용하는 것이 일반적이며, 이는 밀에 함유된 단백질 분자들이 물을 흡수하여 점탄성을 갖게 되어 빵, 국수, 과자 등 여러 식품의 원료로 쉽게 활용할 수 있기 때문이다(Oh et al. 2019). 국내에서 1인당 연간 밀 소비량은 33 Kg으로 약 200만 톤 가량이 소비되고 있으나 자급률은 약 12%로 낮은 수준이다(Kim et al. 2012, Oh et al. 2019). 밀 자급률 향상을 위하여 다양한 밀 육종에 대한 연구 및 새로운 육종신소재 개발을 진행하고 있으나 밀의 복잡한 염색체 구조로 인하여 연구 성과가 부진하

며 그로 인하여 유전체 분석 기술 및 육종집단의 보환을 통하여 육종 연구를 진행하고자 한다(Kim et al. 2020, Choi et al. 2018).

보통밀은(*Triticum aestivum* L., 2n=6×=42, AABBDD) 유전체 크기가 17 Gbp로 매우 방대한 이질 6배체의 거대 배수성 작물이며 반복된 DNA서열로 구성된 WGD (Whole Genome Duplication)을 전체 유전체의 85% 정도 포함하고 있기 때문에 타작물에 비하여 유전체 데이터 수집과 전장유전체 염기서열 분석에 제약이 많다(Choi et al. 2018, Paux et al. 2006). 또한 우리나라는 지리적으로 재배할 경작지가 협소하고 밀의 종 다양성이 풍부하지 않기 때문에 밀 자원관리 및 육종 소재가 되는 유전자원의 관리가 필요하며 연구자가 실제 활용하는 유전자원 집단의 연구 목표에 부합성이 높은 정확한 육종집단의 구축이 필요하다(Ma et al. 2018, Brown & Spillane 1999).

최근 국내외 유전체 기반 육종 연구로 핵심집단 구축 및 NAM (Nested Association Mapping) 집단을 구축하려는 연구가 진행되고 있다(Choi et al. 2018). 콩 유전자원의 다양성에 관한 연구로

형태적 특성, 동위효소의 변이연구를 수행하여 왔으며 후속연구로 한국 콩 재래종의 핵심집단에 관한연구(Cho et al. 2008)를 통하여 260계통을 선정하였으며, 야생콩에서도 493점을 핵심집단으로 확보하였다(Park et al. 2011). 또한, IA3023품종을 모본으로 하여 40개의 부분과 교배하여 총 5,600개의 RIL (Recombinant Inbred Lines)계통으로 이루어진 NAM집단을 양성하였다(Song et al. 2017). 토마토의 경우에는 국립유전자원센터 등 국내외에서 수집한 약 400점 이상의 유전자원 중 형태적 특성과 유전적 다양성 및 유연관계를 분석하여 95점을 선발하였고 국립원예특작과학원이 보유한 100점, 민간 육종회사 3점을 포함하여 총 198점으로 핵심집단을 구축하였다(Sim et al. 2018). 고추는 대학 보유 유전자원과 국립농업유전자원센터에서 보유한 유전자원을 합한 3,821개의 자원에 대하여 ‘UnWeighted Neighbor Joining’으로 계통도를 작성한 후 Powercore1.0 소프트웨어를 활용하여 252 점의 핵심자원을 선발하였다(Kwon et al. 2015). 또한, 해외에서는 밀의 유전자형과 표현형 정보를 수집하여 가공 적성에 맞게 활용하고 있으며 스위스, 캐나다, 인도 등의 연구 현황으로는 2000년부터 2007년 사이에 가장 많은 것으로 보고 되었으나 국내에서는 밀 유전자원으로 핵심집단을 구축한 사례가 없었다(Son et al. 2013).

분자 유전학 기술의 발전으로 다양한 분자 표지 인자가 개발되어 있으며 특히 SSR (Simple Sequence Repeat) 마커는 염기서열의 반복이 DNA 복제 과정에서 반복구간의 절단 및 삽입에 의해 변화가 일어나 그 차이를 활용하여 분석하며, 식물체 게놈 전체에 널리 분포하기 때문에 재현성 높은 마커로 인정받고 있다(Hong et al. 2013, Kim et al. 2014). 또한 SSR마커는 공우성 마커로서 반복염기서열 확인 및 프라이머 제작 등의 초기 개발비용과 시간이 소요되는 어려움이 존재하나 loci마다 많은 alleles를 가지고 있으며 Polymorphic Information Content (PIC)가 높으며 식물 집단들 내에 유전자좌마다 많은 대립인자들을 가지면서도 대립유전자의 특성을 명확하게 나타낼 수 있어 유전적 다양성, 계통유연관계 및 집단구조를 분석하기에 수월하다고 알려져 있다(Hu et al. 2009, Bang et al. 2011, Kim et al. 2015).

