A THESIS FOR THE DEGREE OF MASTER OF SCIENCE

Genetic Mapping and QTL Analysis for Capsaicinoid Content in Pepper (*Capsicum* spp.)

고추 캡사이시노이드 함량을 조절하는 양적 형질 유전자좌 분석 및 유전자 지도 작성

FEBRUARY, 2020

DO-GYEONG LEE

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

Genetic Mapping and QTL Analysis for Capsaicinoid Content in Pepper (*Capsicum* spp.)

UNDER THE DIRECTION OF DR. BYOUNG-CHEORL KANG
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF SEOUL NATIONAL UNIVERSITY

BY
DO-GYEONG LEE

MAJOR IN HORTICULTURAL SCIENCE AND BIOTECHNOLOGY
DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

FEBRUARY, 2020

APPROVED AS A QUALIFIED DISSERTATION OF
DO-GYEONG LEE
FOR THE DEGREE OF MASTER OF SCIENCE
BY THE COMMITTEE MEMBERS

CHAIRMAN _____
Doil Choi, Ph.D.

VICE-CHAIRMAN _____
Byoung-Cheorl Kang, Ph.D.

MEMBER _____
Cecile Segonzac, Ph.D.

Genetic Mapping and QTL Analysis for Capsaicinoid Content in Pepper (*Capsicum* spp.)

DO-GYEONG LEE

DEPARTMENT OF PLANT SCIENCE
THE GRADUATE SCHOOL OF SEOUL NATIONAL UNIVERSITY

ABSRACT

The genus *Capsicum* displays various levels of pungency due to the accumulation of capsaicinoid. The biosynthesis of capsaicinoids in pepper is determined by *Pun1*, *Pun2*, and *pAMT* genes. The capsaicinoid contents are regulated by QTLs. This study was conducted to reveal additional genetic factors controlling capsaicinoid biosynthesis in *Capsaicum* spp. In the first chapter, QTL analysis was performed using an F_2 population derived from crossing the pungent *Capsicum chinense* 'Jolokia' and the non-pungent *Capsicum chinense* 'SNU11-001'. Since *C. chinense* 'SNU11-001' carries *pAMT* mutation causing no capsaicinoid accumuation, Kompetitive allele specific PCR (KASP) analysis, SNP genotyping method analysis was conducted for genotyping the *pAMT* gene for F_2 plants and the

F_2 population was grouped into the whole and the *pAMT* normal populations. In case of the whole population, all detected QTLs were clustered on chromosome 3. Some of QTL regions corresponded to the *pAMT* gene. In case of the *pAMT* normal population, QTLs were detected on chromosome 5 and chromosome 11 for dihydrocapsaicin trait. In the second chapter, genetic mapping of a novel pungeny gene in *Capsicum chacoense*. The non-pungent pepper *C. chacoense* 'PI260433' which carries the *Pun1* gene and the recessive *pun2* gene and accumulates no capsaicinoids and capsinoids. The pungent pepper *C. annuum* 'Jeju' which carries *Pun1* and *pAMT* genes showed accumulation of both capsaicinoids and capsinoids. A complementation test revealed that loss of pungency in *C. chacoense* PI260433-np may be due to a mutation at a novel pungency locus. Through QTL analysis, QTLs were detected on chromosome 3 and chromosome 9. The QTLs detected on chromosome 3 may correspond to the location of the *pAMT* locus. The QTLs detected on chromosome 9 may contain the *Pun2* locus. In conclusion, genes controlling capsaicinoid accumulation on chromosome 3, 5, and 11 were revealed and genetic mapping of *Pun2* gene was conducted.

Keywords: capsaicinoid, NGS, kompetitive allele specific PCR (KASP), pepper, quantatitive trait locus (QTL)

Student number: 2018-20331

CONTENT

CHAPTER II. Genetic mapping of a novel pungeny gene in *Capsicum chacoense*

LIST OF TABLES

CHAPTER I

CHAPTER II

.LIST OF FIGURES

GENERAL INTRODUCTION

CHAPTER I

CHAPTER II

LIST OF ABBREVIATIONS

pAMT	Putative aminotransferase
BLAST	Basic local alignment search tool
QTL	Quantitative trait locus
SNP	Single nucleotide polymorphism
GBS	Genotyping-by-sequencing
NGS	Next generation sequencing
cM	centi Morgan (the unit of genetic distance)
LOD	Logarithm of the odds
GATK	GenomeAnalysisTK
HPLC	High-performance liquid chromatography
QTL	Quantitative trait locus
CS	Capsaicin synthase
KASP	Kompetitive allele specific PCR
LG	Linkage group
CTAB	Cetyl trimethylammonium bromide
BCAT	Branched-chain amino acid transferase

GENERAL INTRODUCTION

Pepper is one of the economically important fruit consumed all over the world. It is known that pepper has unique characteristic, pungency. Capsaicin, an alkaloid derived from fatty acid and phenylpropanoid biosynthetic pathway, is responsible for pungency (Bennett and Kirby, 1968; Leete and Louden, 1968; Nelson, 1919a; Nelson 1919b). The biosynthesis of capsaicinoid is restricted to the genus Capsicum. Although more than ten different capsaicinoid structures exist (Mazourek *et al*. 2009), capsaicin and dihydrocapsaicin the most predominant, accounting for almost 90% of all capsaicinoids (Kozukue *et al*. 2005; Choi *et al*. 2006). For the majority of Capsicum species, it is known that capsaicinoids start accumulating in fruits approximately 20 days post-anthesis (DPA) (Iwai *et al*. 1979). The biosynthesis of capsaicinoid occurs in the placental epidermis cells. It is known that secreted towards the outer cell wall, and finally accumulate within structures named blisters located on the placenta surface (Suzuki *et al*. 1980; Stewart *et al*. 2007). There are several conditions required for capsaicinoid synthesis, species of capsicum fruits, developmental stage and growth conditions. The mechanisms by which the capsaicinoid amounts are regulated in chili pepper fruits are still unknown and studies are still working. Capsaicinoids are mostly used in foods, medical, cosmetic, and dietary.

The general capsaicinoid biosynthetic pathway was established at the end of the 1960s, finding that the vanillyamine moiety was synthesized from phenylalanine, and that the branched-chain fatty acid was derived from valine (Bennett and Kirby 1968; Leete and Louden 1968). The most important molecular biology approaches to understand the capsaicinoid biosynthesis pathway started with Curry *et al*. (1999) (Figure 1).

Controlling presence and absence of synthesis in capsaicin are depend on single gene, *Pun1*, *Pun2*, *CaKR1*, *pAMT* gene. The *Pun1* gene, a single genetic locus has been know to be responsible for pungency. The mutation of gene cause loss of pungency as a deletion of AT3, which encodes an acyltransferase protein belonging to the BAHD family of acyltransferase (Stewart *et al*, 2005). The *pun2* gene, known to be the ortholog of *cap*, a QTL that controls capsaicin content (Blum *et al*., 2003). A putative ketoacyl-ACP reductase (*CaKR1*) gene involved in fatty acid biosynthesis was recently found to control the pungency trait in *C. chinense* (Koeda *et al*., 2018). The *pAMT* gene is studied to catalyzes the formation of vanillylamine from vanillin in the phenylpropanoid pathway. One of the characteristic of *pAMT* gene is that peppers that harbor a non-functional *pAMT* allele synthesize capsinoids instead of the pungent capsaicinoids (Lang *et al*., 2009; Tanaka *et al*. 2010a, b).

The contents of capsaicinoids are decided by QTLs. The studies of QTLs were done by using biparental population and combination of QTL mapping and GWAS study. QTLs controlling capsaicinoid contents on chromosome 3,4, and 7

were identified by using biparental population (Blum *et al*., 2003; Ben Chaim *et al*., 2006; Yarnes *et al*., 2013). Recent studies by using QTL mapping and Genome-wide association study (GWAS) identified QTLs controlling capsaicinoid contents on chromosome 1,3,6, and 10 (Han *et al*., 2019).

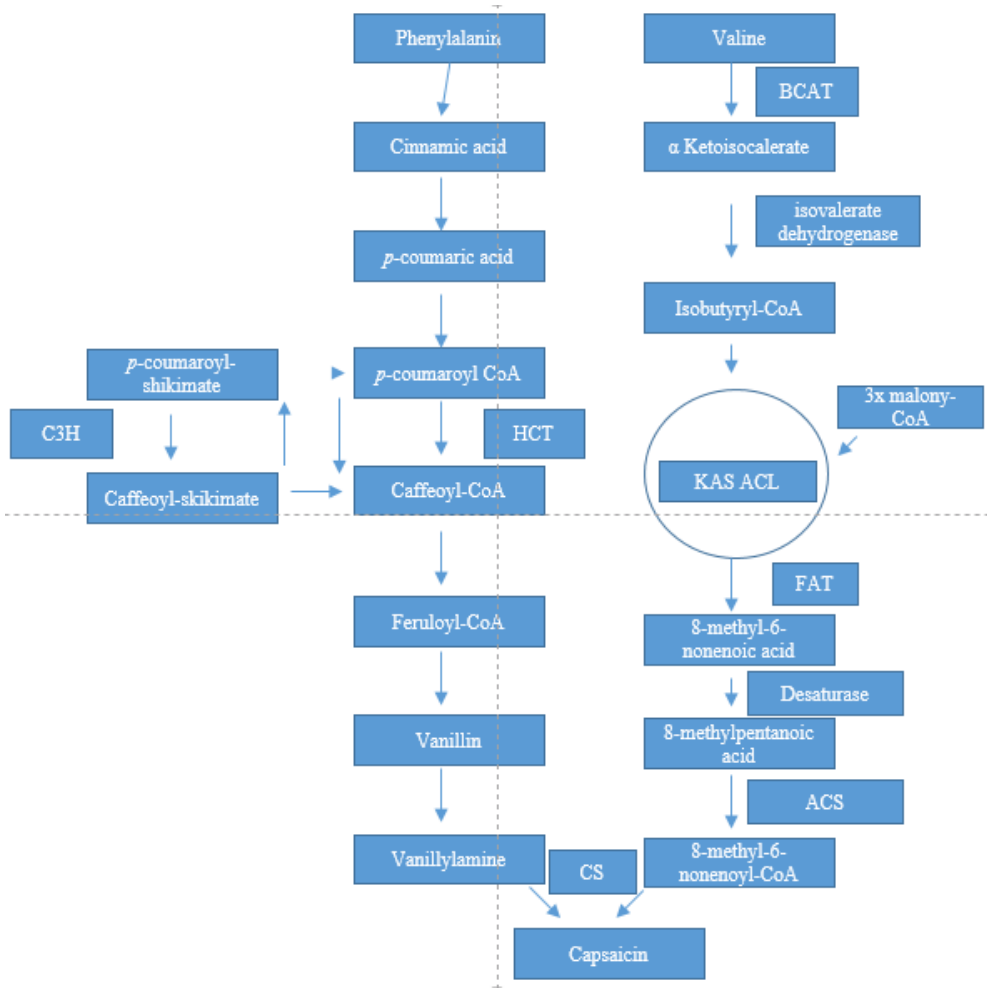


Figure 1. Capsaicin biosynthesis pathway in pepper.

REFERENCES

Ben-Chaim A, Borovsky Y, Falise M, Mazourek M, Kang BC, Paran I, Jahn M (2006) QTL analysis for capsaicinoid content in *Capsicum*. *Theor Appl Genet* 113:1481

Bennett D, Kirby G (1968) Constitution and biosynthesis of capsaicin. *J Chem Soc C* 442-446

Blum E, Mazourek M, O'connell M, Curry J, Thorup T, Liu K, Jahn M, Paran I (2003) Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum*. *Theor Appl Genet* 108:79-86

Curry J, Aluru M, Mendoza M, Nevarez J, Melendrez M, O'Connell MA (1999) Transcripts for possible capsaicinoid biosynthetic genes are differentially accumulated in pungent and non-pungent *Capsicum* spp. *Plant Sci* 148:47-57

Han K, Jang S, Lee J-H, Lee D-G, Kwon J-K, Kang B-C (2019) A MYB transcription factor is a candidate to control pungency in *Capsicum annuum*. *Theor Appl Genet* 132:1235-1246

Iwai K, Suzuki T, Fujiwake H (1979) Formation and accumulation of pungent principle of hot pepper fruits, capsaicin and its analogues, in *Capsicum annuun* var. annuun cv. karayatsubusa at different growth stages after flowering. *Agric Biol Chem* 43:2493-2498

Koeda S, Sato K, Saito H, Nagano AJ, Yasugi M, Kudoh H, Tanaka Y (2019). Mutation in the putative ketoacyl-ACP reductase *CaKR1* induces loss of pungency in *Capsicum*. *Theor Appl Genet* 132:65-80

Kozukue N, Han J-S, Kozukue E, Lee S-J, Kim J-A, Lee K-R, Levin CE, Friedman M (2005) Analysis of eight capsaicinoids in peppers and pepper-

containing foods by high-performance liquid chromatography and liquid chromatography− mass spectrometry. *J Agr Food Chem* 53:9172-9181

Lang Y, Kisaka H, Sugiyama R, Nomura K, Morita A, Watanabe T, Tanaka Y, Yazawa S, Miwa T (2009) Functional loss of *pAMT* results in biosynthesis of capsinoids, capsaicinoid analogs, in *Capsicum annuum* cv. CH-19 Sweet. *Plant J* 59:953-961

Leete E, Louden MC (1968) Biosynthesis of capsaicin and dihydrocapsaicin in *Capsicum frutescens*. *J Am Chem Soc* 90:6837-6841

Mazourek M, Pujar A, Borovsky Y, Paran I, Mueller L, Jahn MM (2009) A dynamic interface for capsaicinoid systems biology. *Plant Physiol* 150:1806-1821

Nelson E (1919a) The constitution of capsaicin, the pungent principle of *capsicum*. *J Am Chem Soc* 41:1115-1121

Nelson E (1919b) Vanillyl-acyl amides. *J Am Chem Soc* 41:2121-2130

Stewart JC, Kang BC, Liu K, Mazourek M, Moore SL, Yoo EY, Kim BD, Paran I, Jahn MM (2005) The *Pun1* gene for pungency in pepper encodes a putative acyltransferase. *Plant J* 42:675-688

