

수박계통간 염색체수준의 유전적변이 분석

김윤성 · 고찬섭 · 양희범 · 강순철

Genome-wide analysis of sequence variations in eight inbred watermelon lines

Youn-Sung Kim · Chan-Sup Ko · Hee-Beom Yang · Sun-Chul Kang

Received: 16 June 2016 / Revised: 20 June 2016 / Accepted: 20 June 2016
© Korean Society for Plant Biotechnology

Abstract To investigate the genetic basis of phenotypic differences, sequence variations were analyzed in 8 inbred watermelon lines by re-sequencing. The number of sequence variations differed depending on the chromosome. Only 12.9% of SNPs were found within genes, whereas the rest were detected in promoter or intergenic regions. SNP density analysis showed that there was a highly variable region at the end of chromosome 6, which is similar to previously published findings. However, this region with high SNP density did not show much variation between the lines. In contrast, highly conserved regions with a size of 6.5-10 Mb were found in chromosomes 10 and 11. Pathway analysis suggested that the DIMBOA (a natural antibiotic)-glucoside degradation pathway was significantly different between the lines, indicating that the eight lines may have different levels of pathogen resistance. Among the carbohydrate-related genes, the alpha-galactosidase gene was the most variable among the lines. Information from this study will be helpful in understanding the watermelon breeding process at the molecular level.

Keywords Watermelon, NGS, SNP, DIMBOA, alpha-galactosidase

서론

수박은 박과에 속하는 주요 과채류 중 하나이다. 전세계 채소류 생산량의 약 9.5%를 차지하는 것으로 알려져 있으며 연간생산량은 약 109M 톤(2013년 기준) 정도이다(FAOSTAT). 국내에서는 재배면적이 약 1만5천헥타 정도되며 연간 생산량은 64만톤(2012년 기준)으로 알려져 있다(Crop production survey). 국내 수박종자 시장은 과채류 중에서 6번째로 큰 것으로 알려져 있다.

수박에는 라이코펜 등 유용성분을 많이 포함하고 있어 국내에서는 대표적으로 많이 소비되는 과일중 하나이다(Perkins-Veazie et al. 2006; Yoo et al. 2012). 따라서 종자회사에서는 다양한 수박품종을 육종하고 있다. 이러한 수박 육종이 지속되면서 각 회사가 보유하고 있는 재료의 다양성이 많이 감소된 것으로 알려져 있다. 심지어 미국 계통인 Charleston Gray와 중국 계통인 9710을 비교해 보면 평균1430 bp 마다 1개의 SNP가 발견되었다(Guo et al. 2013). 그러나 야생수박은 육성된 품종들에 비해 유전적 다양성이 더 컸다. 브라질에서 확보된 야생형 수박들은 지역에 따라 다양성이 다른 결과를 보여주었다(Gama et al. 2013) 이러한 수박의 유전적 다양성은 옥수수, 콩, 벼에 비해 상당히 낮은 것으로 평가되었다(Guo et al. 2013).

Guo 등은(2013) 수박의 유전체 염기 서열을 밝혀내고 20개의 수박 계통(야생형, 재배형 등)의 유전적 다양성을 re-sequencing을 통해 상세하게 분석하였다. 특히 야생형인 *C. lanatus* subsp. *Mucospermus*와 주요 품종의 계통인 *C. lanatus* subsp. *Vulgaris*을 비교 분석하였다. 두 계통에서 가장 차이가 많이 나는 상위 1% 지역은 약 7.78 Mb 정도 되고 이 지역에는 741개의 유전자가 있었다. 이들 유전자들에 대한 Gene Ontology 분석을 통해 수박육종 과정 중에 선택적으로 많아진 대사과정을 알 수 있었다.

Y.-S. Kim (✉) · C.-S. Ko · H.-B. Yang · S.-C. Kang
농협종묘센터
(Nong-Hyup Seeds, An-Seong-si, 17558, Korea)
e-mail: yskim0907@hanmail.net

C.-S. Ko
고려대 생명과학대학
(College of Life science and biotechnology, Korea University,
Seoul, 136-713, Korea)

Table 1 Characteristics of the eight inbred lines studied

Line No.	Fruit shape	Color of fruit	Stripe	Flesh color
163	Circular	Black	No	Red
434	Broad elliptic	Green	Yes	Dark red
435	Elliptic	Green	Yes	Orange
437	Elliptic	Yellow	Yes	Dark orange
438	Circular	Green	No	Dark orange
439	Circular	Black	Yes	Dark red
442	Circular	Green	Yes	Red
443	Circular	Green	Yes	Dark pink

Nimmakayala 등은(2014) 아프리카, 아시아, 유럽 등에서 사용되는 183개 수박 계통의 유전적 다양성을 genotyping-by-sequencing을 통해 분석하였다. 각 수박계통은 3개의 그룹으로 분리되었고 그 중 두 그룹은 아프리카 계통들로만 구성되어 있었다. 이러한 결과는 아시아와 유럽에서 아직 품종육성에 사용하지 않은 아프리카 계통이 있음을 시사하는 것이다. 또한 아프리카 및 아메리카 계통이 아시아와 유럽 계통보다 유전적 다양성이 더 크다는 것을 확인하였다. 선발과정 중에 특히 염색체 3,9번이 선택적으로 선발된 증거도 발견하였다.

한편 ISSR (inter-simple sequence repeat) 마커를 활용하여 국내 품종의 유전적 유연관계를 분석한 결과 품종간 유사도가 아주 높게 나왔다(권용삼 외, 2006). 그러나 국내에서 육성된 재료의 유전적 다양성에 대한 연구는 대체적으로 부족한 편이다. 본 연구에서는 본격적으로 마커 개발을 하기 전에 수박 8개 계통에 대해 re-sequencing을 수행하여 계통간 유전적 다양성을 지능 수준에서 분석하였다. 계통간 상당히 상이한 지역 뿐 아니라 잘 보존된 지역도 발견할 수 있었으며 대사 경로 분석을 통해 의미있게 차이나는 대사 과정을 확인할 수 있었다.