벼는 약 30,000점의 유전자원을 보유하고 있는 농촌진흥청 국립농업과학원 국립유전자원센터에서 기초정보를 활용하여 4,400 점에 대하여 분자유전학적 DNA information이 구축되어 있으며 이중 166점을 핵심집단으로 선발하고 SSR마커를 활용하여 집단의 유전적 구조분석을 수행한 바 있다(Park & Jeong 2009). 또한, 참깨의 경우 2016년에 유전자원센터에 제공받은 전체 핵심집단 277점에서 SSR마커를 활용하여 112점의 미니 핵심집

단을 구축 한 바 있다(Min et al. 2016) 해외에서는 팽나무버섯(*Flammulina velutipes*)에 대하여 SSR 마커를 활용하여 유전적 다양성을 확인하고 19개의 품종과 13개의 야생계통을 핵심집단으로 선발한 사례가 있다(Liu et al. 2018).

최근에는 NGS (Next Generation Sequencing) 기술을 이용하여 핵심집단 구축 및 검정에 활용되고 있으나 밀은 A, B, D계놈의 매우 높은 반복염기서열과 거대한 유전체 크기로 인하여 전장유전체 염기서열 분석을 통해 얻은 데이터를 구분하기 어려워 SSR 마커를 활용하였다(Moon et al. 2015, Safar et al. 2010).

해당 연구는 이러한 유전적 다양성 및 집단구조를 분석하기 수월한 SSR마커를 활용하여 농촌진흥청에서 보유중인 밀 유전자원과 국내 고세대 계통을 포함한 유전자원 1,967점에 대하여 유전체 분석과 농업 특성 평가를 통하여 국내 환경 적응 밀 핵심집단을 구축하고자 한다.

재료 및 방법

밀 유전자원 수집 및 DNA 추출

본 연구를 진행하기 위하여 유전자원센터에서 분양 받은 수집 유전자원 1,267점과 진흥청 고세대 계통인 CB700점을 포함하여 1,967점(Table 1)을 확보하였고 농촌진흥청 식량과학원 포장에서 생육되었으며, 성숙기 이전의 신선한 잎을 채취하여 액화질소로 곱게 마쇄한 후 Genomic DNA Prep Kit (Biofact, Daejeon, Korea)를 활용하여 Genomic DNA를 수집하였으며 농도는 20 ng/ul로 조절하여 -20℃에 보관하였다.

농업특성평가

1,967점의 유전자원에 대하여 공주대학교 연구포장 및 국립식량과학원 연구 포장에서 증식을 진행하였으며 국립식량과학원 연구포장(위도 35°, 경도 127°)에 입모, 엽이, 초형과 작물의 생육에 밀접한 관계를 가진 출수기, 성숙기, 도복 등의 21가지 농업특성에 대해 2018년, 2019년에 걸친 현장평가를 2회 실시하였으며 축적된 데이터를 통하여 형질별로 수치화하여 정리하였다(RDA 2012). 출수의 경우 파종 후 각 계통 처리구의 이삭이 40%이상 출수된 날짜, 성숙의 경우 계통의 60%이상의 이삭이 익었을 때의 날짜를 기록하였으며 간장은 지표면에서부터 이삭 목까지를 기준으로 길이를 측정하였다.

SSR마커를 이용한 PCR

해당 연구에서는 Balfouier et al. (2007)을 참조하여 핵심집단

Table 1. Genetic resources of origin country in this study.

No.	Country	Resources	No.	Country	Resources	No.	Country	Resources
1	Afghanistan	72	21	France	16	41	Pakistan	11
2	Argentina	12	22	Germany	14	42	Peru	3
3	Armenia	4	23	Greece	31	43	Poland	14
4	Australia	11	24	Hungary	18	44	Portugal	43
5	Austria	7	25	India	48	45	PRT	12
6	BGR	7	26	Iran	3	46	Romania	6
7	Bosnia & Herzegovina	20	27	Iraq	8	47	Russia	49
8	BRA	2	28	ISR	2	48	South Africa	6
9	Bulgaria	30	29	Italy	17	49	Spain	16
10	Canada	18	30	JPN	24	50	SVN	1
11	CHN	159	31	KGZ	4	51	SWE	1
12	CHL	4	32	Kenya	5	52	Syria	13
13	COL	6	33	South Korea	342	53	THK	3
14	Croatia	19	34	Lebanon	6	54	TUN	3
15	CUB	1	35	Macedonia	15	55	Turkey	44
16	Ecuador	2	36	Mexico	367	56	Ukraine	17
17	Egypt	8	37	Mongolia	18	57	USA	155
18	ETG	6	38	Morocco	4	58	Uzbekistan	5
19	Ethiopia	76	39	Nepal	2	59	Yemen	1
20	FIN	2	40	NLD	5	60	YUG	1

을 선발하기 위한 37쌍의 SSR primer (Table 2)를 확보 1,967점에 대하여 PCR을 수행하였다. PCR 증폭에서 용액 조성은 총 25 uL로 20 ng의 genomic DNA, 1× PCR buffer, 0.1 mM의 forward와 reverse primers, 0.2 mM dNTPs, 1 unit의 Taq Polymerase (Biofact, Seoul)로 구성하였다. PCR 반응은 처음 95℃에서 2분간 최초 denaturation을 한 후, 다음 95℃에서 30초간 denaturation, 각 프라이머의 annealing temperature의 조건에 맞게 50~60℃에서 40초간 annealing, 72℃에서 2분간 extension을 하여 35 cycle을 반복한 후, final extension을 72℃에서 5분간 실시하였다. 증폭된 PCR product의 경우 QIAxcel Advanced Instrument (Qiagen, Hilden, Germany)를 활용하여 디지털 전기영동을 실시하였다.