Stewart JrC, Mazourek M, Stellari GM, O'Connell M, Jahn M (2007) Genetic control of pungency in *C. chinense* via the *Pun1* locus. *J Exp Bot* 58:979-991

Tanaka Y, Hosokawa M, Miwa T, Watanabe T, Yazawa S (2009) Newly mutated putative-aminotransferase in nonpungent pepper (*Capsicum annuum*) results in biosynthesis of capsinoids, capsaicinoid analogues. *J Agr Food Chem* 58:1761-1767

Tanaka Y, Hosokawa M, Miwa T, Watanabe T, Yazawa S (2010) Novel loss-of-function putative aminotransferase alleles cause biosynthesis of capsinoids, nonpungent capsaicinoid analogues, in mildly pungent chili peppers (*Capsicum chinense*). *J Agr Food Chem* 58:11762-11767

Yarnes S.C, Ashrafi H, Reyes-Chin-Wo S, Hill TA, Stoffel KM, Van DeynzeA (2012) Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific *Capsicum* RIL population. *Genome* 56:61-74

CHAPTER I

QTL analysis for capsaicinoid contents in pepper

ABSTRACT

Chili peppers characterized by pungency are one of the important vegetable crops with a wide variety of uses including food additives and pharmaceuticals. The level of pungency is dependent on genetic factors, developmental stages of the fruit, and environmental conditions of cultivation. The pungency principle of pepper is known as capsaicinoids. Capsaicinoids are synthesized by the condensation of vanillylamine with a branched-chain fatty acid in the placental tissue. Increasing capsaicinoid content is one of the important objectives in pepper breeding. To identify genetic factors that control content of capsaicinoids, 173 F_2 plants derived from a cross between the non-pungent *Capsicum chinense* 'SNU11-001' with the null *pAMT* allele and the extremely pungent *C. chinense* 'Bhut Jolokia' were used to construct a genetic linkage map by using the genotyping-by-sequencing (GBS). A total of 1,718 SNPs were identified in the GBS data sets and 615 bin markers along the 12 chromosomes. The map covers a total length of 1,150.3 cM with an average

bin marker distance of 1.92 cM. The result of QTL analysis for the whole 'SJ' population revealed that two QTLs for capsaicin. In case of capsaicin, QTLs were all detected on chromosome 3 in the whole population. In case of dihydrocapsaicin, three QTL was detected on chromosome 3 in the whole population and on chromosome 5 and chromosome 11 each in the normal *pAMT* population. In case of QTLs controlling total capsaicinoid content were all detected in whole population on chromosome 3. Furthermore, the plants with the null *pAMT* allele were removed, and QTL analysis was performed only for plants with the normal *pAMT* allele. The share QTL region on chromosome 3 in the whole population contains *pAMT* gene. The QTLs detected in the normal *pAMT* population didn't share same region of previously studies QTLs. According to validation analysis, the markers were highly associated with the dihydrocapsaicin content in the placenta of these plants.

INTRODUCTION

Chili peppers characterized by pungency is one of the agriculturally important vegetables. Peppers that contain high capsaicinoids, carotenoids and vitamins are consumed not only as vegetables and but also as industrial and pharmaceutical ingrediaents (Zhu *et al*., 2019; Guzman *et al*., 2011; Geleta and Labuschagne 2006).

Capsaicin and dihydrocapsaicin are two major compounds explaining 80-90% of the total capsaicinoids, and precursors are synthesized by two main pathways, phenylpropanoid and branched-chain fatty acid pathway (Aza-Gonzalez *et al*., 2011). Capsaicinoids are produced by the condensation of vanillylamin, derived from phenylalanine, with a branched-chain fatty acid, derived from either valine or leucine (Bennett and Kirby, 1968; Leete and Louden, 1968; Sukrasno and Yeoman, 1993; Suzuki *et al.,* 1981). Studies on capsaicinoid biosynthestic genes have been done and the single dominant gene encoded by *Pun1* was discovered to be a putative acyltransferase, the last enzyme in the capsaicinoid biosynthesis pathway.

The level of pungency is determined by by quantitative trait loci (QTLs). Quantitative trait loci (QTL) analysis is a statistical method that links phenotypic data and genotypic data to explain the genetic basis of variations of complex traits (Falconer and Mackay, 1996; Kearsey, 1998; Lynch and Walsh, 1998). The ultimate goal of QTL analysis has been to answer the question of whether phenotypic differences are primarily due to a few loci with fairly large effects among many loci

each with minute effects. It appears that a substantial proportion of the phenotypic variation in many quantitative traits can be explained with few loci of large effect, with the remainder due to numerous loci of small effects (Remington and Purugganan, 2003; Mackay, 2004; Roff, 2007). Until now, QTL mapping was used as powerful method for identifying regions of genome that co-segregate with specific trait. Despite this success, QTL mapping suffers from two fundamental limitation; only allelic diversity that segregates between the parents of the particular F_2 cross or within the RIL population can be assayed. The other is the amount of recombination that occurs during the creation of the RIL population places a limit on the mapping resolution. (Korte and Farlow, 2013). To overcome this drawbacks, Genome-wide association study (GWAS) was used.

Many QTLs studies on capsaicinoid content have been done. The QTLs controlling capsaicinoids content, five QTLs for capsaicin content, *cap3.1*, *cap4.1*, *cap4.2*, *cap7.1* and *cap7.2* and four QTLs for dihydrocapsaicin content, *dhc4.1*, *dhc4.2*, *dhc7.1* and *dhc7.2* were identified Four of the QTLs were common between capsaicin and dihydrocapsaicin content (Ben-Chaim *et al*., 2006) and a major QTL was identified as *cap*, mapped on chromosome 7 explaining 34-38% of the phenotypic variation of capsaicinoid content (Paran *et al*., 2010).

In this study, QTL analysis was done for capsaicinoid contents using a bi-parental F_2 population. An F_2 population derived from a cross between the non-pungent *Capsicum chinense* 'SNU11-001' with the null *pAMT* allele and the

extremely pungent *C. chinense* 'Bhut Jolokia' were used to construct a genetic linkage map by using the genotyping-by-sequencing (GBS). In additions to the whole population, QTL analysis was performed only for plants with the normal *pAMT* allele by removing plants with the null *pAMT*. We were able to detect five QTLs in whole population and two QTLs on the normal *pAMT* population. Identified QTLs for capsaicin, dihydrocapsaicin and capsaicinoid content using Jolokia will contribute to accelerating breeding high highly pungent pepper cultivars.

MATERIALS AND METHODS

Plant materials and mapping population construction

Two *Capsicum chinense* cultivar containing different levels of capsaicinoids, a non-pungent 'SNU11-001' and a pungent 'Bhut Jolokia' (Jolokia) were used as parental lines. To construct mapping population these parental lines were used for interspecific F_2 population, 'SNU11-001' x 'Jolokia' ('SJ'). A total of 'SNU11-001' x 'Bhut Jolokia' F_2 173 plants were grown in Anseong, Republic of Korea in 2016 (Park *et al*., 2019).

Genomic DNA extraction

Healthy young leaves from plants were sampled and DNA was extracted with cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). Homogenization of leaf tissues were performed using 3mm steel beads with the aid of TissueLyserII (Qiagen, Netherlands). Measurement of DNA concentration and purity were done using Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and diluted to a final concentration of 20 ng/μL for molecular marker genotyphing using 0.1 M TE buffer (pH 7.0)

KASP marker development and analysis

KASP markers are consisted with two allele specific forward primers and one common reverse primers. Markers were developed based on 8bp insertion on 6th exon of pAMT gene sequence and designed at LGC Genomics (KBD assay, LGC). The KASP assay primer mix consisted of common primer (5'-GTAGGTGAAGATGGTGTGGTATTACA-3'), and FAM-labeled (5'-GAAGGTGACCAAGTTCATGCTCTTTTTCTGTGGCCTCCCAA-3') and Hex-labeled (5'- GAAGGTCGGAGTCAACGGATTCTTTTTCTGTGGCGGTGTGGC -3') that enables differential amplification. Using these primers, KASP analysis was done on two 96-well plates and all reagents were votexed briefly prior to use. For PCR reaction, mixture consists of 5.0 µL 10 ng/µL genomic DNA, 5 µL 2 x KASP reaction mixture (LGC Genomics, Hoddesdon, UK), 0.06 µL 25 mM MgCl2 and 0.14 µL KASP assay primer mixture. Thermal cycling program was used for KASP reaction run: 94℃ for 15 min, followed by 10 cycles of touchdown PCR of 94℃ for 20 s, 60 °C decreasing by 0.6 °C in each cycle for 60 s, followed by 26 cycles of 94 °C for 20 s, 55 °C for 60 s, and 37 °C for 60 s, followed by plate reading at 37 °C for 1 s. To get distinct genotyping clusters, additional thermal cycles, including 3 cycles of 94 °C for 20 s and 57 °C for 1 min, were performed at the end.

HPLC analysis for capsaicinoid content measurement

For HPLC analysis, three fully matured pepper fruits were harvested from each individual. Each matured pepper fruit was separated into placenta tissue and pericarp tissue. The placenta tissue of pepper fruit was used for freeze drying (Park *et al*., 2019)Capsaicinoids were extracted following the method of Han *et al*., (2013) and HPLC was performed at the National Instrumentation Center for Environment Management (Seoul, Republic of Korea).

GBS library preparation and sequencing

GBS was constructed in the same way of previous study (Park *et al*., 2019). The number of 124 'SJ' F_2 individuals and two replications of each parents and 400-ng samples of genomic DNA were used to construct Illumina sequencing library for GBS followed by Truong *et al*., (2012). Genomic DNAs were digested with EcoRI and MseI. Single-end sequencing was performed on four lines of an Illumina Hiseq 2000 (Illumina, San Diego, CA, USA) at Macrogen Inc. (Seoul, Republic of Korea).

Analysis of SNPs

SNP analysis was done in the same way of previous study (Park *et al*., 2019). The adapter and barcodes were removed using CLC genomic workbench software version 8.0 (CLC Bio, Aarhus, Denmark). The reference genome, C. chinense scaffold version 1.2 (http://peppergeneome.snu.ac.kr), was used to align trimmed

reads, using Burrows-Wheller Aligner version 0.7.12 (Li, 2013). Genome Analysis Toolkit (GATK) UnifiedGenotyper version 3.3, with the criteria of a QUAL value larger than 30 and a minimum depth of 3, was used to futher sort and filter SNPs.

Construction of Bin map and linkage map

A bin map and linkage map was constructed in the same way of previous study (Park *et al*., 2019). A bin map was constructed using a slightly modified sliding-window approach in purpose to reduce variant calling error (Han *et al*., 2016). The linkage map was constructed using the Carthagene software (De Givery *et al*., 2005) and the criteria for linkage group were a LOD score threshold of 3.0 and maximum distance of 50 Cm. The MapChart2.3 software (Vorrips, 2002) was used to draw the resulting linkage maps.

QTL analysis

High-density genetic map from GBS data and phenotype data for capsaicin and dihydrocapsaicin content in pericarp were used for QTL analysis. Composite interval mapping (CIM) was performed by using Windows QTL cartographer v2.5 (Wang *et al*., 2012). LOD threshold was determined using 1,000 permutations with a 5% probability for each chromosomes and trait. Phenotypic variation proportion explained by each QTL was explained and estimated by using the R^2 (%) value.

RESULTS

Capsaicinoid content measurement in the biparental population

The capsaicinoid content of placenta tissues of two varieties of pepper, SNU11-001 which is non –pungent pepper, Jolokia which is pungent pepper were measured. Capsaicin and dihydrocapsaicin content were measured by HPLC analysis and total capsaicinoids were calculated by adding the amounts of capsaicin and dihydrocapsaicin. The capsaicin, dihydrocapsaicin, total capsaicinoid content in the placenta tissue of fruits from the pungent parent 'Jolokia' were 82,882 μg/g DW, 50,190 μg/g DW, 13,3072 μg/g DW, respectively. No capsaicinoids were detected from placenta tissue of non-pungent pepper 'SNU11-001' due to the null allele of *pAMT* (Figure 1). 'Jolokia' and 'SNU11-001' were used to develop a mapping population. A total of 172 F_2 plants were used to measure the content of capsaicinoid with HPLC analysis. The average capsaicinoid content of the placental tissues from biparental population was 59,058 μg/g DW (Table 1). The distribution of total capsaicinoid content in the whole population and pAMT normal population both showed wide phenotypic variation and normal distribution in their placenta capsaicinoid content (Figure 2; Figure 3; Figure 4). The individuals in both population showed bimodal distribution in capsaicinoid phenotype. It explains the

relation of one major QTL or tightly linked gene cluster in biosynthesis in capsaicinoid contents (Chee *et al*., 2001). As a result, it indicated the capsaicinoid biosynthesis in the placenta is a quantitative trait.

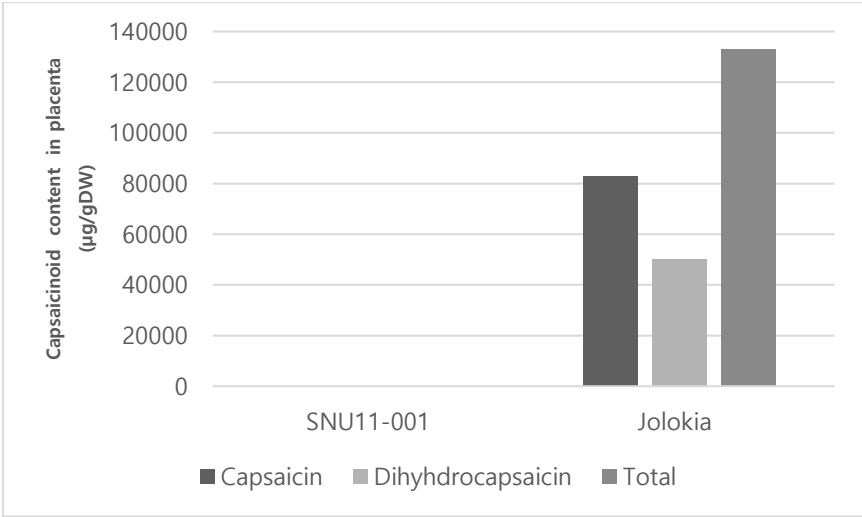


Figure 1. Capsaicinoid content (μg/g DW) in the placenta of the cultivars used in this study.