재료 및 방법

수박품종

농협종묘가 보유하고 있는 특성이 서로 다른 8개 계통(163, 434, 435, 437, 438, 439, 442,443)을 선택하여 사용하였다. 8개 계통의 대표적인 특성을 Table 1에 정리하였다.

NGS

Next Generation Sequencing (NGS)를 수행하기 위해 수박 잎에서 kit(바이오니아, 한국)를 사용하여 genomic DNA를 추출하였다. 추출방법은 kit내 사용방법을 따랐다. 추출된 DNA

는 agarose gel 전기영동으로 분해여부를 확인하였다. NGS는 테라젠바이오에 의뢰하여 진행하였으며 일차적인 SNP 및 indel 추출 또한 테라젠바이오에서 수행되었다. 얻어진 SNP, indel 중 모든 계통에서 호모인 것만 선발하였고 또한 각 계통의 depth가 10 이상인 것들만 선발하여 이후 분석에 사용하였다.

서열분석

서열분석은 R을 활용하여 수행하였다. 계통간 유사성 분석 및 클러스터링은 R에서 dist, hclust를 사용하여 분석하였다. SNP밀도 계산은 1 kb 당 SNP의 개수를 측정하였으며 그 결과는 R을 이용하여 그래프로 표시하였다. SNP분포를 색으로 표시하기 위해 먼저 각 계통별 SNP를 이용하여 클러스터링을 수행한 후 기준서열과 동일한 SNP는 연두색, 서열이 다르면 빨강색으로 표시하였다.

수박 유전자에 대한 정보는 CuGenDB (www.icugi.org)의 유전정보(version 1)를 활용하였다.

Pathway 분석은 Plant MetGenMAP (bioinfo.bti.cornell.edu/cgi-bin/MetGenMAP)에서 사용방법에 따라 데이터를 입력하여 분석하였다. P value 0.05를 기준으로 분류하고 이어서 multitest correction방법으로 FDR을 적용하였다.

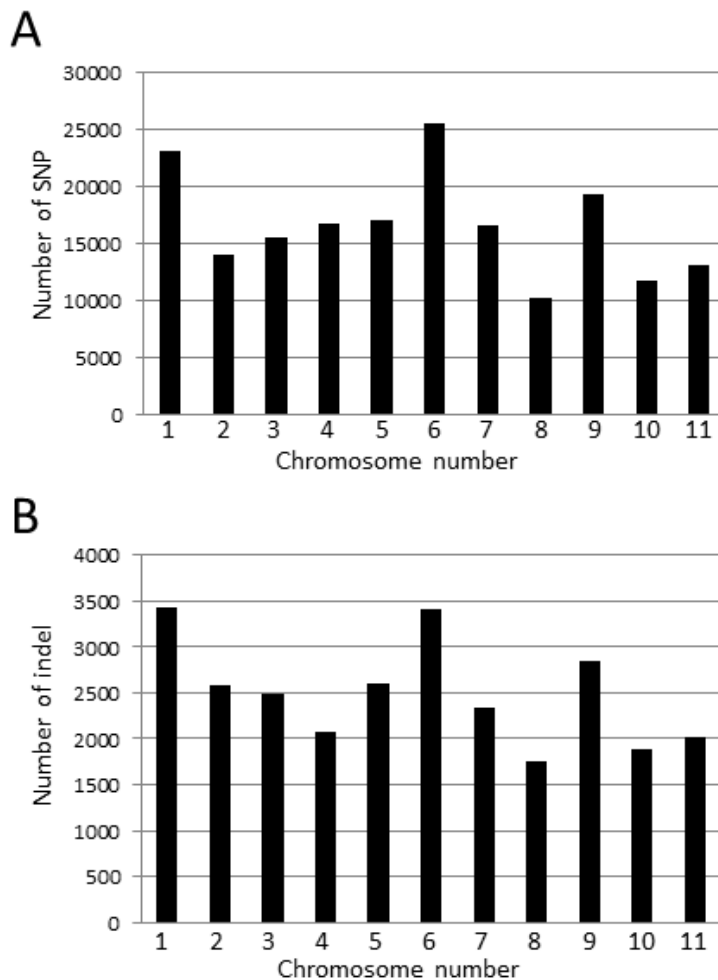
결과 및 고찰

수박 re-sequencing

마커 개발 및 특성분석에 활용하기 위해 수박 8개 계통에 대해 re-sequencing을 수행하였다. 총 읽은 read 수는 약 689백만 개 이고 총 약 68.9 Gb가 얻어졌다(Table 2). 맵핑률은 평균 92%이상이었으며 depth는 평균 22배 정도였다. 얻어진 서열결과를 이용하여 얻은 SNP는 총 183,277개였으며 indel은 27,419개였다. 염색체별로 보면 1만 ~ 2.5만개, indel은 1700 ~ 3400개 정도였다(Fig. 1). 염색체에 따른 분포를 보면 SNP

Table 2 Summary of NGS data from the 8 watermelon lines

Sample name	Total sequenced reads	Mapped reads	Mapping rate (%)	Properly paired reads	Properly paired read rate (%)	Mapping depth (x)
163	74,511,622	69,688,412	93.53	54,125,759	72.64	21.1
434	80,328,753	73,732,835	91.79	54,398,403	67.72	22.3
435	86,791,832	80,894,231	93.20	61,198,194	70.51	24.5
437	81,478,323	76,766,686	94.22	58,406,163	71.68	23.3
438	100,682,695	95,141,853	94.50	72,902,668	72.41	28.8
439	85,939,493	79,362,127	92.35	57,378,761	66.77	24.0
442	89,863,239	83,042,565	92.41	60,266,949	67.07	25.2
443	89,370,475	83,214,827	93.11	62,515,685	69.95	25.2