Bioinformatics tool을 활용한 근연관계 정립 및 핵심집단 구분

QIAxcel ScreenGel Software (Qiagen, Hilden, Germany)를 이용하여 증폭된 PCR screenig을 진행하였으며, band의 유무에 따라 마커별로 0 또는 1로 표기하여 기록하였으며 이를 토대로 MEGA-X를 활용하여 Maximum Likelihood method를 Tamura-

Neimodel로 근연관계 조사를 수행하였다(Nei 1973). 사용된 수식은 아래와 같다.

$$\text{Gene Diversity (GD)} = 1 - \sum P_{ij}^2$$

P_{ij} 는 i번째 마커에서 j번째 대립유전자의 빈도수이다.

또한, Locus당 대립단편 수를 계산하여 CorehunterII (Beukelaer et al. 2012), Genocore (Jeong et al. 2017) 등의 Bioinformatics tool을 활용하여 핵심집단 후보군을 선발하였다(Kim et al. 2007). 사용된 수식은 아래와 같다.

$$\text{Coverage (\%)} = 1/m \sum_{j=1}^m \frac{Dc}{De} * 100$$

또한, structure 2.3.4와 structure harvester를 활용하여 PCR결과를 바탕으로 하여 Length of Burn in period를 10,000으로 설정 MCMC repeats를 5000으로 설정하여 분석을 수행하여 delta K와 최적의 그룹 및 clustering을 계산하였다(Falush et al. 2003, Perrier & Jacquemoud-Collet 2006, Earl et al. 2012).

Table 2. List of used SSR marker primers in this study. Each primers were designed as described in Balfourier et al. (2007).

Primer	Forward (5'→3')	Reverse (5'→3')
Xgwm2-3A	CTGCAAGCCTGTGATCAACT	CATTCTCAAATGATGGAACA
Xgwm5-3A	GCCAGCTACCTCGATACAACTC	AGAAAGGGCCAGGCTAGTAGT
Xgwm11-1B	GGATAGTCAGACAATTCTTGTG	GTGAATTGTGTCTTGTATGCTTCC
Xgwm44-7D	GTTGAGCTTTTCAGTTCGGC	AGTGGCATCCACTGAGCTG
Xgwm46-7B	GCACGTGAATGGATTGGAC	TGACCCAATAGTGGTGGTCA
Xgwm99-1A	AAGATGGACGTATGCATCACA	GCCATATTTGATGACGCATA
Xgwm120-2B	GATCCACCTTCCTCTCTCTC	GATTATACTGGTGCCGAAAC
Xgwm135-1A	TGTCAACATCGTTTGTAAAAGG	ACACTGTCAACCTGGCAATG
Xgwm149-4B	CATTGTTTTCTGCCTCTAGCC	CTAGCATCGAACCTGAACAAG
Xgwm196-6A	ACCACTGCAGAGAACACATACG	GTGCTCTGCTCTAAGTGTGGG
Xgwm186-5A	GCAGAGCCTGGTTCAAAAAG	CGCCTCTAGCGAGAGCTATG
Xgwm190-5D	GTGCTTGCTGAGCTATGAGTC	GTGCCACGTGGTACCTTTG
Xgwm219-6B	GATGAGCGACACCTAGCCTC	GGGGTCCGAGTCCACAAC
Xgwm233-7A	TCAAAACATAAAATGTTTCATTGGA	TCAACCGTGTGTAATTTTGTCC
Xgwm234-5B	GAGTCCTGATGTGAAGCTGTTG	CTCATTGGGGTGTGTACGTG
Xgwm251-4B	CAACTGGTTGCTACACAAGCA	GGGATGTCTGTTCCATCTTAG
Xgwm257-2B	AGAGTGCATGGTGGGACG	CCAAGACGATGCTGAAGTCA
Xgwm260-7A	GCCCCCTTGACACAATC	CGCAGCTACAGGAGGCC
Xgwm261-2D	CTCCCTGTACGCCTAACGC	CTCGCGCTACTAGCATTG
Xgwm272-5D	TGCTCTTTGGCGAATATATGG	GTTCAAAACAAATTAAAAAGGCC
Xgwm285-3B	ATGACCCTTCTGCCAAACAC	ATCGACCGCGATCTAGCC
Xgwm312-2A	ATCGCATGATGCACGTAGAG	ACATGCATGCCTACCTAATGG
Xgwm337-1D	CCTCTTCCTCCCTCACTTAGC	TGCTAACTGGCCTTTGCC
Xgwm372-2A	AATAGAGCCCTGGGACTGGG	GAAGGACGACATTCCACCTG
Xgwm400-7A	GTGCTGCCACCACTTGC	TGTAGGCACTGCTTGGGAG
Xgwm413-1B	TGCTTGCTAGATTGCTTGGG	GATCGTCTCGTCTTGGCA
Xgwm415-5A	GATCTCCCATGTCCGCC	CGACAGTCGTCACTTGCCTA
Xgwm427-6A	AAACTTAGAACTGTAAATTCAGA	AGTGTGTTCAATTTGACAGTT
Xgwm437-7D	GATCAAGACTTTTGTATCTCTC	GATGTCCAACAGTTAGCTTA
Xgwm469-6D	CAACTCAGTGACACACAACG	CGATAACCACTCATCCACACC
Xgwm480-3A	TGCTGCTACTTGTACAGAGGAC	CCGAATTGTCCGCCATAG
Xgwm539-2D	CTGCTCTAAGATTCATGCAACC	GAGGCTTGTGCCCTCTGTAG
Xgwm566-3B	TCTGTCTACCCATGGGATTTG	CTGGCTTCGAGGTAAGCAAC
Xgwm610-4A	CTGCCTTCTCCATGGTTTGT	AATGGCCAAAGGTTATGAAGG
Xgwm626-6B	GATCTAAAATGTTATTTTCTCTC	TGACTATCAGCTAAACGTGT
Xgwm642-1D	ACGGCGAGAAGGTGCTC	CATGAAAGGCAAGTTCGTCA
Xgwm664-3D	CAGTCAGTGCCGTTAGCAA	AGCTTTGCTCTATTGGCGAG