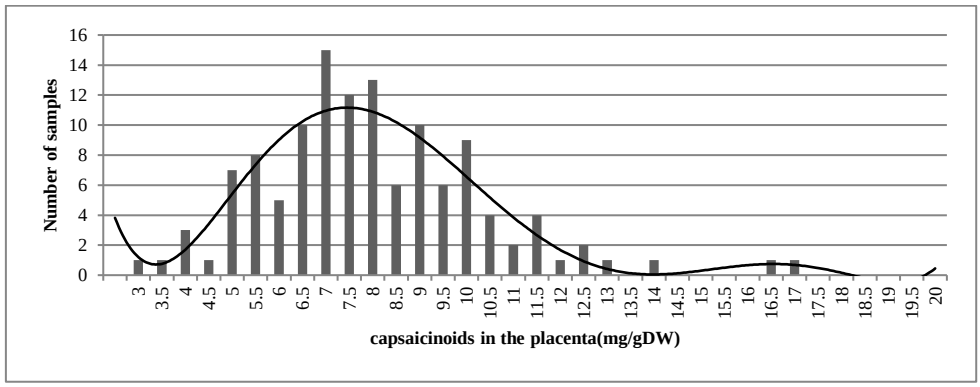


Figure 2. Frequency distribution of capsaicinoid content (mg/g DW) in placenta tissues of plants from 'SJ' F_2 population.

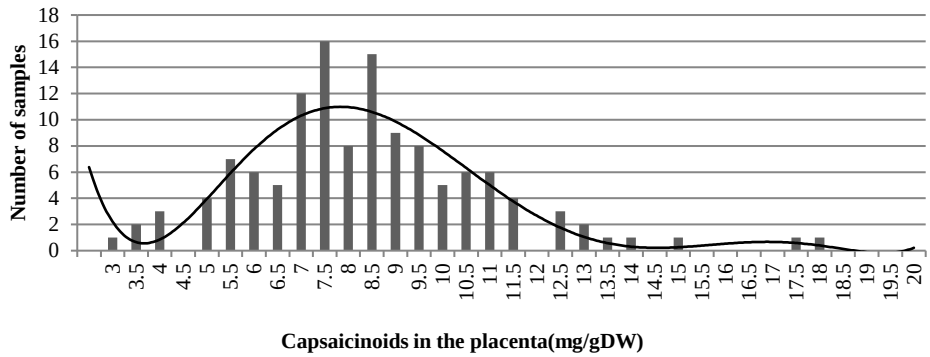


Figure 3. Frequency distribution of capsaicinoid content (mg/g DW) in placenta tissues of plants from the normal *pAMT* F_2 population.

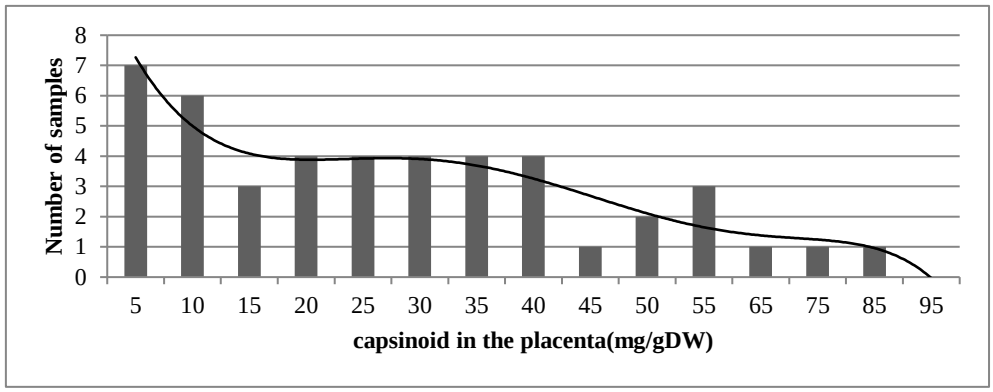


Figure 4. Frequency distribution of capsinoid content in placenta tissues of plants from the *pAMT* mutant F_2 population.

Table 1. Phenotypic characteristic of the estimated total capsaicinoid content in placenta of the parent and progeny of 'SNU11-001' x 'Jolokia' (SJ)'.

Trait	Population	Generation/ Year	Parent		Biparental mapping population	
			SNU11-001	Jolokia	Mean±S.D	Range
Capsaicinoid content in the placenta (ug/gDW)	SJ	F_2/ 2016	ND	133,072	59058±42431	0-179273

Genotype analysis

pAMT genotypes of 'SJ' population was performed to classify F_2 plants for QTL analysis. As a result, 41 samples showed dominant *pAMT* homozygous genotype (*pAMT*/*pAMT*), 85 showed heterozygous genotype (*pAMT*/*pamt*), and 47 showed recessive *pAMT* homogenous genotype (*pamt*/*pamt*) (Table 2; Figure 5). Based on the results, QTL analysis for two groups: one group is consisted with samples that have the normal *pAMT* (*pAMT*/*pAMT* and *pAMT*/*pamt*) allele and the other is the whole population including plants with the null *pAMT* allele. According to the result, the ratio was well corresponded to an expected ratio of 1:2:1. In case of the normal *pAMT* samples, capsaicinoids are expected to be synthesized normally whereas the whole population contains plants which cannot synthesize capsaicinoids due to the nonfunctional allele of *pAMT*.

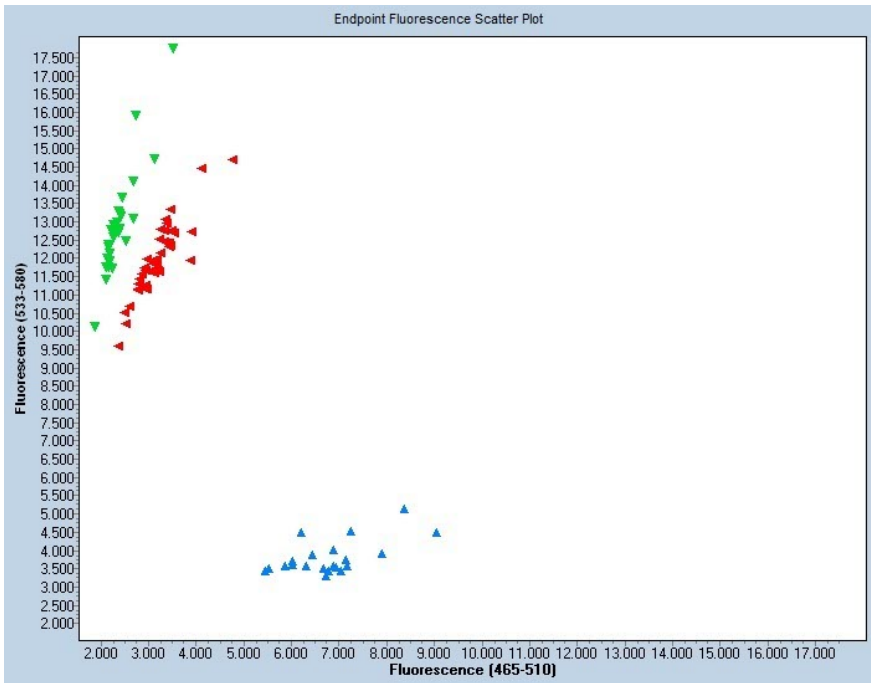


Figure 5. Genotype data from the KASP analysis. Samples clustered near the x-axis (blue dots) shows the dominant *pAMT* homozygous genotype (*pAMT*/*pAMT*). On the contrary, clustered samples near the y-axis (blue dots) shows the recessive *pAMT* homozygous genotype (*pamt*/*pamt*)

Table 2. Segregation of *pAMT* genotypes and phenotypes in an F_2 population.

No. of plants	Pungent		Non-pungent	Expected Ratio
	pAMT/*pAMT*	*pAMT*/*pamt*	*pamt*/*pamt*	
173	41	85	47	1:2:1

Construction of linkage map

As a genotyping method, GBS was used. Using 172 of the F_2 progeny and parents, 306,521,312 single-end reads were obtained and utilized to construct high-density SNP linkage map. For linkage map construction, a total of 8,297 identified SNPs were used. To determine the physical position of each SNP, the sequencing reads were aligned to the reference genome sequence (*C. chinense* chromosome version 1.2). For the accurate SNP linkage map construction, a modified sliding-window approach were used (Huang *et al*., 2009; Chen *et al*., 2014; Han *et al*., 2016). A total of 8,297 SNPs were identified in the GBS datasets and combined into 1,925 bin markers along the 12 chromosomes. As a result, linkage map was constructed with an average chromosome length of 136.78 cM, and an average of 160.4 bin markers per linkage group (Park *et al*., 2019).

QTL analysis for capsaicinoid content

QTL analysis was done with two groups, one for the 'SJ' whole population, the other for the 'SJ' normal *pAMT* population. The name of QTLs were named based on the abbreviation of the population, trait, and the chromosome number. For 'SJ' whole population, 5 QTLs were detected (Table 3; Figure 7). Of these, all QTLs were clustered on chromosome 3. QTLs for controlling capsaicin were on chromosome 3. *SJ_Plcapsaicin3.1* and *SJ_Plcapsaicin3.2* were detected and *SJ_Plcapsaicin3.1* had higher LOD value than the other QTL. QTL for controlling dihydrocapsaicin was detected on chromosome 3. QTLs for controlling total capsaicinoid contents were on chromosome 3. *SJ_Plcapsaicinoid3.1* and *SJ_Plcapsaicinoid3.2* were detected. *SJ_Plcapsaicinoid3.1* had highest LOD value and highest R^2 value which explains phenotypic variation and position of each QTL were shown (figure 8). In case of 'SJ' normal *pAMT* population, two QTLs for dihydrocapsaicin were detected with LOD thresholds 5.7 and 4.3, respectively (Table 3; Figure 7). No QTLs for capsaicin and total capsaicinoid were detected. QTLs controlling dihydrocapsaicin content were *SJnor_Pldihydrocapsaicin5* on chromosome 5 and *SJnor_Pldihydrocapsaicin11* on chromosome 11 which explains phenotypic variation 28.4% and 0.2% each. The position of each QTLs are shown in Figure 9.

In 'SJ' whole population, QTLs for capsaicin, dihydrocapsaicin and total capsaicinoid were all detected on chromosome 3 and some QTLs shared same location on 39.5 to 40.4 cM and 48.8 to 53 cM on chromosome 3. The *pAMT* gene

sequence was blasted on *C. chinense* scaffold version 1.2 to compare the regions of QTLs. The results showed that the shared region on 39.5 to 40.4 cM contains *pAMT* region. To validate the effect of the QTLs detected on the normal *pAMT* population, individual plants in the 'SJ' population were sorted into three groups according to their genotypes ('SNU11-001' or 'Jolokia' and heterozygote) at bin markers located within the QTL. The box plot was drawn with SJ5_bin98 for QTL detected on chromosome 5 and SJ11_bin85 for QTL detected on chromosome 11. All the markers were highly associated with the dihydrocapsaicin content in the placenta of these plants.

Table 3. Quantitative trait loci (QTL) for capsaicinoid content in the placenta detected in 'SJ' whole population and the normal *pAMT* population.

	Trait	QTL	Chromosome	Position	Location (cM)	LOD	R^2
Whole population	Capsaicin	*SJ_Plcapsaicin3.1*	3	40.11	39.5-40.4	43.9	0
	Capsaicin	*SJ_Plcapsaicin3.2*	3	49.81	48.8-53	29.3	0
	Dihydrocapsaicin	*SJ_Pldihydrocapsaicin3*	3	40.11	39.7-40.4	52.3	0
	Capsaicinoid	*SJ_Plcapsaicinoid3.1*	3	40.11	39.7-40.3	50.4	90.4
	Capsaicinoid	*SJ_Plcapsaicinoid3.2*	3	50.81	48.8-53.4	33.5	78
Normal *pAMT* population	Dihydrocapsaicin	*SJnor_Pldihydrocapsaicin5*	5	87.31	86.5-89.8	5.7	28.4
	Dihydrocapsaicin	*SJnor_Pldihydrocapsaicin11*	11	114.81	108.5-123.3	4.3	0.2

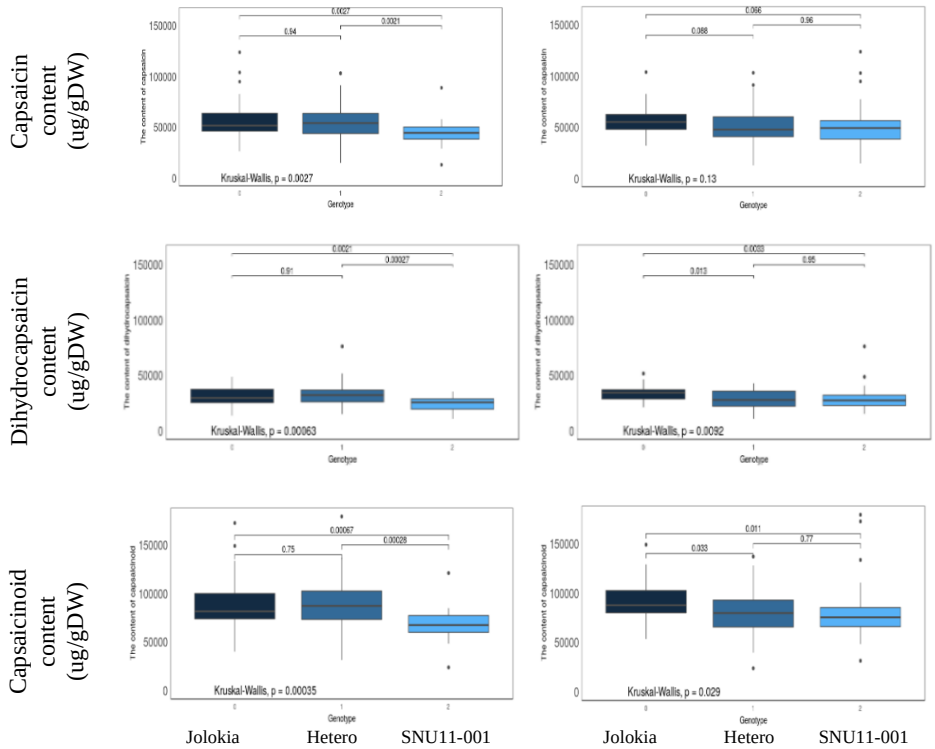


Figure 6. Comparison of QTLs and SNPs associated with dihydrocapsaicinoid content and box plots of capsaicin, dihydrocapsaicin and capsaicinoid content regulated by two markers representing the two QTLs. Bin markers used for box plot were SJ5_bin98 (a),, and SJ11_bin85 (b) 'SJ' *pAMT* normal population, respectively.