**Fig. 1** Number of sequence variations in each chromosome. (A) Number of SNPs. (B) Number of indels

및 indel 모두 비슷한 패턴을 보여주었다. 염색체 1번과 6번에서 SNP와 indel이 많이 발견되는 반면 염색체 8번에서는 가장 적은 수가 발견되었다. 이러한 차이는 염색체의 크기와는 상관이 없어 보인다. 염색체 1번과 6번의 크기는 약 34 M bp와 27 M bp이며 8번은 26 M bp로 6번과 8번이 크기가 비슷하나 SNP, indel의 수는 상당한 차이가 있었다.

동아시아지역 6개 계통을 대상으로 한 NGS 결과에서는

약 59만개의 SNP와 10만 개의 indel을 찾을 수 있었다(Guo et al. 2013). 동일연구에서 미국의 5계통에서는 37만 개의 SNP, 9만4천개의 indel이 찾아져서 본 연구에서 얻은 SNP나 indel 보다 많은 수가 발견되었다.

각 염색체별 SNP분포를 보다 세밀하게 조사하기 위해 SNP밀도를 계산해 보았다. 전체적으로 보면 1 kb 당 0.55개의 SNP가 발견되었고 indel의 경우 0.083개가 발견되어 SNP

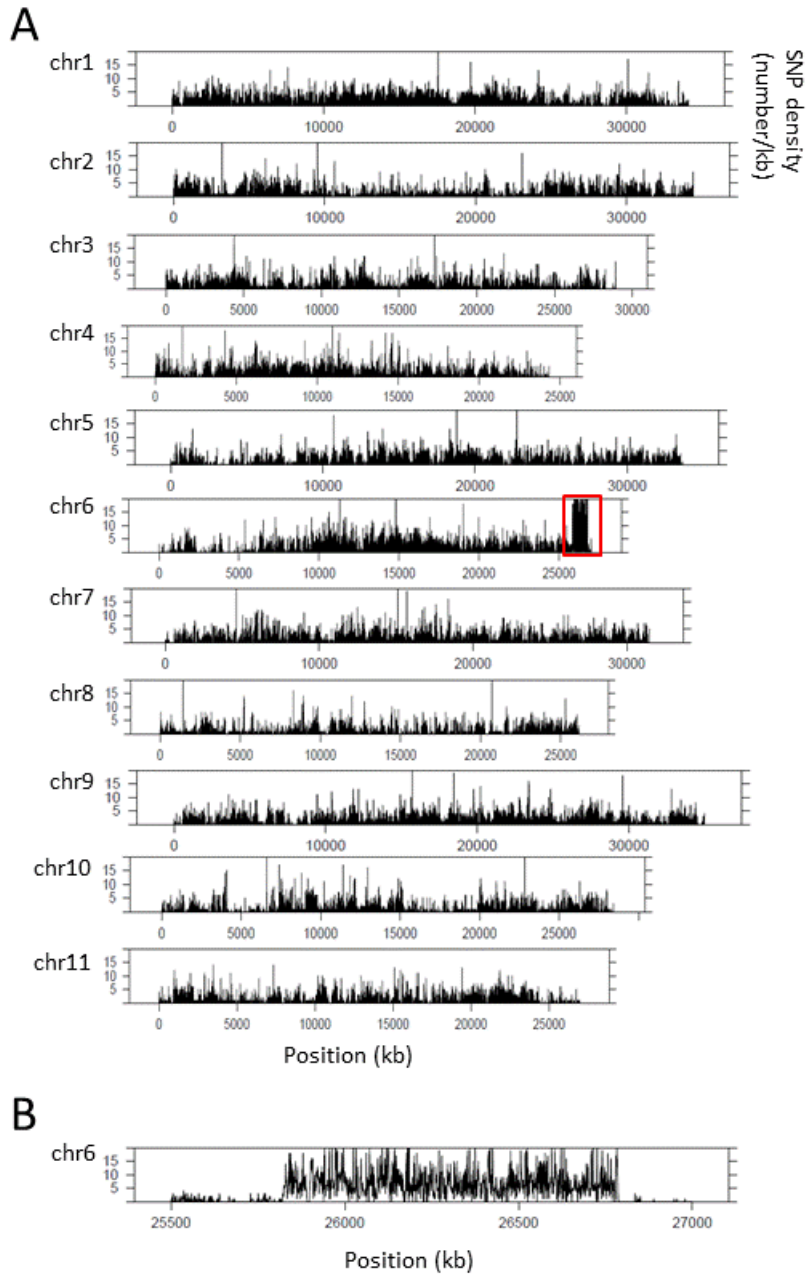


Fig. 2 SNP density map. SNP density was measured as the number of SNPs per 1 kb. (A) SNP density in each chromosome. The red square indicates a region with particularly high SNP density. (B) Enlarged view of the high-SNP density region in chromosome 6

와 indel의 밀도는 6.6배 정도 차이가 났다. 1개의 SNP가 발견되는 평균길이는 1,804 bp로서 미국품종과 중국품종간 SNP밀도와(1,430 bp당 1개의 SNP) 유사하였다(Guo et al. 2013). 한편 각 염색체별로 SNP 밀도가 상당히 높은 곳을 발견할 수 있었다. 염색체 8번은 대체적으로 낮은 SNP 밀도를 보여주었으며 염색체 6번의 끝부분은 특이하게도 SNP 밀도가 높은 곳이 집중되어 있었다(Fig. 2A). Guo 등(2013)이 *C. lanatus* subsp. *vulgaris*, subsp. *mucospermus*에서 염색체별 다양성을 조사한 결과를 보면 염색체 3번의 중간지역과 6번