Marker trait analysis를 이용한 single marker analysis

Phenotype과 genotype data를 회귀분석한 marker-trait analysis를 진행하여 각 마커의 Loci는 통계적 모델(Lynch & Walsh 1998, Soller et al. 1976)에 의해 분석을 수행하였다. 분석모델은 아래와 같으며

$$\log(y_i+1) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + e_i$$

방정식에서 y_i+1 값은 각각의 농업특성 형질을 의미하며, β_0 은 방정식의 절편, β_1 은 추가효과 x_{1i} 는 설명변수, β_2 는 우월효과, e_i 는 확률 오차이다. 이러한, Single-marker analysis를 통하여

21가지 형질에 대하여 LOD값 5를 기준으로 하여 형질별 연관마커 후보군을 선발하였다.

결과 및 고찰

농업특성평가 및 밀 유전자원 생육비교

국립식량과학원에서 제공 받은 밀 유전자원에 대하여 원산지 조사를 실시한 결과 전체 1,967자원 중 멕시코 367점과 국내 기원 342점이 포함되어 있으며 멕시코의 경우 국제 옥수수밀연구소(CIMMYT)이 있어 다양한 밀 유전자원을 수집할 수 있을 것이며 국내유전자원의 경우 한국의 재배환경에 적합한 계통을 탐색하고자 국내에서 육성중인 CB계통 및 재래종을 포함하였을 것으로 사료된다(Table 1). 중국 159점과 미국 155점으로 중국 다음으로 높은 자원 수를 차지하였으며 수집한 유전자원들에 대하여 2018년, 2019년에 걸쳐서 증식 및 농업특성평가를 수행하여 국내 환경에 적합한 계통들을 고려하였다.

국내에서 육성된 품종들은 조숙이면서도 대립, 다수성이 육종의 목표였기 때문에 1990년대에 육성한 금강밀과 우리밀의 경우에도 이러한 특징을 가지고 있으며 전체 유전자원의 출수기와 성숙기 및 수발아 저항성 관련된 적립계 탐색을 위하여 종피색을 조사하였다(Table 3). 농촌진흥청 농업 과학 기술 조사 분석 기준을 참고하여 각 유전자원에 대하여 총경수의 40%가 출수한 날을 출수기로 기록하고 계통의 이삭 대부분이 황화된 시기를 성숙기로 지정하여 구분하였으며 파종일부터 출수기까지의 영

양생장기간이 출수기부터 성숙기까지의 생식생장기간보다 길다. 또한, 국내 밀의 평균 출수기는 2018년도 2019년도 각각 평균적으로 5월 3일, 5월 2일로 비슷하게 확인 되었으며 성숙기는 6월 11일로 같았다(RDA 2012, Son et al. 2021). 금강밀의 경우 출수기가 4월 말, 성숙기가 6월 초로 조숙성을 띄고 있으며 수량성도 다른 품종들에 비해 뛰어나므로 현재까지도 재배 장려 품종으로 평가 받고 있으며 국내 밀 생산의 약 70%를 차지하고 있다(Kim et al. 2020, Cho et al. 2018). 이렇게 국내에서 품종화된 계통의 경우 국내 재배환경에 맞춰 출수기와 개화시기가 빠른 것이 특징인데 수집한 전체 유전자원의 약 3분의 1인 659계통이 출수기가 4월 말 이전이므로 국내에서 생육하기에 적합한 장점을 가지고 있다고 사료된다. 또한, 약 80%의 유전자원의 종피색이 붉은색을 띄고 있는데 국내에서는 적립계 수발아 저항성 과자용 밀인 ‘고소’를 육성하였으며 국수용으로는 금강밀, 빵용으로는 조정밀을 재배하고 있다(Kang et al. 2015).