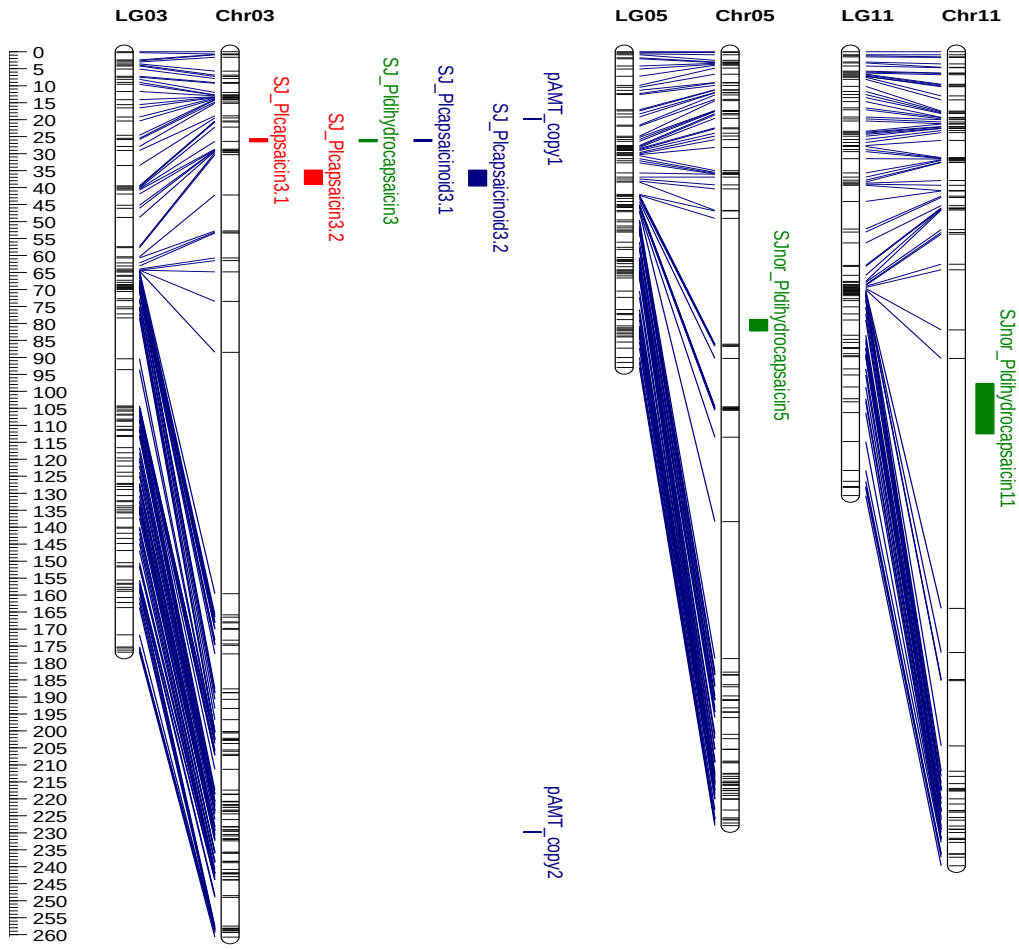


Figure 7. Result of QTL analysis in both population. Physical and linkage map containing the QTL position on chromosome 3, chromosome 5 and chromosome 11 associated with the capsaicin, dihydrocapsaicin and total capsaicinoid content in the placenta based on bin marker in 'SJ'

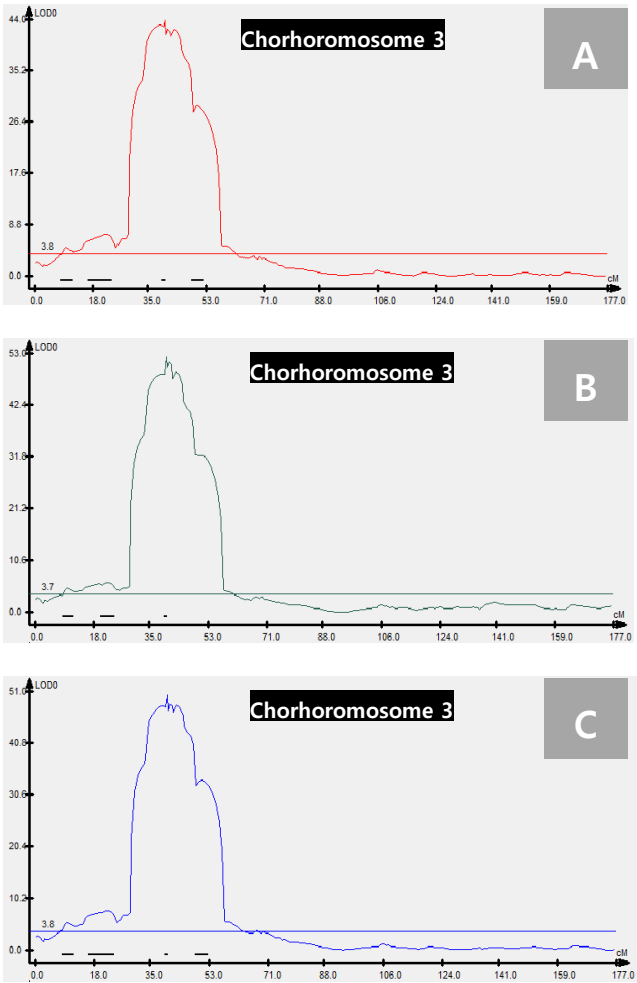


Figure 8. Position of QTLs for capsaicin, dihydrocapsaicin and total capsaicinoid content in placenta tissues for 'SJ' whole population.

Four QTLs were identified for capsaicin in the placenta on chromosome 3. This QTL is located nearby SJ3_bin27 marker. Three QTLs for dihydrocapsaicin conents in the placenta were detected on chromosome 3 located nearby SJ5_bin19 marker. Four QTLs for total capsaicinoid content were detected on chromosoe 3 located nearby SJ_ marker.

A: QTLs for capsaicin, B: QTLs for dihydrocapsiacin, C: QTLs for total capsaicinoid

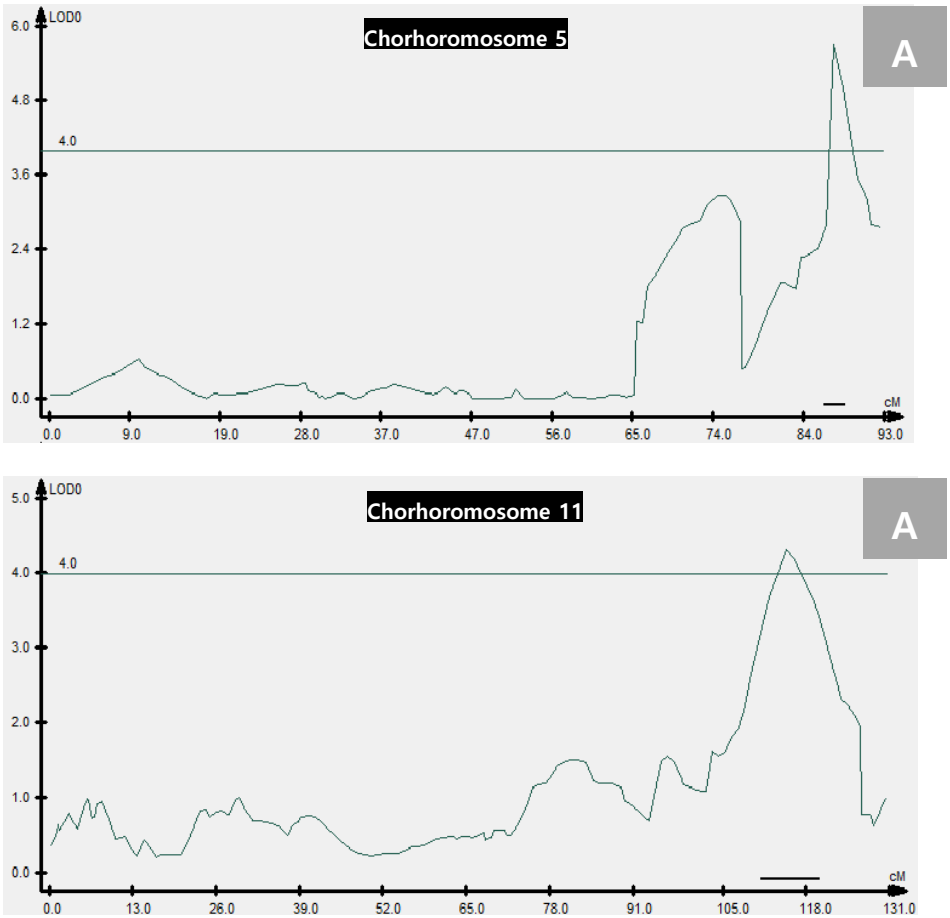


Figure 9. Position of QTLs for dihydrocapsaicin content in placenta tissues for 'SJ' *pAMT* normal population. Two QTLs were identified for dihydrocapsaicin in the placenta on chromosome 5 located nearby SJ5_bin98 marker and chromosome 11 located nearby SJ11_bin58 marker each.

A: QTL for dihydrocapsaicin content

DISCUSSION

Pungency is one of the attractive attributes in peppers and affected by genetic and environmental factors (Harvell and Bosland, 1997). Capsaicin, the alkaloid responsible for the spicy flavor, used for medical, culinary and military purpose, resulted in a desire to breed high content capsaicinoid pepper. The responsibility of pungency in pepper was known as capsaicinoid compound and capsaicinoid biosynthetic pathway was published with several candidate genes. Presence of pungency is determinded by single gene, *Pun1*, *Pun2*, and *pAMT*. However, the biosynthetic pathway of capsaicinoid remains still unclear and more studies about candidate genes are needed.

The capsaicinoid content in pepper depends on QTLs and many studies were done. Previous studies identified QTLs affecting capsaicinoid content on chromosome 1, 3, 4, 6, 7 and 10 (Blum *et al*., 2003; Ben Claim *et al*., 2006; Yarnes *et al*., 2013; Han *et al*., 2018). There are many candidate genes controlling production of capsaicinoid. Using combination of QTL mapping and GWAS study, five candidate genes for controlling capsaicinoid contents in pepper: pAMT, C4H, 4CL and CSE from the phenylpropanoid pathway, and FatA, from the fatty acid pathway were identified (Han *et al*., 2018). Of those, *pAMT* gene, encoding an aminotransferase which produces vanillylamine from vanillin, is precursor of capsaicinoids (Curry *et al*., 1999).

There are benefits of using F_2 population for QTL analysis in that a population can be easily constructed in short time. However, the limitation for using an F_2 population

is no replication of phenotype evaluation because F_2 can be grown only once. Even though some drawback of using F_2 population as QTL analysis, still used as plant material for studies.

Due to the fast development of sequencing technologies, genotyping became more and more affordable and accessible to the scientific community. The GBS strategy is by far the most widespread technique for high-throughput genotyping, allowing simultaneous variant calling and genotyping for thousands of SNPs without the need for a reference genome. Although many QTL mapping studies using less dense linkage maps have confirmed their effectiveness, increasing the number of markers allows full exploitation of the recombination events in the population, improving the resolution of the QTLs. Therefore we used GBS analysis for construction of high-density genetic map. The selection of restriction enzymes is one of the key factors for determining the quality of GBS analysis. We used *Eco*RI and *Mse*I to construct the 'SJ' linkage map. The sliding window approach was used to impute missing SNP data, where adjacent SNPs with the same genotype are combined into a bin marker (Poland and Rife, 2012; Chen *et al*., 2014) resulting in 12 linkage groups containing 1,925 bin markers converted from 8,297 SNPs. The high density genetic was constructed and the result of comparison with *C. chinenese* scaffold version 1.2 reference genome showed collinear form.

QTL analysis were done with two groups according to the *pAMT* genotype. One for the 'SJ' whole population, the other for the normal *pAMT* population. In case of the whole population, the QTLs location on chromosome 3 were compared with previously reported QTLs. Among the shared QTL locations, *pAMT* gene was located on 39.5 to 40.4

cM. In case of the normal *pAMT* population, no QTLs sharing with the previously reported QTLs were detected.

With the development of sequencing technique and advent of NGS, we were able to better define the genomic region controlling pungency, identifying new SNP markers and genomic regions not previously identified. Comparison with other QTLs studies before, we have identified several new QTLs controlling content of dihydrocapsaicin. This result may help to identify genes controlling pungency in pepper. Also, development of highly pungent pepper varieties can be possible.