의 끝 부분이 가장 큰 다양성을 보여주었다. 본 연구에서는 염색체 3번의 SNP 밀도가 다른 염색체와 큰 차이가 없었던 반면 염색체 6번의 끝부분에서는 Guo 등의 연구결과와 유사한 결과를 얻을 수 있었다(Fig. 2). 염색체 6번의 SNP고밀도지역은 1.1M bp 정도의 크기이며 이 지역의 전후로 SNP 밀도가 급격하게 떨어졌다(Fig. 2B). 이 지역에 있는 유전자 수는 133개로 육종과정에 큰 영향을 받지 않는 유전자로 생각된다.

염색체 6번 SNP고밀도지역에서 계통간 SNP 분포를 살펴

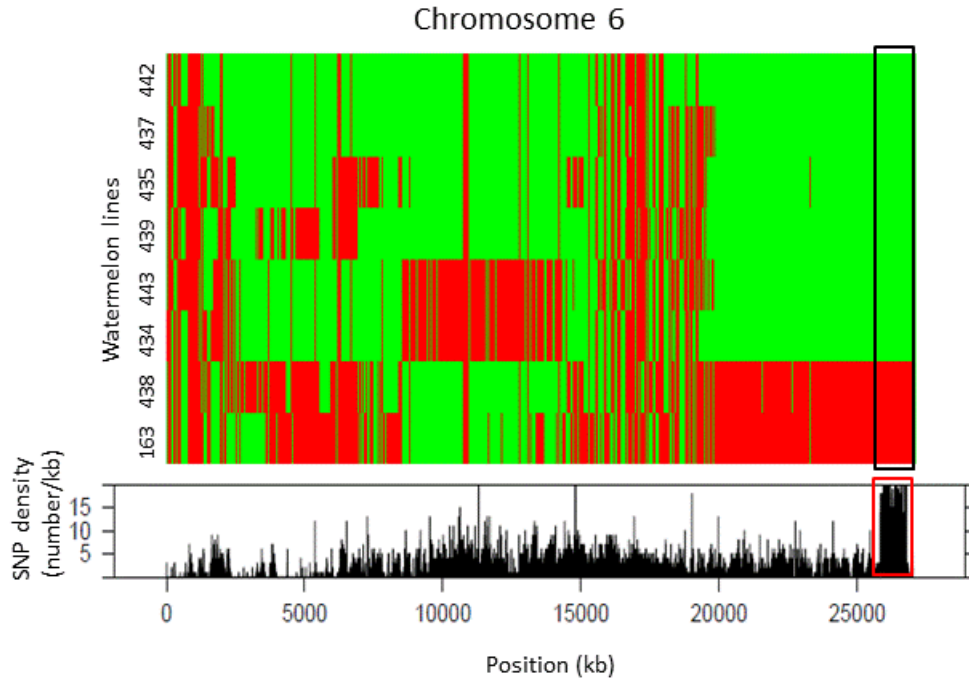


Fig. 3 Graphical representation of SNP distribution in chromosome 6. Upper, SNP distribution; green indicates agreement with the sequence of the reference genome and red indicates differences from the reference genome. The X and Y axes represent the position of SNPs and watermelon lines, respectively. Lower, SNP density as described in Fig. 2. The boxed region indicates the region with high SNP density

보기 위해 염색체 6번의 SNP분포를 색으로 표시하여 보았다(Fig. 3). 염색체 6번 SNP 고밀도 지역의 SNP는 크게 두 그룹으로 나눌 수 있었다. 163번, 438번 계통이 한 그룹을 형성하였고 나머지 계통은 두 번째 그룹으로 나뉘어 졌다. SNP 고밀도지역에서 그룹 내 서열차이가 없고 그룹간 차이만 있었다. 즉 SNP고밀도 지역이 계통간 다양성이 높은 지역은 아니었다.

나머지 염색체에 대해서도 SNP분포를 색으로 표시하여 보았다. 특이하게도 염색체 10번과 11번의 경우 계통간 차이가 없는 큰 지역을 확인할 수 있었다(Fig. 4). Indel도 동일한 지역에서 차이가 전혀 없는 큰 지역을 확인할 수 있었다. 이 큰 보존지역의 크기는 염색체 10번은 약 10M bp, 11번은 약 6.5M bp 정도였다. 이러한 보존지역이 다른 수박 계통에서도 일반적으로 발견되는지, 이렇게 보존된 특별한 이유가 있는지는 연구가 더 필요하다고 생각된다. 이 지역에 있는 SNP나 indel은 품종판별용으로 사용하기에 적합하지 않을 것이다.

계통간 유사성 분석

SNP나 indel/SSR은 품종판별, 순도검정, QTL 맵핑 등에 널리 사용되고 있다. 또한 계통 혹은 품종간 유전적 유사성을 조사할 때도 사용되고 있다. 일반적으로는 정해진 수의 마커만을 사용한다. 본 연구에서는 전체 서열을 확보하였으

므로 이를 바탕으로 계통간 유사성을 조사해 보았다. 먼저 염색체별로 조사하여 본 결과 염색체별로 서로 다른 결과를 보여주었다. 예를 들어 염색체 1번에서는 434번 계통과 443계통이 가장 가깝게 나오나 2번 염색체에서는 상당히 멀게 나타났다(Fig. 5A). 모든 염색체의 SNP 결과를 모아서 유사성을 분석한 결과를 Fig. 5B에서 보여주고 있다. 염색체 1번의 결과와 비교해 볼 때 동일한 결과도 있으나 서로 다른 결과도 보여준다. 전체적으로는 다른 결과를 보여주고 있다. 따라서 마커를 활용하여 계통간 유사성을 분석할 때 마커의 수가 결과에 상당한 영향을 미칠 것으로 생각된다.