SSR 마커를 활용한 유전자원의 genetic diversity 확인

A계능 15개, B계능 12개, D계능 10개로 구성된 총 37세트의 SSR마커를 제작하여 밀 유전자원 1,967계통에 대해서 PCR을 수행을 통하여 genetic diversity를 분석하였다(Balfouier et al. 2007). 디지털 전기영동을 실시하여 마커당 평균 allele 수를 정밀 검정한 결과 평균 1.58, 최소값은 0.54 최댓값은 2.29로 나타났다(Table 4, Fig. 1). 유럽 중심의 유전자원으로 빵밀 핵심집단을 구축한 연구에서는 프랑스 계통의 유전자원 비율이 매우 높았고 모집단 3,942자원에 대해 평균 23.9 allele를 분석한 것으로 보고 되었는데(Balfouier et al. 2007) 반면 이 연구에서 평균 allele가 낮게 검정된 이유는 유럽 중심의 밀 핵심집단이 국내 자원이 다수 포함된 수집 자원에 비하여 해외 자원 중심으로 개발된 SSR의 적합도가 높기 때문으로 사료된다. 그러나 각각의 SSR마커는 A, B, D 계능에 고르게 분포되어 있으며 이를 활용하여 독립적인 loci 분석이 가능하고 Single marker analysis에 적용할 수 있으며 농업형질별로 구분한다면 좀 더 세밀한 육종이 가능할 것으로 예상된다.

Bioinformatics tool을 활용한 핵심집단 선발

수집 유전자원의 SSR 마커 분석 결과를 분석하여 Omics 프로그램으로 핵심집단을 구축하였다. CorehunterII 분석으로 393점, Genocore 분석으로 293점을 핵심집단을 각각 선발하였으며, 중복된 선발 자원을 제거한 후 총 614점을 선발하였다(Beukelaer et al. 2012, Jeong et al. 2017). 또한, 선발된 614점의

Table 3. Evaluation of ratio for 1,967 genetic resources to 3 trait of 19 agricultural trait.

Character	Character state	Frequency	Percentage (%)
Seed coat color	White	341	17.33
	Yellow	56	2.84
	Red	1555	79.05
	Purple	5	0.25
	Black	1	0.05
	White&Red	3	0.15
	Yellow&Red	6	0.03
Heading date	4/15~4/29	659	33.5
	4/30~5/12	1130	57.44
	5/13~5/26	178	9.04
Ripening date	5/28~6/5	204	10.38
	6/6~6/14	1468	74.63
	6/15~6/25	295	14.99

Table 4. Part of genotype scoring total 1,967 genetic resources.

Primer	No.												
	1	2	3	4	5	6	7	8	9	10	11	12	13
High Value	6	5	5	9	9	4	7	4	6	6	4	5	7
Average	1.25	2.26	1.46	2.29	1.49	1.46	2.70	1.15	1.64	2.01	1.11	2.10	1.99

Primer	No.												
	14	15	16	17	18	19	20	21	22	23	24	25	26
High Value	4	7	6	7	13	5	3	7	6	5	7	4	5
Average	0.86	2.44	1.99	2.15	2.27	1.21	0.76	1.96	1.60	0.54	1.93	0.92	1.16

Primer	No.											
	27	28	29	30	31	32	33	34	35	36	37	
High Value	6	4	4	6	4	3	3	3	6	8	6	
Average	1.42	1.18	1.18	1.83	1.12	1.52	1.34	1.24	1.57	1.83	1.56	

Fig. 1. Part of PCR band (1-95 accession × 6 SSR markers) using QIAxcel (digital electrophoresis). The top line and bottom line are the alignment markers provided by the company to correct each PCR products (alignment marker).

계통들에 대하여 근연관계를 검정하기 위하여 ML (maximum Likelihood) tree를 MEGA-X 패키지로 분석하여 선발된 각 핵심 집단 간의 phylogenetic tree를 작성하였다(Fig. 2). 품종간 교배를 통한 계통들도 존재하며 고세대 계통이 포함되어 있어 집단의 다양성을 확인하였다. 또한, SSR마커를 활용한 PCR결과를 토대로 structure 분석을 수행하여 그룹을 나눌 수 있는 delta K 값을

탐색하였으며 7개의 그룹으로 나눌 수 있었다(Table 5, Fig. 3). 본 실험에서는 값이 낮아 구분하기 어려울 수 있으나 추후 분석된 SNP genotyping을 통하여 재검증되었다(Unpublished data).

Single Marker Analysis를 통한 형질별 마커 탐색
19가지의 농업형질과 출수기, 성숙기에 대하여 LOD값 5를

Fig. 2. Phylogenetic tree for the core collection 614 accession.

Table 5. Calculating Delta K for 10 repeat performance by structure harvester.

K	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	Ln''(K)	Delta K
1	-80793.590000	1.780418	—	—	—
2	-74567.380000	14.638746	6226.210000	2181.940000	149.052383
3	-70523.110000	28.942451	4044.270000	1743.270000	60.232287
4	-68222.110000	421.193710	2301.000000	640.640000	1.521010
5	-66561.750000	618.831959	1660.360000	262.870000	0.424784
6	-65164.260000	74.034274	1397.490000	383.900000	5.185436
7	-64150.670000	210.582188	1013.590000	3278.760000	15.569978
8	-66415.840000	8169.266024	-2265.170000	5325.450000	0.651888
9	-63355.560000	1775.067841	3060.280000	1450.830000	0.817338
10	-61746.110000	407.291652	1609.450000	—	—

Fig. 3. Calculating delta K of core collection (a), structure (b), simple neighbor joining tree (c) and clustering (d). Using the structure program at based PCR results discovery optimal values and separate groups. Discovering optimal value delta K.