REFERENCES

Aza-González C, Núñez-Palenius HG, Ochoa-Alejo N (2011) Molecular biology of capsaicinoid biosynthesis in chili pepper (*Capsicum* spp.). *Plant Cell Rep* 30:695-706

Bennett DJ, Kirby GW (1968) Constitution and biosynthesis of capsaicin. *J Chem Soc C* 442-446

Ben-Chaim A, Borovsky Y, Falise M, Mazourek M, Kang BC, Paran I, Jahn M (2006) QTL analysis for capsaicinoid content in *Capsicum*. *Theor Appl Genet* 113:1481

Chee-Sanford JC, Aminov RI, Krapac IJ, Garrigues-Jeanjean N, Mackie RI (2001) Occurrence and diversity of tetracycline resistance genes in lagoons and groundwater underlying two swine production facilities. *Appl Environ Microbiol* 67:1494-1502

Chen Z, Wang B, Dong X, Liu H, Ren L, Chen J, Hauck A, Song W, Lai J (2014) An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F_2 maize population. *BMC Genomics* 15:433

Chen Z, Wang B, Dong X, Liu H, Ren L, Chen J, Hauck A, Song W, Lai J (2014) An ultra-highdensity bin-map for rapid QTL mapping for tassel and ear architecture in a large F_2 maize population. *BMC Genomics* 15:433

Curry J, Aluru M, Mendoza M, Nevarez J, Melendrez M, O'Connell MA (1999) Transcripts for possible capsaicinoid biosynthetic genes are differentially accumulated in pungent and non-pungent *Capsicum* spp. *Plant Sci* 148:47-57

De Givry S, Bouchez M, Chabrier P, Milan D, Schiex T (2004) CARHTA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* 21:1703-1704

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue Phytochemical Bulletin 19:11-15

Falconer DS (1960) Introduction to quantitative genetics. Introduction to quantitative genetics

Geleta LF, Labuschagne MT (2006) Combining ability and heritability for vitamin C and total soluble solids in pepper (*Capsicum annuum* L.). *J Sci Food Agric* 86:1317-1320

Guzman I, Bosland PW, O'Connell MA (2011) Heat, color, and flavor compounds in *Capsicum* fruit. In: Gang D (Ed) The Biological Activity of Phytochemicals pp. 109-126

Han K, Jeong HJ, Sung J, Keum YS, Cho MC, Kim JH, Kwon JK, Kang BC (2013) Biosynthesis of capsinoid is controlled by the *Pun1* locus in pepper. *Mol Breeding* 31:537-548.

Harvell KP, Bosland PW (1997). The environment produces a significant effect on pungency of chiles. *HortScience* 32:1292-1292

Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang T, Dong G, Sang T, Han B (2009) High-throughput genotyping by whole-genome resequencing. *Genome Res* 19:1068-1076

Kearsey MJ (1998) The principles of QTL analysis (a minimal mathematics approach). *J Exp Bot* 49:1619-1623

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9:29

Leete E, Louden MC (1968) Biosynthesis of capsaicin and dihydrocapsaicin in *Capsicum frutescens*. *J Am Chem Soc* 90:6837-6841

Li J, Pandeya D, Jo YD, Liu WY, Kang BC (2013) Reduced activity of ATP synthase in

mitochondria causes cytoplasmic male sterility in chili pepper. *Planta* 237:1097-1109

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits Vol. 1, pp. 535-557

Mackay IM (2004) Real-time PCR in the microbiology laboratory. Clin Microbiol Infect 10:190-212

Paran I, Akler T, Jones G (2010) QTLs for capsaicinoids content in *Capsicum. Advances in genetics and breeding of Capsicum and eggplant.* Universidad Politécnica de Valencia, Valencia, pp. 273-278

Poland JA, Rife TW (2012). Genotyping-by-sequencing for plant breeding and genetics. *The Plant Genome* 5:92-102.

Remington DL, Purugganan MD (2003) Candidate genes, quantitative trait loci, and functional trait evolution in plants. Int J Plant Sci 164:S7-S20

Roff DA (2007) A centennial celebration for quantitative genetics. Evolution Int J 61:1017-1032

Sukrasno N, Yeoman MM (1993) Phenylpropanoid metabolism during growth and development of *Capsicum frutescens* fruits. *Phytochemistry* 32:839-844

Suzuki T, Iwai K (1984) Constituents of red pepper species: chemistry, biochemistry, pharmacology, and food science of the pungent principle of *Capsicum* species. In *The alkaloids: Chemistry and pharmacology* Vol. 23, pp. 227-299

Truong HT, Ramos AM, Yalcin F, de Ruiter M, van der Poel HJ, Huvenaars KH, Hogers RC, van Enckevort LJ, Janssen A, van Orsouw NJ, van Eijk MJ (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. *PloS One* 7:e37565

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77-78

Wang S, Basten CJ, Zeng Z.B (2012) Windows QTL Cartographer 2.5. Raleigh, NC: Department of Statistics, North Carolina State University

Yarnes SC, Ashrafi H, Reyes-Chin-Wo S, Hill TA, Stoffel KM, Van DeynzeA (2012) Identification of QTLs for capsaicinoids, fruit quality, and plant architecture-related traits in an interspecific *Capsicum* RIL population. *Genome* 56:61-74

Zhu Z, Sun B, Cai W, Zhou X, Mao Y, Chen C, Wei J, Cao B, Chen C, Chen G, Lei J (2019) Natural ariations in the MYB transcription factor MYB31 determine the evolution of extremely pungent peppers. *New Phytol* 223:922-938

CHAPTER II

Genetic mapping of a novel pungeny gene in *Capsicum chacoense*

ABSTRACT

The production of chili pepper worldwide in 2014 was 3.91 million hectares, with 1.78 t/ha average productivity of dry chili pepper pods (FAO, 2013). *Pun1*, *Pun2* and, *pAMT* are major genes responsible for capsaicinoid contents variation in pepper. Nonetheless, numerous non-pungent pepper cannot be explained by variations of these three genes. To discover the genetic factor controlling capsaicinoid contents of pepper *C. annuum* 'Jeju' and *C. chacoense* 'PI260433-np' F_2 population was used. Allelism test which indicates possibility of novel gene controlling capsaicinoid contents in pepper was conducted by previous researcher Koen Han. Genotyping was conducted with GBS analysis. The *C. annuum* 'Dempsey' reference genome (unpublished) was used to align the reads and a total of 10,136 SNPs were obtained from GBS analysis. A high-density bin

map of 'PJ' F_2 population was constructed with 1,181 bin markers with an average genetic distance of 3.48 cM. The total genetic map length was estimated to be 4,112 cM. QTL analysis with high-density genetic map was performed and QTLs were detected on chromosome 3 and chromosome 9. Through comparison of physical location between physical location of *pAMT* gene location and QTLs detected on chromosome 3, they shared same location each other. Complementation test was conducted to figure out whether the mutation in two strains are in different genes. The loss of pungency in *C. chacoense* PI260433-np was result of mutation at a novel pungency locus. The possibility that QTLs detected on chromosome 9 likely to contain *Pun2* locus.

INTRODUCTION

Capsicum species characterized by pungency is one of the important vegetable crops. Pungency of pepper is determined by capsaicinoids which are mainly composed of capsaicin and dihydrocapsaicin. Capsaicin and dihydrocapsaicin differ only by the saturation of the acyl moiety. The understanding of the capsaicinoid biosynthetic pathway has been important studies as valuable target for vegetable crop improvement through plant breeding.

Capsaicinoid biosynthesis is regulated by single genes and the content is regulated by QTLs. For single genes, *Pun1*, *Pun2* and *pAMT* were studied. *Pun1* which encodes a putative acyltransferase, the last enzyme in the capsaicinoid biosynthesis pathway, was discovered to control the presence and absence of pungency. The mutant allele was identified with a large 2.5 kb deletion spanning 1.8 kb of the putative promoter and 0.7 kb of truncated 1st exon is following. The conservation of the deletion between Bell and Jalapeno peppers indicated that the deletion was widespread throughout *C. annuum*. Furthermore, four additional mutantation alleles were discovered later in other *Capsicum* spp. The second *Pun1* allele in *Capsicum chinense* has a 4 bp deletion in the 1st exon region that creates an early stop codon. The third *Pun1* allele in *Capsicum frutescens* has a large deletion in the 2nd exon region that creates truncation in the 2nd exon. The fourth *Pun1* allele has one insertion in the 2nd exon that causes a frameshift mutation.

pAMT is known to catalyze the formation of vanillyamin from vanillin in the metabolic pathway of capsaicinoids and encodes aminotransferase which produces vanillylamine, a precursor of capsaicinoid (Curryet *et al*., 1999; Blumet *et al*., 2003; Abraham-Jua´rezet *et al*., 2008). A series of studies showed that loss-of-function of *pAMT* causes low content of capsaicinoids and accumulates capsinoids (Lang *et al*., 2009; Tanaka *et al*., 20010a). Capsinoids have similar structure with capsaicinoids but differs in presence of ester group instead of an amino group (Yazawa *et al*., 1989; Kobata *et al*., 1998; Lang *et al*., 2009). To date, ten non-functional *pAMT* alleles are identified and use as plant material for breeding non-pungent pepper (Lang *et al*., 2009; Tanaka *et al*., 2010a,b, 2015, 2017; Koeda *et al*., 2014; Park *et al*., 2015; Tsurumaki *et al*., 2019).

In the undomesticated *Capsicum chacoense*, a novel locus regulating presence of pungency was discovered named *Pun2* (Stellari *et al*., 2010). The loss of pungency in *Capsicum chacoense* was due to mutation at a novel locus. By bulked segregation analysis, Hpms1-172 marker was identified and COSⅡ markers known to map near Hpms1-172 were screened. This novel gene was mapped on chromosome 7 (Ben-Claim *et al*., 2006), however, the identity of *Pun3* is unknown. *Pun3* encodes the transcription factor CaMYB31, which regulates the expression of *Pun1* and other capsaicinoid biosynthetic genes, especially fatty acid-related genes (Han *et al*., 2018).

In this study, in order to find the genetic factor controlling contents of pungency in pepper, both quantitative and qualitative view were considered. QTL analysis was

performed to discover genetic factor controlling capsaicinoid contents in 'PJ' F_2 population. Discovered QTLs on chromosome 3 indicates shared region of *pAMT* gene. In case of new-founded QTLs on chromosome 9 indicates possibility of inclusion of *Pun2* gene.

MATERIALS AND METHODS

Plant materials and mapping population construction

Non-pungent *Capsicum chacoense* 'PI260433-np' and *Capsicum annuum* 'Jeju' were used as parental lines. Three developmental stages, mature green stage, breaker stage and mature red stage of fruits were harvested and mature green stage samples were used for HPLC analysis (Figure 1). These parental lines were used to construct a F_2 population, 'PI260433-np' x 'Jeju' ('PJ'). A total of 171 plants were grown in the field of Seoul National University farm (Suwon, Republic of Korea) in 2019.

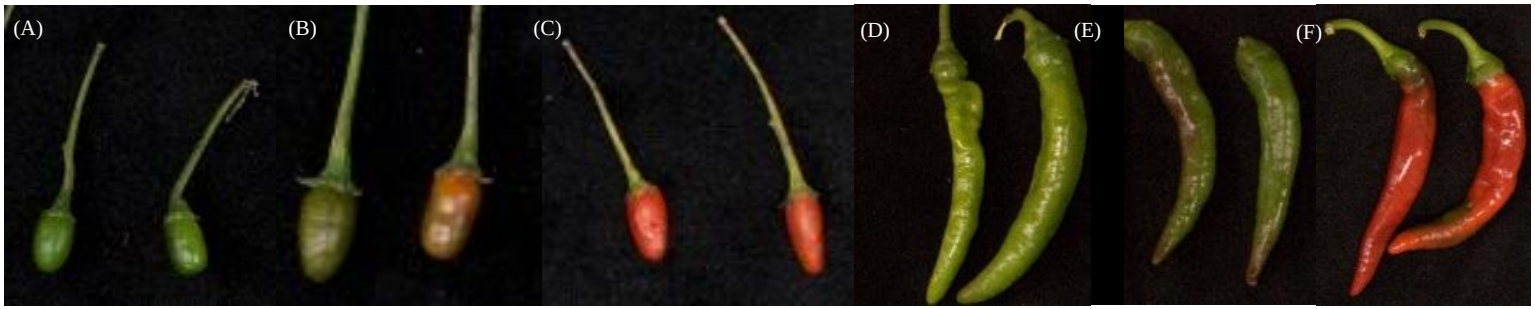


Figure 1. Fruits of 'PI260433-np' and 'Jeju' used as parental lines in this study.

Mature green stage 'PI260433-np' (A), breaker stage 'PI260433-np' (B), mature red stage 'PI260433-np' (C), mature green stage 'Jeju' (D), breaker stage 'Jeju' (E) and mature red stage 'Jeju'

Allelism test

An allelism test was carried out by crossing C. *chacoense* 'PI260433-np' with nonpungent cultivars C. *annuum* 'YCM334' and C. *annuum* 'ECW30R'. If two recessive genes are allelic, they will fail to complement each other in the F_1 hybrids

Phenotyping with Gibbs's screening and HPLC analysis

Mature green fruits from each individual were sampled. Placenta tissues were used for Gibb's analysis (Jeong *et al*., 2012). Placenta tissues were placed on filter paper and the same volume of 2,6-dichloroquinon-4-chloroimide(Gibb's reagent; Sigma-Aldrich, Saint Louis, Missouri, USA) were sprayed. The filter paper with sprayed spots was steamed for 30 s with ammonia gas. If the color changes to blue, the samples were determined to be pungent. For accurate capsaicinoid and capsinoid content analysis, HPLC analysis was done. The placenta tissue of pepper fruit was used for freeze drying. capsaicinoids and capsinoids were extracted following the method of Han *et al*.,(2013) and HPLC was performed at the National Instrumentation Center for Environment Management (Seoul, Republic of Korea). In short, only one biological replications were done by freeze-drying placenta tissue and then ground using a hand blender (HR2860; Koninklijke Philips, Amstrerdam, the Netherlands). 0.1g of pepper powder was placed in 2-ml microcentrifuge tube and dried using a centrifugal speed vacuum concentrator SVQ-70 (Operon, Gimpo, Republic of Korea). To dissolve the pellet, 1 ml methyl alcohol was mixed and voltexed. The mixture was filtered with 0.2-mm syring filter (PN4450; Pall

Corporation, Port Washington, NY, USA). The filtered extracts were transferred to a high-performance liquid chromatography (HPLC) vial (5182-0715; Agilent Technologies, Santa Clara, CA, USA).

Genomic DNA extraction

Leaves with healthy and young plants were sampled. DNA was extracted using cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1987). Leaf tissue homogenization was performed using two 3 mm steel beads with the aid of TissueLyserII (Qiagen, Netherlands). Purity and DNA concentration were measured with Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). DNA was diluted to a final concentration of 80 ng/μL by 0.1 M TE buffer (pH 7.0) for genotyping-by-sequencing(GBS).