유전자 변이

각 계통간 변이가 주로 어디에서 일어나는지 조사해 보기 위해 유전자 내 SNP 수 및 PIC 값을 조사해 보았다. Cucurbit 지놈 데이터베이스(CuGenDB)에서 얻은 수박 유전자는 총 22,682 개였으며 SNP가 1개 이상 발견된 유전자는 7,178개로 전체 유전자의 31.6%였다. 한편 SNP가 10개 이상 발견된 유전자는 총 377개로서 전체의 1.7%에 불과하였다. SNP가 단 한 개도 발견되지 않은 유전자는 총 15,504개로 전체의 68.4%에 달하였다. 유전자 내에서 발견된 SNP수는 23,663 개로 전체 SNP의 12.9%를 차지하였다. 따라서 대다수의 SNP가 프로모터 및 유전자 사이의 지역에서 발견되는 것을 알 수 있었다.

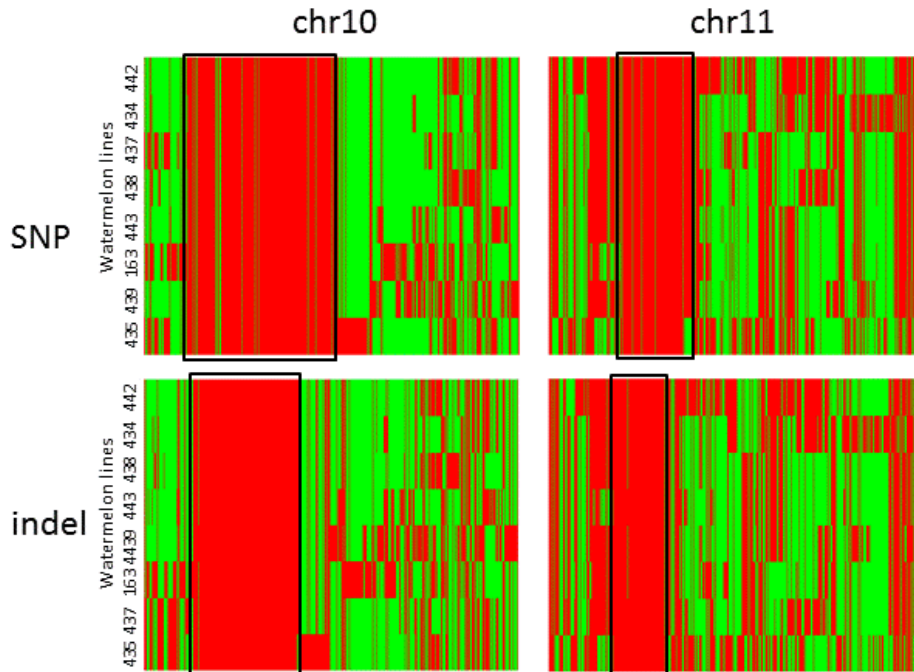


Fig. 4 Graphical representation of SNP and indel distribution in chromosomes 10 and 11. The X and Y axes represent the position of SNPs and watermelon lines, respectively. Green indicates SNPs with the same sequence as the reference genome, and red indicates differences. Black boxes indicate the regions with highly conserved sequences between the lines

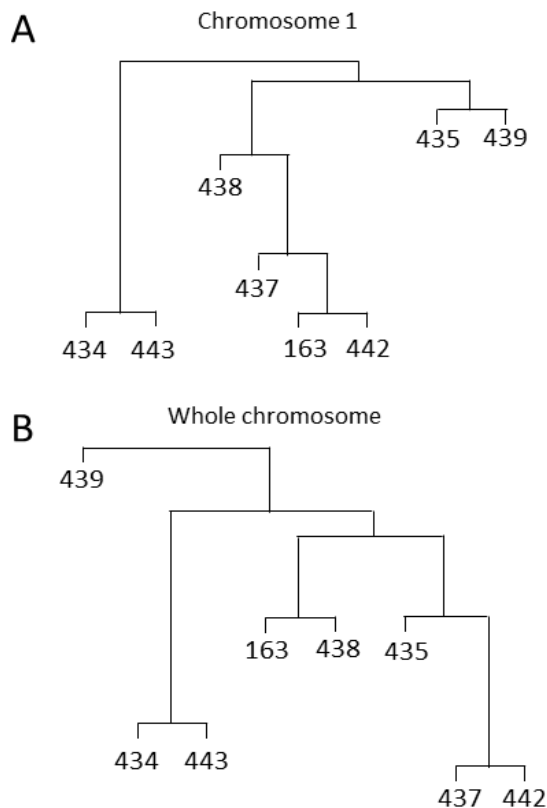


Fig. 5 Dendrogram of the eight watermelon lines. Hierarchical clustering was performed using the R statistical programming language. (A) Dendrogram based on a SNP in chromosome 1. (B) Dendrogram based on SNPs from all chromosomes

염색체 6번의 SNP고밀도 지역에 있는 유전자들의 SNP 개수를 조사해 보았다. 염색체 6번에 있는 유전자 내에 있는 SNP 개수의 평균은 2.1개였으나 고밀도 지역의 SNP 개수 평균은 17개로 8배 이상의 SNP가 있는 것을 확인할 수 있었다. SNP가 가장 많이 발견되는 유전자 10개 중 6개가 염색체 6번에 위치하고 있으며 그 유전자 6개 모두 SNP 고밀도 지역에 있는 유전자였다(Table 3, Fig. 6A).