기준으로 하여 총 1,450 개의 연관마커 후보군을 선발하였으며, 도복과 한해와 같은 양적형질에 대해서는 158 (선발LOD 평균: 18.7), 32(17.3)개의 많은 마커를 발견하였으며, 옆폭, 간색의 경우에는 각각 25 (6.9), 18 (6.9)개가 탐색되었다. SDS 침전가와

Table 6. Single-marker-analysis by agricultural trait for 614 resources core collection candidate.

Trait	Single-marker No.	Average of selected LOD
Cold damage (한해)	132	17.3
Plant type (초형)	64	8.8
Upright habit degree of flag leaf (지엽수부)	79	15.1
Seed coat color (종피색)	45	10.2
Seed (종자량)	81	10.7
Seedling stand (입모)	59	11.6
Leaf width (옆폭)	25	6.9
Leaf length (옆장)	66	8.7
Auricle (엽이)	50	9.1
Ear length (수장)	85	10
Ear color (수색)	43	10.9
Awn length (망장)	44	7.9
Awn color (망색)	48	10.9
lodging (도복)	158	18.7
Culm length (간장)	52	16
Culm color (간색)	18	6.9
SDS	36	6.8
Protein	29	7.6
ASH	134	15.4
Heading date (출수기)	97	10.8
Ripening date (성숙기)	105	12.6

단백질 함량과 연관된 마커는 36 (6.8)개와 29 (7.6)개의 마커만 탐색되었으나 회분함량(ASH)는 134 (15.4)개의 마커가 발견되었다. 국내 재배환경의 영향을 받는 출수기와 성숙기의 경우 각각 97(10.8), 105(12.6)개의 마커가 탐색되었다(Table 6, Fig. 4). 90개가 넘는 마커가 탐색된 한해, ASH, 출수기, 성숙기와 같은 형질에서는 추후 중요한 loci에서 SNP를 탐색하여 기존의 마커(SSR, RFLP, AFLP, RAPD 마커 등)과 유사하지만 표현형과 직접적으로 관련있는 유전자 자체의 존재 여부를 판단할 수 있는 유전자 특이 마커를 개발 하여 향후 밀 분자유종에 기여할 것으로 사료된다(Choi et al. 2018). 또한, 단백질 함량과 연관인 있는 protein, ASH의 경우에는 SNP가 속한 유전자를 발굴하여 유전자의 발현량 조절을 통해 좀더 고품질 밀 육성에 기여할 것으로 사료된다. 최근에는 밀가루의 단백질 함량은 가곡 적성을 결정하기 위하여 핵심집단 후보군 614점에 대하여 NIR 을 활용하여 침전가 및 단백질 함량간의 상관관계를 분석하여

Fig. 4. A total of 21 agricultural traits for single marker analysis at the core collection 1,967 accession. The horizontal axis is the marker name and vertical axis is the size of the value. Red line mean: cut off LOD 5.

박력분, 중력분, 강력분을 구분하였다(Yang et al. 2021). 이렇게 구축된 핵심집단 유전자원들의 다른 농업형질과 탐색한 마커를 활용하면 세밀한 육종이 가능할 것으로 사료되며 추후 SNP

chip 제작 및 GWAS (전장유전체연관분석)에 활용 될 것으로 기대된다.

적 요

본 연구는 해외에서 구축된 밀 핵심집단을 바탕으로 국내 재배환경에 적합한 밀 핵심집단 구축을 위하여 37세트의 SSR마커를 활용하여 핵심집단을 구축하였고 이를 바탕으로 국내 밀 수집 유전자원을 대표할 수 있는 육종집단을 구축하고자 하였다. 국내 밀 유전자원 및 계통의 모든 다양성을 포함하게 된다면 유전적 거리가 먼 야생 근연종 위주로 구성될 가능성이 있으므로 실제 연구에 필요한 유전자원을 대표하는 계통들의 일관성이 가능한 높아야 한다(Brown & Spillane 1999, Ma et al. 2018). 핵심집단의 경우 개체수를 모집단의 10%, 유전자좌의 70% 이상을 보유하고 있어야 하나 본 연구에서는 밀 유전체의 크기가 17 Gbp에 달하는 거대 배수성 작물이기 때문에 핵심집단의 크기가 모집단의 30%의 규모로 구축되었으며 사용자의 목적에 따라 SSR마커를 차용하여 핵심집단 후보를 선발하고 농업특성 평가를 조사하였다. 추후 정밀한 SNP genotyping을 통하여 mini core collection 구축이 가능하며 가공적성, 수발아 저항성 등 다양한 목적형질에 따른 맞춤형 밀 핵심집단의 구축으로 인하여 밀 육종 및 생산력 증대에 기여할 것으로 예상된다.

사 사

본 논문은 농촌진흥청 연구사업(과제명: 밀 핵심집단을 이용한 내염성 및 내한발생 소재 개발, 연구번호: PJ015964022021)에 의해 지원 받아 작성되었음.