GBS library preparation and sequencing

Genomic DNA of F_2 individuals and three replications of each parent were used to construct Illumina sequencing library for GBS followed by Truong *et al*., (2012). Genomic DNAs were digested with *Eco*RI and *Mse*I. After digestion, *Mse*I adapters and *EcoR*I adapters with different barcodes were used to ligate DNA fragments. Each sample with the same quantity of adapter-ligated DNA fragments was pooled for sequencing. Single-end sequencing was performed on one lines of an Illumina Hiseq 2000 (Illumina, San Diego, CA, USA) at Macrogen Inc. (Seoul, Republic of Korea).

SNP analysis

The adapter and barcodes were removed using CLC genomic workbench software version 8.0 (CLC Bio, Aarhus, Denmark). The reference genome, *Capsicum annuum* 'Dempsey'(not published) was used to align trimmed reads, using Burrows-Wheller Aligner version 0.7.12 (Li, 2013). To convert the alignment files into BAM files, Sequence Alignment/Map (SAM) tools version 1.1 was used. Then, Picard Tools version 1.119 was used to manipulate the SAM files and performed duplicate marking and sorting. Genome Analysis Toolkit (GATK) UnifiedGenotyper version 3.3, with the criteria of a QUAL value larger than 30 and a minimum depth of 3, was used to further sort and filter SNPs.

Bin map construction

The construction of bin map was done by using modified sliding-window approach to reduce variant calling error and calculation of recombination breakpoints (Han *et al*., 2016). SNPs with non-polymorphic and missing data were removed. The ratio of SNPs with both parental genotype was calculated for each window, definded as 25 linked SNPs, and the overall genotype of each window was decided. The SNPs with over 0.7 ratio were defined was paternal and maternal genotype. The SNPs ratio between 0.3 and 0.7 were defined as heterozygous genotype. Construction of linkage maps were conducted with Carthagene softwere (De Givry *et al*., 2005). The criteria for construction of linkage map were a LOD score threshold of 4.0 and a maximum distance of 50 cM. The calculation

of distance between bin markers was done with Kosambi mapping function. The final linkage map was drawn using MapChart2.3 software (Voorips, 2002). The *C. annuum* 'Dempsey' reference genome was used to construct bin map. The comparison with physical location of bin markers was done with MapChart2.3 software.

Genetic mapping of *Pun2*

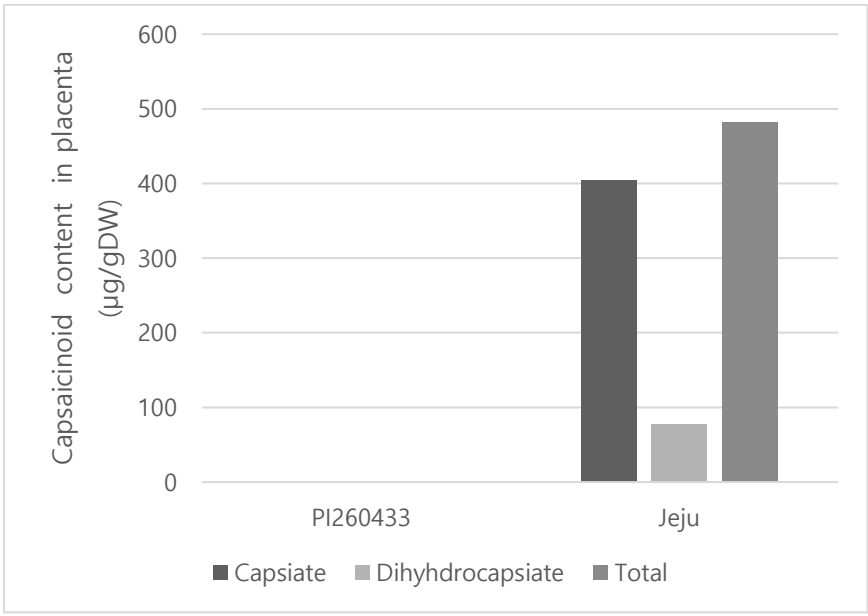
High-density genetic map from GBS data and phenotype data for capsaicin and dihydrocapsaicin content in pericarp were used for QTL analysis. Composite interval mapping (CIM) was performed by using Windows QTL cartographer v2.5 (Wang *et al*., 2012). LOD threshold was determined using 1,000 permutations with a 5% probability for each chromosomes and trait. Phenotypic variation proportion explained by each QTL was explained and estimated by using the R^2 (%) value.

RESULT

Phenotype analysis

Two types of phenotyping methods, Gibb's and HPLC analyses were used. Non-pungent pepper 'PI260433-np' didn't show blue color from Gibb's analysis demonstrating no capsaicinoid and capsinoid accumulation in 'PI260433-np'. By contrast, pungent pepper, 'Jeju' showed blue color in Gibb's reagent test. In HPLC analysis, capsaicinoid content was 5417.5 μg/g dry weight of placenta tissue and capsinoids content was 482.2 μg/g DW. 'Jeju' contained small amount of capsinoids (Figure 2).

For evaluation for pungency in 'PJ' F_2 population, Gibb's analysis was first conducted at least two times, using mature green stage fruits. If the color of samples turned into blue, it was determined to be pungent, while samples with no color changes were non-pungent. Out of 172 F_2 samples, 114 samples were pungent and 57 samples were non-pungent. The segregation ratio of non-pungent vs pungent peppers in the F_2 population were 2:1. For more accurate phenotype analysis, HPLC analysis was conducted to confirm Gibb's analysis. The mean capsaicinoid content of the placenta tissue of F_2 population was 2067 ug/g DW and the mean capsinoid content was 583ug/g DW, respectively (Table 1). The distribution of both capsaicinoid and capsinoid content showed positive skew. The positive skew means the mean is greater than the median (Figure 3).



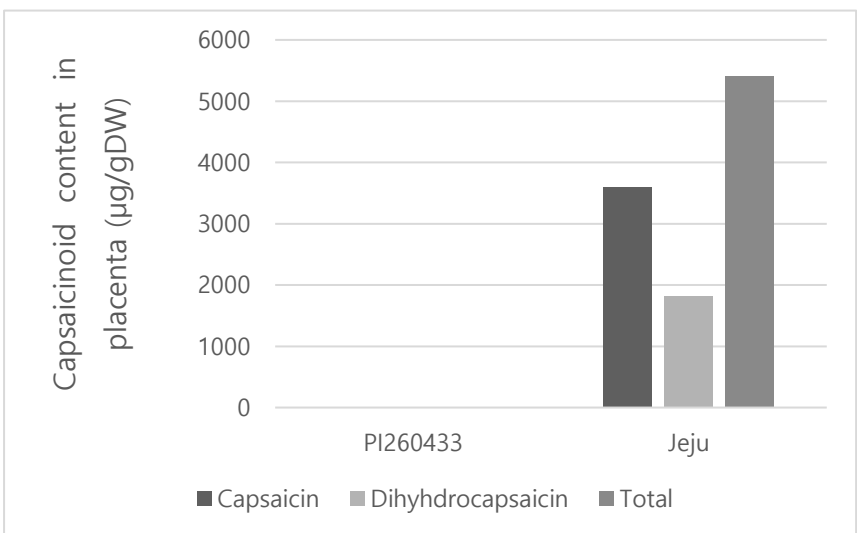


Figure 2. Capsaicinoid and capsinoid content in the placenta tissues of the parental lined used in this study.

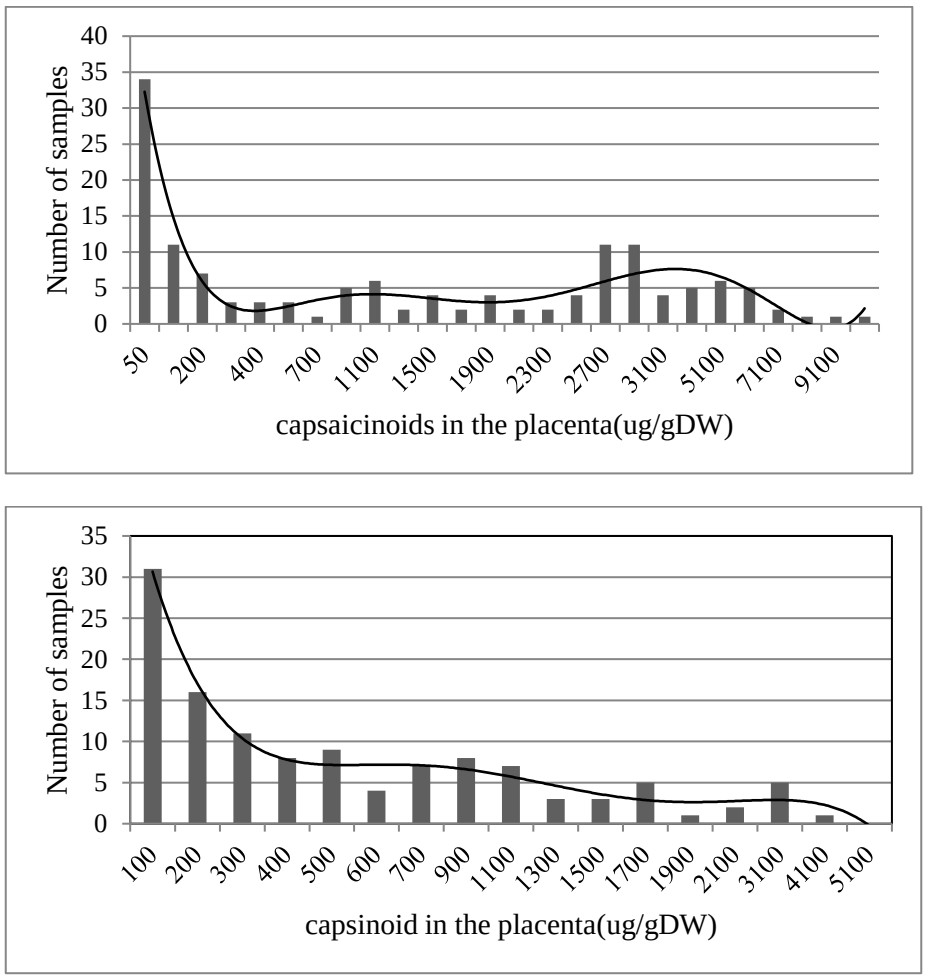


Figure 3. Frequency distribution of capsaicinoid and capsinoid content (ug /g DW) in the placenta tissues of the lines used in this study.

Table 1. The estimated total capsaicinoid and capsinoid content in placental tissues of the parents and progeny of PJ population.

Line	Number of line		Pungent : non-pungent	Capsaicinoid content (ug/g dry weight of placenta)			Capsiniod content (ug/g dry weight of placenta)		
	Pungent	Non-pungent		Capsaicin	Dihydrocapsaicin	Total capsaicin	Capsiate	Dihydrocapsiate	Capsinoid
PI260433-np	0	1	0 : 1	0	0	0	0	0	0
Jeju	1	0	1 : 0	3594	1823	5417	404	77	482
F_2	114	57	2 : 1	876±1195	639±937	2067±2067	388±467	194±267	583±693

Allelism test

Allelism test was conducted to confirm whether *Pun2* is a novel gene (Koen Han, unpublished). F_1 plants from a cross between 'PI260433-np' and 'ECW30R' with the non-functional *pun1* allele accumulated capsaicinoids with 4,135 ug/gDW (Table 2). F_1 plants from a cross between 'PI260433-np' and 'YCM334' were also pungent with 4,314 ug/gDW of capsaicinoid. These results demonstrated that the *pun2* could complement both *pun1* and *pun3* alleles. F_1 plants from the cross between 'Jeju' and 'PI260433-np' were all pungent with 804.9 ug/g dry weight.

Genotyping-by-sequencing and bin map construction

The genotyping of the 'PJ' F_2 population was done with GBS. *EcoR1/Mse1*-digested DNA were used to construct GBS libraries. The average number of reads per sample was 2,803,611 and the reads were aligned to *C. annuum* 'Dempsey' reference genome (unpublished). A total of 10,136 SNPs were obtained from GBS analysis (Table 2). The density of SNPs showed difference among the chromosomes. The SNPs were distributed across chromosomes in each chromosome (Figure 4).

To construct a linkage map, the modified sliding window approach was used to correct missing data and genotyping error (Huang *et al*., 2009; Chen *et al*., 2014; Han *et al*., 2016). To determine recombination breakpoints, 25 consecutive SNPs were considered as one sliding window. A single recombination bins were defined by adjacent SNPs with the same genotype (Figure 6). A high-density bin map of 'PJ' F_2 population was

constructed (Figure 5). CarthaGene software was used to construct a genetic linkage map. The map consisted of 1,604 bins with an average genetic distance of 3.48 cM. The total genetic map length was estimated to be 4,112 cM. Among the 12 linkage groups, the genetic distance of chromosome 1 was longest with 333 cM and chromosome 8 and 11 were the shortest with 174 cM (Table 3). The bin map developed using 1604 bin markers was align to *C. annuum* 'Dempsesy' reference genome to compare the physical positon of each bin (Figure 5).

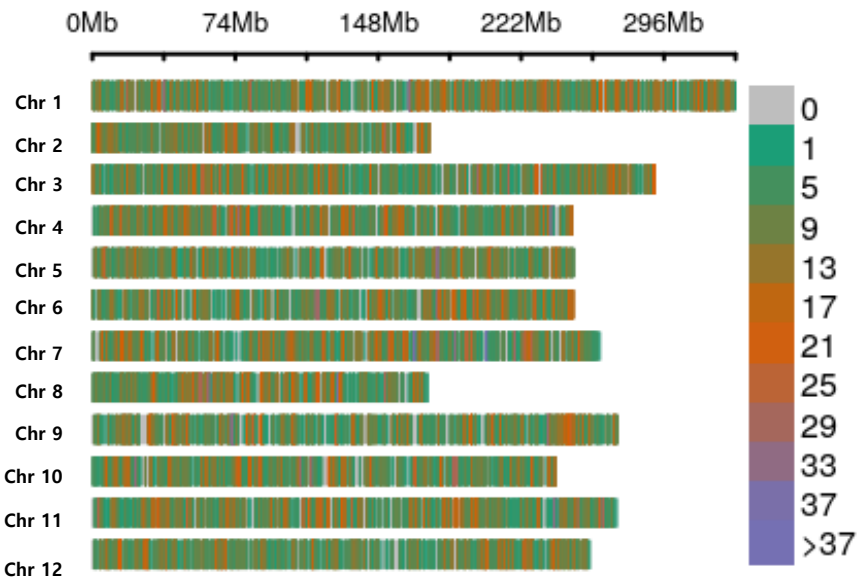


Figure 4. SNP densities of the 'PJ' F_2 population.