SNP가 10개 이상인 유전자 리스트를 이용하여 SNP가 특정 대사경로 유전자에 많이 나타나는지 조사해 보았다. P value 0.05 기준으로 DIMBOA-glucoside degradation pathway (p value=0.023)만 통계적으로 의미있는 경로로 얻어졌다. DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxane-3-one)는 hydroxamic acid로 박테리아 뿐 아니라 곰팡이, 곤충에 대해서도 작용하는 항생제로 알려져 있다(Couture et al. 1971; Long et al. 1975). 이러한 결과는 수박 계통간 DIMBOA 양에 차이가 있을 가능성을 나타내는 것이며 더불어 병해충에 대한 저항성도 차이가 날 가능성이 있음을 시사한다.

수박에서 당도는 가장 중요한 특성 중 하나이다. 수박의 당도는 설탕, 포도당, 과당 등의 함량에 의해 결정된다. Guo 등(2015)은 재배종 수박과 야생형 수박의 유전자 발현을 비교 분석하여 당관련 유전자 중에서는 45개 유전자의 발현이 의미 있게 변한 것을 확인하였다(Guo et al. 2015). 이러한 유전자에는 invertase, sucrose synthase, sugar transport protein 등이 포함되어 있다. 이들 유전자들의 변이는 품종간 당함량 차이를 설명해 줄 수 있는 단서가 될 수 있으므로 앞서

Table 3 Number of SNPs in selected genes

Gene	Number of SNP	Chromosome	Function
Cla019196	239	6	Serine/threonine protein kinase ATM
Cla019198	148	6	Unknown
Cla019258	146	7	Light dependent short hypocotyls 10
Cla013440	93	2	Strawberry notch
Cla019342	71	3	Ankyrin repeat containing protein
Cla019217	65	6	Unknown
Cla019297	65	6	Unknown
Cla017464	61	10	ADP-ribosyl ation factor GTPase activating protein
Cla019268	55	6	Unknown
Cla019261	53	6	Transcription factor GTE10

Table 4 Number of SNPs in carbohydrate-related genes

Gene	Number of SNP	Annotation	Gene	No SNP	Annotation
Cla001115	0	Sugar transporter	Cla013113	0	Sugar transporter
Cla001346	2	Sugar transporter	Cla013521	0	Sugar transporter
Cla001496	1	Sugar transporter	Cla013571	0	Sugar transporter
Cla002328	1	Acid invertase	Cla013902	0	UDP-galactose/glucose pyrophosphorylase
Cla002358	0	Fructokinase	Cla015835	0	Sugar transporter
Cla002641	1	Sugar transporter	Cla015836	0	Sugar transporter
Cla004116	2	Sugar transporter	Cla015850	0	Neutral invertase
Cla004909	1	Sugar transporter	Cla015944	0	Sugar transporter
Cla006068	0	Sugar transporter	Cla015947	1	Sugar transporter
Cla006123	1	Alpha-galactosidase	Cla017392	1	Sucrose synthase
Cla007136	3	Sugar transporter	Cla017674	0	Cell wall invertase
Cla007286	12	Alpha-galactosidase	Cla018637	0	Sucrose synthase
Cla009124	3	Sucrose synthase	Cla019238	28	Alpha-galactosidase
Cla009350	0	Sugar transporter	Cla019300	4	Sugar transporter
Cla009857	0	UDP-glucose 4-epimerase	Cla019566	1	Sugar transporter
Cla011131	0	Sucrose synthase	Cla020872	0	Cell wall invertase
Cla011361	0	Sugar transporter	Cla021693	0	Sugar transporter
Cla011923	1	Sucrose-phosphate synthase	Cla021809	1	Neutral invertase
Cla012211	0	Alpha-galactosidase	Cla022734	0	Alpha-galactosidase
Cla012668	6	Sugar transporter	Cla022735	0	Alpha-galactosidase
Cla012809	0	UDP-glucose 4-epimerase	Cla022883	1	Alpha-galactosidase
Cla012829	0	Sugar transporter	Cla022885	5	Alpha-galactosidase
Cla012830	0	Sugar transporter			

NGS를 수행한 8개 품종간 변이 정도를 조사해 보았다. 45개 유전자 중 25개에서는 SNP가 전혀 발견되지 않았으며 가장 많은 SNP가 발견된 유전자는 Cla019238(28개) 였으며 그 다음으로 Cla007286(12개), Cla012668(6개), Cla022885(5개) 순이었다. 특이한 점은 SNP가 가장 많은 4개 유전자 중 3개가

alpha-galactosidase (Cla019238, Cla007286, Cla022885)이라는 점이다(Table 4, Fig. 6B).

Alpha-galactosidase는 galactose를 가지고 있는 올리고당으로부터 끝에 있는galactose를 분해하는 효소로 알려져 있다 (Gaudreault and Webb 1983). 이 효소는 박과에서 phloem으로