REFERENCES

- Balfourier F, Roussel V, Strechelnko P, Exbrayat-Vinson F, Sourdille P, Boutet G, Koenig J, Ravel C, Beckert M, Charmet G. 2007. A worldwide bread wheat core collection arrayed in a 384-well plate. *Theor Appl Genet* 114: 1265-1275.
- Bang KH, Jo IH, Chung JW, Kim YC, Lee JW, Seo AY, Park JH, Kim OT, Hyun DY, Kim DH, Cha SW. 2011. Analysis of genetic polymorphism of Korean ginseng cultivars and foreign accessions using SSR markers. *Korean J Medicinal Crop Sci* 19: 347-353.
- Beukelaer HD, Smykal P, Davenport GF, Fack V. 2012. Core Hunter II: Fast core subset selection based on multiple genetic diversity measures using mixed replica search. *BMC Bioinformatics* 13: 312.
- Brown AHD, Spillane C. 1999. Implementing core collections-principles, procedures, progress, problems and promise. pp 1-9. In: Johnson RC, Hodgkin T. (Eds) *Core Collections for Today and Tomorrow*. IPGRL, Rome, Italy.
- Cho GT, Yoon MS, Lee JR, Baek HJ, Kang JH, Kim ST, Paek NC. 2008. Development of a core set of Korean soybean Landrace [*Glycine max* (L.) Merr.] *Journal of Crop Sci Biotech* 11: 157-162.
- Cho SW, Kang TG, Kang SW, Kang CS, Park CS. 2018. Assessment of DNA markers related to days to heading date, tiller number, and yield in Korean wheat populations. *Korean J Breed Sci* 50: 211-223.
- Choi CH, Yoon YM, Son JH, Cho SW, Kang CS. 2018. Current status and prospect of wheat functional genomics using Next Generation Sequencing. *Korean J Breed Sci* 50: 364-377.
- Choi MS, Kim DL, Ohk HC, Jeong NH, Lee NR, Kang ST, Kim ES, Ko HM, Kim KH, Kim MH, Kwak HJ, Song JY, Joo HE, Jeong JY, Jeong SH, Baek IS, Kim JH, Lee KH, Jeong CR. 2018. Establishment of soybean genomic breeding platforms by genome-wide association studies and practical application. Rural Development Administration (RDA) National Institute of Crop Science. Korea.
- Chung JW, Park YJ. 2009. Genetic diversity and population structure of Korean rice core collection. *Journal of Korean Soc Int Agric* 21: 282-288.
- Falush D, Stephens M, Pritchard JK. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164: 1567-1587.
- Earl DA, vonHoldt BM. 2012. Structure harvester: A website and program for visualizing structure output and implementing the evanno method. *Conserv Genet Resour* 4: 359-361.
- Hong WJ, Khaing AA, Park YJ. 2013. Cultivar identification of chrysanthemum (*Dendranthema grandiflorum* Ramat.) using SSR Markers. *Journal of Korean Soc Int Agric* 25: 385-394.
- Hu X, Wang J, Lu P, Zhang H. 2009. Assessment of genetic diversity in broomcorn millet (*Panicum miliaceum* L.) using SSR markers. *J Genet Genomics* 36: 491-500.
- Jeong SM, Kim JY, Jeong SC, Kang ST, Moon JK, Kim NS. 2017. GenoCore: A simple and fast algorithm for core subset selection from large genotype datasets. *PLoS ONE* 12: e0181420.
- Kang CS, Cheong YK, Kim BK. 2016. Current situation