Table 2. Number of sequencing reads from GBS and SNPs from QTL mapping.

	PJ F_2 population
Number of accession	176
Genotyping method	GBS
Average number of reads per sample	2,803,611
Total number of SNPs	101,136
Average distance between SNPs (bp)	29,906

Table 3. Summary of the sequencing data and the linkage map constructed from 'PJ' F_2 population.

Chr.	Number of SNPs	Number of bins	Physical length of bin (Mb)		Genetic distance of bin (cM)	
			Mean	Total	Mean	Total
1	1115	166	161.3	319.6	176.1	333
2	639	109	115.8	190.9	83	175
3	965	168	157.3	351.3	136.9	291
4	839	149	242.4	431	132.7	249
5	703	134	147.1	293	125.9	249
6	803	146	182.6	378	141.4	249
7	989	154	287.9	467.7	140.2	263
8	637	104	226.3	351.3	85.8	174
9	847	127	102.7	194.5	133.7	271
10	844	129	191.7	306.9	120	240
11	969	104	226.3	351.3	85.8	174
12	786	114	57.9	123.9	131.9	258
Total	10136	1604	2099.3	3759.4	1493.4	2926

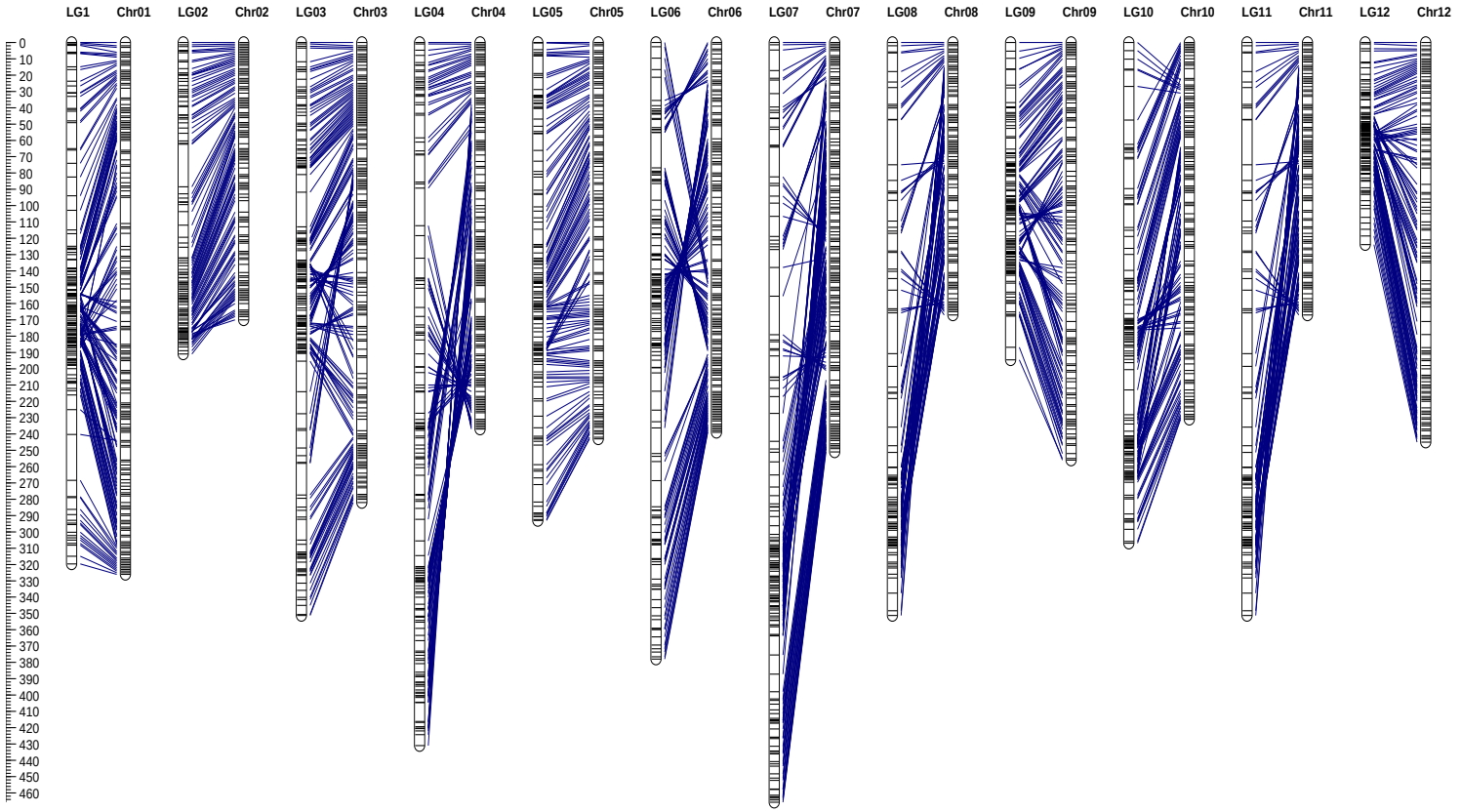


Figure 5. Comparison of the genetic map of the 'PJ' F_2 population with the physical map. Scale bar on the left shows the genetic map position (cM)

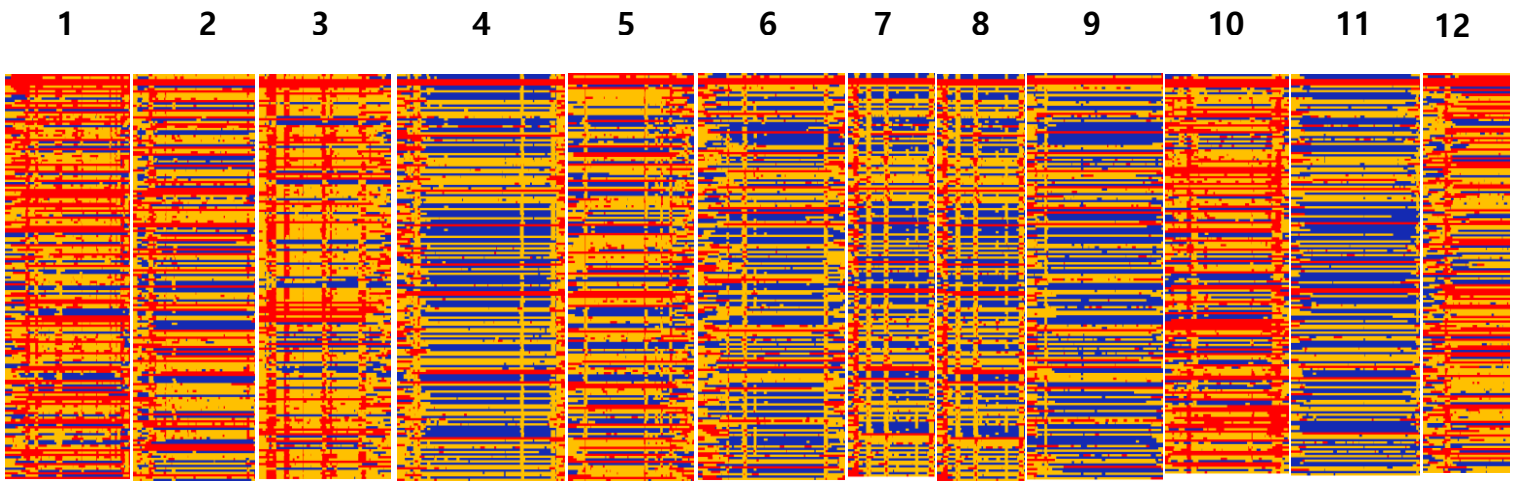


Figure 6. Bin map of the 'PJ' F_2 population. Red region indicates same genotype with 'Jeju', blue means 'PI260433' and yellow means heterozygous genotype

QTLs associated with capsaicinoid content

QTLs controlling the content of total capsaicinoid contents were detected on 'PJ' F_2 population (Table 4). The phenotype data and an highly-density bin map were used to identify genetic factor. The QTLs were detected on chromosome 3 and chromosome 9. On chromosome 3, 7 QTLs were detected and among them *PJ_Gibbs3.6* was located near 250 -262 Mbp with the highest LOD thresholds. Furthermore, *PJ_Gibbs3.4* was located on 189 – 205 Mbp showing the highest R^2 score explaining 46% of total phenotypic variation. On chromosome 9, 2 QTLs were detected. Among those, *PJ_Gibbs9.2* located on 103 -108 Mbp showed the highest LOD score and *PJ_Gibbs9.1* showed higher R^2 score explaining 29.4% of total phenotypic variation. The position of each QTLs are shown in Figure 8. The physical position of QTLs detected on chromosome 3 and chromosome 9 were estimated. The *C. annuum* 'Dempsey' reference genome was used to compare the physical locations of QTLs. In case of QTLs detected on chromosome 3, *pAMT* gene was positioned on *PJ_Gibbs3.5* located on 263.8 – 284.8 Mbp (Figure 7).

To validate the effect of the QTLs, individual plants in the 'PJ' F_2 population were sorted into three groups according to their genotypes ('PI260433-np' or 'Jeju' and heterozygote) at bin markers located within the QTL. The box plot was drawn with Bin7_25_1M_249000000 for QTL detected on chromosome 3 and Bin10_25_1M_204000000 for QTL detected on chromosome 9. All the markers were highly associated with the capsaicinoid content in the placenta of these plants (Figure 9).

Table 4. Quantitative trait loci (QTL) for capsaicinoid content in the placenta detected in 'PJ' F_2 population.

Trait	Chromosome	QTL	Position	LOD	Additive	Dominant	R^2	Location
Capsaicinoid	3	*PJ_Capsaicinoid3.1*	136.31	4.2	-0.3	0.5	16.5	135.4-136.9
Capsaicinoid	3	*PJ_Capsaicinoid3.2*	141.71	4.3	-0.3	0.5	17.6	140.2-144.6
Capsaicinoid	3	*PJ_Capsaicinoid3.3*	175.01	4.5	-0.3	0.5	18.8	173.4-178.6
Capsaicinoid	3	*PJ_Capsaicinoid3.4*	197.31	4.8	-0.4	0.7	46.4	189.1-205.6
Capsaicinoid	3	*PJ_Capsaicinoid3.5*	274.01	4.6	-0.3	0.7	23.2	263.8-284.8
Capsaicinoid	3	*PJ_Capsaicinoid3.6*	291.91	5.1	-0.5	0.5	13.0	290.6-292.1
Capsaicinoid	3	*PJ_Capsaicinoid3.7*	311.61	19.6	0.6	0.7	1.4	309-312.3
Capsaicinoid	9	*PJ_Capsaicinoid9.1*	100.31	5.1	-0.5	0.6	29.4	99.4-101.2
Capsaicinoid	9	*PJ_Capsaicinoid9.2*	106.91	6.2	-0.4	0.6	25.9	103.3-108.8

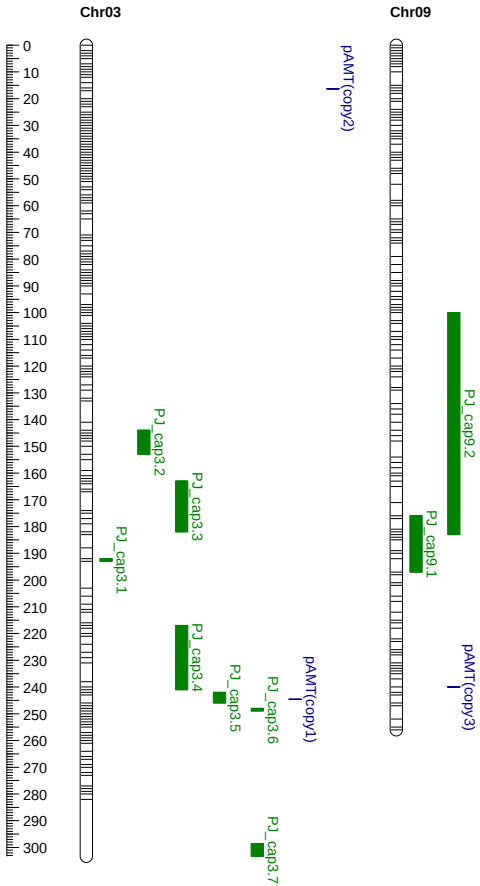


Figure 7. Result of QTL analysis and comparison of QTLs with *pAMT* gene.

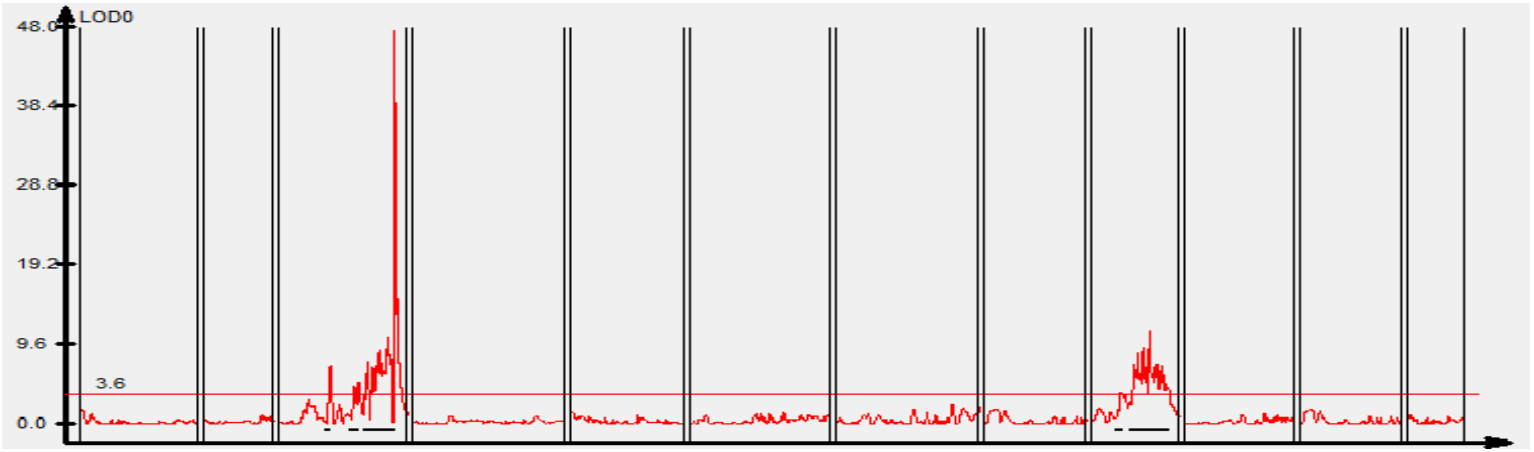


Figure 8. Position of QTLs for total capsaicinoid content in placenta tissues for 'PJ' population on chromosome 3 and chromosome 9.