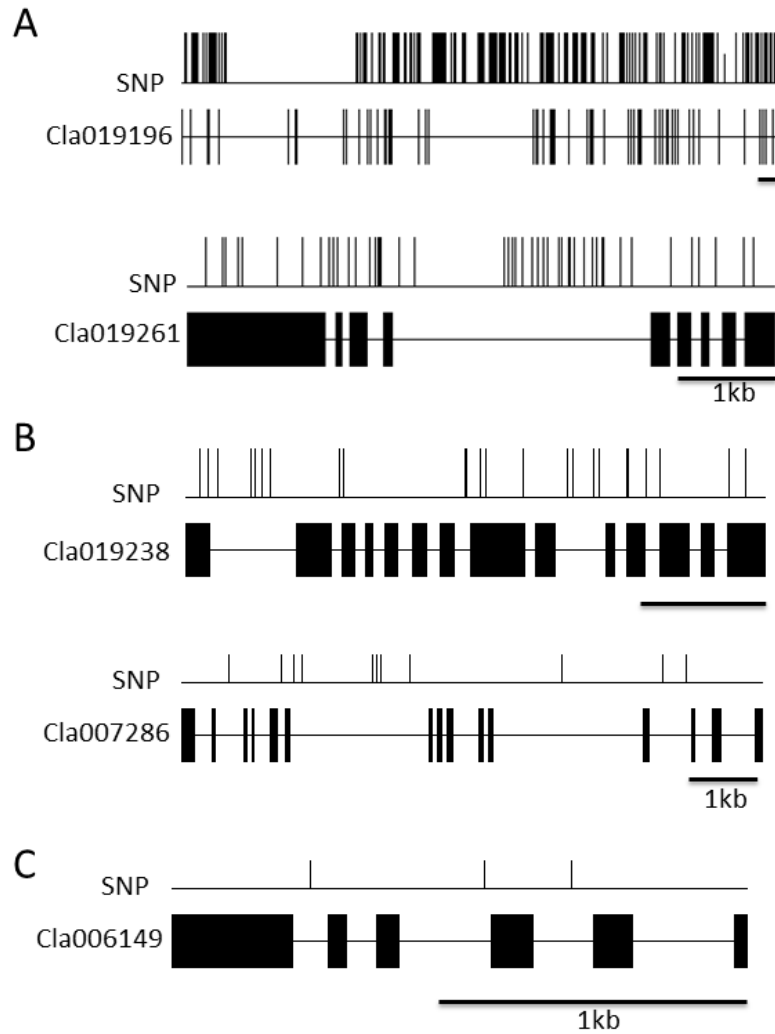


Fig. 6 Genomic structure and position of SNPs in the given genes. Vertical lines indicate the positions of SNPs, and their length indicates the PIC value. Black boxes indicates exons. Regions between the black boxes represent introns. (A) Cla006149 is a putative beta-carotene hydroxylase gene. (B) Cla019238 and cla007286 are alpha-galactosidase genes. (C) Cla019196 is a putative serine/threonine protein kinase ATM gene. Cla019261 is a putative transcription factor GTE10 gene

이송되는 주요 당인 stachyose, raffinose에서 galactose를 분해하는 역할을 한다(Richardson et al. 1982; Schaffer et al. 1996). 멜론의 경우에는 과(fruit)에서 raffinose, stachyose 농도가 아주 낮은 것으로 알려져 있어 phloem에서 이동하면 과로 전달되면서 바로 분해되는 것으로 추정되고 있다(Hubbard et al. 1989; Chrost and Schmitz, 1997). 따라서 alpha-galactosidase는 박과작물에서 당의 분배에 중요한 역할을 수행할 것으로 생각되고 있다(Keller and Pharr 1996). 수박에서도 alpha-galactosidase가 분리가 되었으나 과에서 당의 축적에 어느 정도 중요한 역할을 하는지는 아직 불분명하다(Itoh et al. 1986; Bayraktar and Onal 2013).

Guo 등의(Guo et al., 2015) 결과에 따르면 위의 3개 alpha-galactosidase 유전자의 발현은 과육이 발달하면서 잠깐 증가했다가 점차 감소하는 패턴을 보인다. 또한 재배종 수박과 야생종 수박간 발현차이도 확인되었다. 이러한 결과들

을 종합해 볼 때 alpha-galactosidase 유전자 변이는 품종간 당도차이에 상당한 영향을 줄 가능성이 있다고 생각되며 향후 당도 결정을 하는 유전자 연구에 많은 도움이 될 것으로 사료된다. 본 연구에서 발견된 3종의 alpha-galactosidase가 실제 수박품종의 당도 결정에 어떤 영향을 주는 지 연구해보면 좋을 것으로 보인다. 다만 본 연구에서는 8개 계통만 사용하였으므로 좀 더 많은 수의 계통에서 비교해 볼 필요가 있다.

수박의 과육에는 라이코펜(lycopene)과 같은 카로티노이드 함량이 높아 붉은색을 띤다. 수박에서는 카로티노이드 합성에 관련된 유전자가 26개 알려져 있으며(Guo et al., 2015) 그 중에서 7개 유전자가 재배종과 야생형 수박사이에 발현양 차이를 보여주었다(Guo et al., 2015). 이들 유전자(Table 5)에 대해 8계통에서 변이가 있는지 조사해 보았다. 7개중 4개에서는 SNP가 발견되지 않았고 3개 유전자에서는

Table 5 Number of SNPs in carotenoid-related genes

Gene	Number of SNP	Annotation
Cla005404	0	9-cis-epoxycarotenoid dioxygenase
Cla005453	0	Carotenoid cleavage dioxygenase
Cla006149	3	Beta-carotene hydroxylase
Cla009122	1	Phytoene synthase
Cla009779	0	9-cis-epoxycarotenoid dioxygenase
Cla011420	1	Beta-carotene hydroxylase
Cla015245	0	Carotenoid cleavage dioxygenase

SNP가 1~3개 발견되었다. SNP를 보이는 유전자 3개중 2개는 beta-carotene hydroxylase (Cla006149, Cla011420)였으며 나머지 하나는 phytoene synthase (Cla009122)였다. beta-carotene hydroxylase의 경우 두 유전자 모두 SNP가 intron에서 발견되었고 phytoene synthase는 exon에서 발견되었다(Fig. 6C). 따라서 beta-carotene hydroxylase (Cla006149, Cla011420)는 실험에 사용한 계통간 차이는 사실상 거의 없는 것으로 사료된다. 본 실험에 사용된 계통들간의 카로티노이드 합성 유전자 변이는 상당히 낮다는 것을 보여주는 결과이다.