- and prospect of Korean wheat industry. Food Sci Nutr 21: 20-24.
16. Kang CS, Cheong YK, Kim KH, Kim HS, Son JH, Kim KH, Park JC, Kim DH, Choi JK, Bae JS, Kim KJ, Lee CK, Park KG, Kim BK, Park KH, Park CS. 2015. A wheat variety, 'Goso' with low protein, good cookie, red grain wheat and resistance to pre-harvest sprouting. Korean J Breed Sci 47: 330-338.
  17. Kwon JK, Lee HY, Han KE, Han JW, No NY, Heo OS, Kim SK, Lee KA, Jeong JW, Lee JH. 2015. Analysis of genetic diversity and construction of core collection based on molecular markers and phenotype evaluation in *Capsicum*. TRKO201500010564. Rural Development Administration (RDA), Seoul National University. Korea.
  18. Kim BW, Sa KJ, Park KJ, Park JY, Lee JK. 2015. Genetic analysis of core sets of colored maize and non-colored maize inbred lines using SSR markers. Korean J Breed Sci 47: 54-62.
  19. Kim JY, Moon JC, Baek SB, Kwon YU, Song KT, Lee BM. 2014. Genetic improvement of maize by marker-assisted breeding. Korean J Crop Sci 59: 109-127.
  20. Kim KW, Chung HK, Cho GT, Ma KH, Chandrabalan D, Gwag JG, Kim TS, Cho EG, Park YJ. 2007. PowerCore: A program applying the advanced M strategy with a heuristic search for establishing core sets. Bioinformatics. 23: 2155-2162.
  21. Kim KH, Kang CS, Cheong YK, Park CS, Kim KH, Kim HS, Kim YJ, Yun SJ. 2012. Selection of Korean wheat germplasms with low polyphenol oxidase using polyphenol oxidase-specific DNA markers. Korean J Breed Sci 44: 11-18.
  22. Kim KM, Kang CS, Kim YK, Kim KH, Park JH, Yoon YM, Park HH, Jeong HY, Choi CH, Park JH, Kim YJ, Cheong YK, Han OK, Park TI. 2020. Past and current status, and prospect of winter cereal crops research for food and forage in Korea. Korean J Breed Sci Special Issue: 73-92.
  23. Liu XB, Li J, Yang ZL. 2018. Genetic diversity and structure of core collection of winter mushroom (*Flammulina velutipes*) developed by genomic SSR markers. Hereditas 155.
  24. Lynch M, Walsh B. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Sunderland Ma, Inc.
  25. Ma KH, Yoon MS, Lee KJ, Upadhyaya HD., Kwak DY, Ko JY, Baek HJ. 2018. Establishment and utilization of core and mini core collections of the mandated crops at the international crops research institute for the semi-arid tropics (ICRISAT). J Korean Soc Int Agric 30: 184-192.
  26. Min SK, Choi BW, Park JH, Chung JW, Kim KW, Park YJ. 2016. Assessment of genetic diversity and population structure of the sub core set in sesame (*Sesamum indicum*) using SSR markers. J Korean Soc Int Agric 28: 73-83.
  27. Moon JK, Seo MJ, Kim KH, Jeong HJ, Kang ST, Yoo MH, Kim KR, Kim SM, Park SM, Ko YS, Jeong SC, Baek IS, Lee KH, Lee YK, Kim SH, Kim DS, Kim JB, Ha BK. 2015. Exploration of core collection and useful soybean lines using high-density SNP chip developed from variation of Korean germplasms. TRKO201500010572. Rural Development Administration (RDA) Final research Report. Korea.
  28. Nei M. 1973. Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci U.S.A 70: 3321-3323.
  29. Oh SJ, Choi YM, Yoon HM, Lee, SK, Hyun DY, Shin MJ, Lee MC, Chae BS. 2019. Statistical analysis of protein content in wheat germplasm based on near-infrared reflectance spectroscopy. Korean J Crop Sci 64: 353-365.
  30. Park EH, Toan BD, Kim MJ, Hwang SP, Park KO, Lee GJ, Woo SH, Yoon ST, Lee SG. 2011. Genetic characterization of the soybean nested association mapping population. TRKO201400000479. Rural Development Administration (RDA) Yeongnam university. Korea.
  31. Paux E, Roger D, Ekatherina B, Gay G, Bernard M, Soudille P, Feuillet C. 2006. Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. Plnat J 48: 463-474.
  32. Perrier X, Jacquemoud-Collet JP. 2006. DARwin software. <http://darwin.cirad.fr/>
  33. RDA (Rural Development Administration). 2012. Manual for Standard Evaluation Method in Agricultural Experiment and Research. pp. 316-365.
  34. Safar J, Simkova H, Kubalakova M, Cihalikova J, Suchankova P, Bartos J, Dolezel J. 2010. Development of chromosome specific BAC resources for genomics of bread wheat. Genome Res 129: 211-223.
  35. Sim SC, Na Js, Ngan PT, Thim NT, Nam NN, Kim JE, Lee JB, An SM, Park JH, Kim DY, Kim KN, Jo JH, Lee JY, Park YH, Park TS, No MY, Jo MC, Kim OR, Kim JH, Moon JH, Yang EY, Chae SY, Lee ES, Kwon MS, Lee IS. 2018. Development of genome-based techniques for breeding by design in tomato. TRKO201800043629. Rural Development Administration (RDA), Sejong university Final

- research report. Korea.
36. Soller M, Brody T, Genizi A. 1976. On the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor Appl Genet* 47: 35-39.
  37. Son JH, Kim KH, Shin SH, Kim HS, Kim NS, Hyun JN, Shim SI, Lee CK, Park KG, Kang CS. 2013. ISSR-derived molecular markers for Korean wheat cultivar identification. *Platn Breed Biotech* 1: 262-269.
  38. Son JH, Yang JW, Kang CS, Kim KH, Kim KM, Jeong HY, Park JH, Son JY, Park TI, Choi CH. 2021. Annual analysis of agronomic traits for various wheat germplasm collected from the world in domestic environment. *Korean J Crop Sci* 66: 120-129.
  39. Song Q, Yan L, Quigley C, Jordan BD, Fickus E, Schroeder S, Song BH, An YQC, Hyten D, Nelson R, Rainey K, Beavis WD, Specht J, Diers B, Cregan P. 2017. Genetic characterization of the soybean nested association mapping population. *Platn Genome* 10: 1-14.
  40. Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 10: 512-526.
  41. Yang JW, Park JH, Son JH, Kim KH, Kim KM, Jeong HY, Kang CS, Son JY, Park TI, Choi CH. 2021. Protein and arabinoxylan contents for whole grains from wheat genetic resources 7 cultivated in Korea. *Korean J Crop Sci* 66: 29-36.