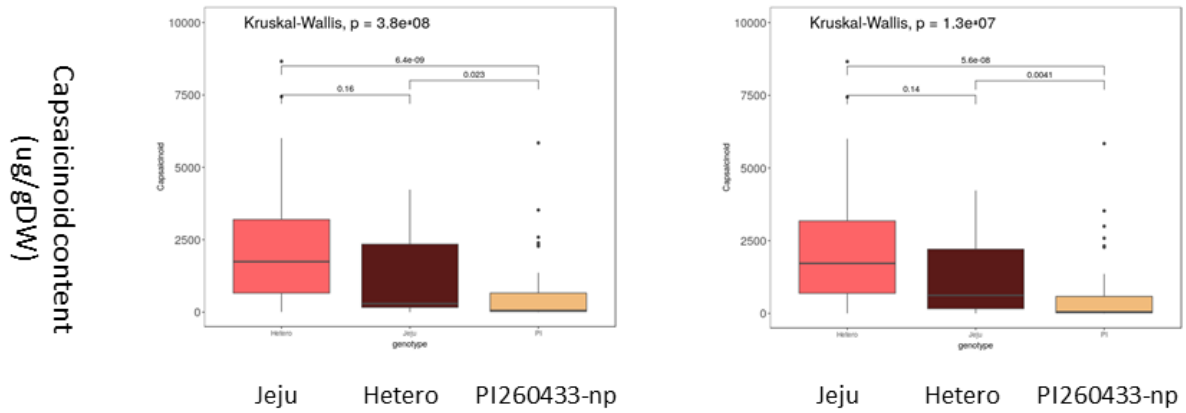


Figure 9. Comparison of QTLs and SNPs associated with capsaicinoid content and box plots of capsaicinoid content regulated by two markers representing the two QTLs. Bin markers used for box plot were Bin7_25_1M_249000000 (a), and Bin10_25_1M_204000000 (b) 'PJ' F_2 population, respectively

DISCUSSION

The *Capsicum* genus contains five domesticated species and *C. annuum* is the most commonly cultivated and economically important of the domesticated species. *Capsicum chacoense* is a wild chilli pepper species with small, evergreen shrub growing about 1 meter tall. It is sometimes harvested from the wild for its fruit and very occasionally cultivated in gardens.

The pungency of pepper fruit is controlled by single genes and QTLs. *pAMT* gene was proposed to catalyze the formation of vanillin to vanillylamine in capsaicinoid biosynthesis pathway (Curry *et al*., 1999; Blumet *et al*., 2003; Abraham-Juarezet *et al*.,2008). The *Pun2* gene which attributes to loss of pungency in *Capsicum chacoense* were demonstrated with mapping and complementation test (Stellari *et al*., 2010). *Pun2* was identified to contribute loss of pungency in *C. chacoense* PI260433-np and may be related to the caps7.1 (Ben-Chaim *et al*, 2006*). C. chacoense* has been used extensively in the breeding of modern pepper cultivars as a source of resistance to viral pathogens. The possibility that caps7.1 QTL could be an ortholog of *pun2* with different expressivity or the result of a historical introgression of the *pun2* allele that has a qualitative rather than quantitative behavior in a 'Habanero' background (Stellari *et al*., 2010). No sequence variation in *Pun2* was identified in *C. chacoense* PI260433 other than a 36 bp insertion in the intron, which is conserved between pungent and non-pungent *C. chacoense* PI260433 lines (Stellari *et al*., 2010).

The plant materials used in this study were *C. annuum* 'Jeju' and *C. chacoense* 'PI260433' and used to map the *Pun2* locus. *C. annuum* 'Jeju' carries functional *Pun1* and *pAMT* genes. However, according to HPLC analysis, 'Jeju' showed slight amount of capsinoid. It demonstrates the *pAMT* gene in *C. annuum* 'Jeju'. By contrast *C. chacoense* 'PI260433' has no mutation in the *Pun1* gene. However, it was suggested that *Pun2* is not functional.

QTL analysis results suggested genetic factors controlling pungency in *C. chacoense* 'PI260433'. Our QTL analysis revealed a total of 9 QTLs associated with capsaicinoid accumulation in the placenta tissue. To compare physical location of detected QTLs, *C. annuum* 'Dempsey' reference genome was used. Among them, several QTLs shared same location and QTLs detected on chromosome 3 shared same location of *pAMT* gene.

REFERENCES

Blum E, Mazourek M, O'connell M, Curry J, Thorup T, Liu K, Jahn M, Paran I (2003) Molecular mapping of capsaicinoid biosynthesis genes and quantitative trait loci analysis for capsaicinoid content in *Capsicum*. *Theor Appl Genet* 108:79-86

Curry J, Aluru M, Mendoza M, Nevarez J, Melendrez M, O'Connell MA (1999) Transcripts for possible capsaicinoid biosynthetic genes are differentially accumulated in pungent and non-pungent *Capsicum* spp. *Plant Sci* 148:47-57

Chen Z, Wang B, Dong X, Liu H, Ren L, Chen J, Hauck A, Song W, Lai J (2014) An ultra-high density bin-map for rapid QTL mapping for tassel and ear architecture in a large F_2 maize population. *BMC Gnomics* 15:433

del Rosario Abraham-Juárez M, del Carmen Rocha-Granados M, López MG, Rivera-Bustamante RF, Ochoa-Alejo N (2008) Virus-induced silencing of *Comt*, *pAmt* and *Kas* genes results in a reduction of capsaicinoid accumulation in chili pepper fruits. *Planta* 227:681-695

De Givry S, Bouchez M, Chabrier P, Milan D, Schiex T (2004) CARHTAGENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics* 21:1703-1704

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19:11-15

Han K, Jeong HJ, Yang HB, Kang SM, Kwon JK, Kim S, Choi D, Kang BC (2016) An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper (*Capsicum annuum*). *DNA Research* 23:81-91

Huang X, Feng Q, Qian Q, Zhao Q, Wang L, Wang A, Guan J, Fan D, Weng Q, Huang

T, Dong G, Sang T, Han B (2009) High-throughput genotyping by whole-genome resequencing. *Genome Res* 19:1068-1076

Jeong HJ, Hwang DY, Ahn JT, Chun JY, Han KE, Lee WM, Kwon JK, Lee YJ, Kang BC (2012) Development of a simple method for detecting capsaicinoids using Gibb's reagent in pepper. *Korean J Hortic Sci* 30:294-300

Kobata K, Todo T, Yazawa S, Iwai K, Watanabe T (1998) Novel capsaicinoid-like substances, capsiate and dihydrocapsiate, from the fruits of a nonpungent cultivar, CH-19 Sweet, of pepper (*Capsicum annuum L.*). *J Agr Food Chem* 46:1695-1697

Koeda S, Sato K, Tomi K, Tanaka V,Takisawa R, Hosokawa M, Doi M, Nakazaki T, Kitajima A (2014) Analysis of non-pungency, aroma, and origin of a *Capsicum chinense* cultivar from a Caribbean island. *J Jpn Soc Hortic Sci* CH-105

Lang Y, Kisaka H, Sugiyama R, Nomura K, Morita A, Watanabe T, Tanaka Y, Yazawa S, Miwa T (2009) Functional loss of *pAMT* results in biosynthesis of capsinoids, capsaicinoid analogs, in *Capsicum annuum* cv. CH-19 Sweet. *Plant J* 59:953-961

Park YJ, Nishikawa T, Minami M, Nemoto K, Iwasaki T, Matsushima K (2015) A low-pungency S3212 genotype of *Capsicum frutescens* caused by a mutation in the putative aminotransferase (*p-AMT*) gene. *Mol Genet Genomics* 290:2217-2224

Stellari GM, Mazourek M, Jahn MM (2010) Contrasting modes for loss of pungency between cultivated and wild species of *Capsicum*. *Heredity* 104:460

Tanaka Y, Hosokawa M, Miwa T, Watanabe T, Yazawa S (2010) Novel loss-of-function putative aminotransferase alleles cause biosynthesis of capsinoids, nonpungent capsaicinoid analogues, in mildly pungent chili peppers (*Capsicum chinense*). *J Agr Food Chem* 58:11762-11767

Tanaka Y, Sonoyama T, Muraga Y, Koeda S, Goto T, Yoshida Y, Yasuba K (2015) Multiple loss-of-function putative aminotransferase alleles contribute to low

pungency and capsinoid biosynthesis in *Capsicum chinense*. *Molcular Breeding* 35:142

Tanaka Y, Nakashima F, Kirii E, Goto T, Yoshida Y, Yasuba K.I (2017) Difference in capsaicinoid biosynthesis gene expression in the pericarp reveals elevation of capsaicinoid contents in chili peppers (*Capsicum chinense*). *Plant Cell Rep* 36:267-279

Truong HT, Ramos AM, Yalcin F, de Ruiter M, van der Poel HJ, Huvenaars KH, Hogers RCJ, Van Enckevort LJG, Janssen A, Van Orsouw NJ, Van Eijk MJ (2012) Sequence-based genotyping for marker discovery and co-dominant scoring in germplasm and populations. *PloS One* 7:e37565

Tsurumaki K, Sasanuma T (2019) Discovery of novel unfunctional *pAMT* allele *pamt*10 causing loss of pungency in sweet bell pepper (*Capsicum annuum L.*). *Breed Sci* 69:133-142

Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered* 93:77-78

Wang S, Basten CJ, Zeng ZB (2012) Windows QTL Cartographer 2.5. Raleigh, NC: Department of Statistics, North Carolina State University

Yazawa S, Ueda M, Suetome N, Namiki T (1989) Capsaicinoids content in the fruit of interspecific hybrids in *Capsicum*. *J Jpn Soc Hortic Sci* 58:353-360

ABSTRACT IN KOREAN

Capsicum 종은 캡사이시노이드의 축적에 의해서 매운 맛 정도의 차이가 나타난다. 고추의 캡사이시노이드 생합성은 *Pun1*, *Pun2* 그리고 *pAMT* 유전자들에 의해 결정이 된다. 캡사이시노이드의 함량은 양적유전자에 의해 조절이 된다고 알려져 있다. 본 연구에서는 고추 종 (*Capsicum* spp.)에서 캡사이시노이드 함량을 조절하는 추가적인 유전적 요인을 구명하기 위하여 진행이 되었다. 첫 번째 장에서는 매운 고추로 알려진 *Capsicum chinense* 'Jolokia'와 맵지 않은 고추로 알려진 *C. chinense* 'SNU11-001'를 교배한 F_2 집단에서 QTL분석을 하였다. *C. chinense* 'SNU11-001'은 *pAMT* 유전자 돌연변이에 의해 캡사이시노이드 생합성 과정이 진행되지 않는 점에서 착안하여, KASP 유전형분석을 통하여 F_2 집단에 대하여 *pAMT* 유전형을 분석하여 전체 F_2 집단과 정상 *pAMT* 유전자를 지니는 집단 두 가지로 나누었다. 전체 F_2 집단에 대하여 QTL분석을 한 결과, 모두 염색체 3번에서 QTL에 분포하였으며, *pAMT* 유전자 지역을 공통적으로 포함하는 것을 알 수 있었다. 정상 *pAMT* 유전자를 지니는 집단에 대하여 QTL분석을 실시한 결과, 디하이드로캡사이신 함량을 조절하는 QTL이 5번 염색체와 11번 염색체에 분포하였다. 두 번째 장에서는 두 가지의 유전자원을 이용하여 *C. chacoense*에서 캡사이시노이드의 유무를 결정하는 새로운 유전자 동정을 하였다. 맵지 않은 고추로 알려 진 *C. chacoense* 'PI260433-np'는 *Pun1*

유전자를 보유하고 있으며, 열성 유전자 *pun2*를 보유하여 캡사이시노이드와 캡시노이드를 생합성하지 않는다. 매운 고추로 알려진 *C. annuum* 'Jeju'는 *Pun1* 유전자와 *pAMT* 유전자를 보유하여 캡사이시노이드와 캡시노이드 합성에 관여한다. 상보성 검사를 통해 *C. chacoense* 'PI260433-np'는 새롭게 발견된 유전자좌에 의해 캡사이시노이드 합성 유무를 결정함을 확인하였다. QTL분석 결과, 염색체 3번과 염색체 9번에서 QTL을 발견하였으며, 염색체 3번에서 발견된 QTL은 *pAMT* 유전자 위치에 해당함을 확인하다. 염색체 9번에서 발견 된 QTL은 *Pun2* 유전자좌를 포함하고 있음을 알 수 있었다. 두 연구 결과로부터, 고추의 캡사이시노이드 함량을 조절하는 QTL들이 염색체 3번, 염색체 5번 그리고 염색체 11번에 위치하고 있음을 확인하였으며, *Pun2* 유전자를 맵핑하였다. 해당 연구 결과를 통하여, 고신미 품종을 육성하고 캡사이시노이드 합성에 관여하는 유전자에 대한 이해를 높일 수 있을 것으로 기대한다.

주요어: 캡사이시노이드, 차세대 염기서열 분석법 (NGS), 형광기반 분석 (KASP), 고추, 양적형질 유전자 (QTL)