적 요

수박의 형태적 변이의 유전적 원인을 분석해 보기 위해 8개 계통에서 re-sequencing을 수행하였다. 유전적 변이의 수는 염색체에 따라 다르게 나왔으며 발견된 SNP의 약 12.9%만이 유전자내에서 발견되었고 나머지는 프로모터나 유전자 사이의 지역에서 발견되었다. SNP 밀도에 대한 분석 결과 염색체 6번의 말단지역에 변이가 집중되어 있는 것을 알 수 있었다. 또한 염색체 10과 11번에 잘 보존된 지역을 발견하였다. Pathway 분석을 통해 DIMBOA(일종의 항생제)-glucoside 분해 대사가 계통간 가장 차이 나는 것으로 확인되었으며 이는 각 계통의 병저항성에서 차이가 날 가능성을 시사하는 것이다. 당대사 관련 유전자 변이를 분석한 결과 alpha-galactosidase 유전자에 가장 변이가 많은 것으로 밝혀졌다. 이러한 연구 결과는 육종을 분자수준에서 이해하는 데 도움을 줄 것으로 생각한다.

사 사

본 연구는 농림수산식품기술기획평가원의 농생명산업기술개발 사업(과제번호: 313043-03-3-HD080)의 지원으로 수행되었다.

References

- Bayraktar H, Onal S (2013) Concentration and purification of alpha-galactosidase from watermelon (*Citrullus vulgaris*) by three phase partitioning. *Separation and Purification Technology*, 118:835-841
- Couture RM, Routley DG, Dunn GM (1971) Role of cyclic hydroxamic acids in monogenic resistance of maize to *Helminthosporium turcicum*. *Physiol Plant Path* 1:515-785
- Chrost B, Schmitz K (1997) Changes in soluble sugar and activity of a-galactosidases and acid invertase during muskmelon (*Cucumis melo* L.) fruit development. *J Plant Physiol* 151: 41-45
- Crop production survey (2012) from <http://kostat.go.kr>
- FAO statistics from <http://faostat.fao.org>
- Gama RNC, Santos CAF, Dias RCS (2013) Genetic variability of watermelon accessions based on microsatellite markers. *Genetics Mol Res* 12(1):747-754
- Gaudreault PR, Webb JA (1983) Partial purification and properties of an alkaline a-galactosidase from mature leaves of *Curcubita pepo*. *Plant Physiol* 71:662-668
- Guo et al (2013) The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature genetics* 45(1):51-8
- Guo S, Sun H, Zhang H, Liu J, Ren Y, Gong G, Jiao C, Zheng Y, Yang W, Fei Z, Xu Y (2015) Comparative transcriptome analysis of cultivated and wild watermelon during fruit development. *PLoS ONE* 10(6):e0130267. DOI:10.1371/journal.pone.0130267
- Hubbard NL, Pharr DM, Huber SC (1989) Sucrose phosphate synthase and acid invertase as determinants of sucrose concentration in developing muskmelon (*Cucumis melo* L.) fruits. *Plant Physiol* 91:1527-1534
- Itoh T, Uda Y, Nakagawa H (1986) Purification and characterization of alpha-galactosidase from watermelon. *J Biochem* 99(1): 243-50
- Keller F, Pharr DM (1996) Metabolism of carbohydrates in sink and sources: galactosyl-sucrose oligosaccharides, p 157-183. In: E. Zamski, A.A. Schaffer (ed.). *Photoassimilate Distribution in Plants and Crops*, Marcel Dekker, New York
- Kwon YS, Lee WS, Cho IH (2006) Assessment of genetic

- relationship among watermelon varieties revealed by ISSR marker. *J Life Sci* 16(2):219-224
- Long BJ, Dunn GM, Routley DG (1975) Relationship of hydroxamic acid content in maize and resistance to northern corn leaf blight. *Crop Sci* 15:333-335
- Nimmakayala P, Levi A, Abburi L, Tomason YR, Saminathan T, Vajia VG, Malkaram S, Reddy R, Wehner TC, Mitchell SE, Reddy UK (2014) Single nucleotide polymorphisms generated by genotyping by sequencing to characterize genome-wide diversity, linkage disequilibrium, and selective sweeps in cultivated watermelon. *BMC Genomics*, 15:767-780
- Perkins-Veazie P, Collins JK, Davis AR, Roberts W (2006) Carotenoid content of 50 watermelon cultivars. *J Agric Food Chem* 54:2593-2597
- Richardson PT, Baker DA, Ho LC (1982) The chemical composition of cucurbit vascular exudates. *J Exp Bot* 33:1239-1247
- Schaffer AA, Pharr DM, Madore MA (1996) Cucurbits, p. 729-757. In: E. Zamski, A.A. Schaffer (ed.). *Photoassimilate Distribution in Plants and Crops*. Marcel Dekker, New York
- Shi L, Peng G, Qianglong Z, Feishi L, Angela RD, Xiaolu W (2016) Development of cleaved amplified polymorphic sequence markers and a CAPS-based genetic linkage map in watermelon (*Citrullus lanatus* [Thunb.] Matsum. And Nakai) constructed using whole-genome re-sequencing data. *Breeding Sci* 66:244-259
- Yoo KS, Bang H, Lee EJ, Crosby K, Patil BS (2012) Variation of carotenoid, sugar, and ascorbic acid concentrations in watermelon genotypes and genetic analysis. *Hort Envir Biotech* 53(6):552-560