

## 당근 종모 형질 관련 cDNA Library 작성 및 EST 분석

오규동<sup>1</sup> · 심은조<sup>1</sup> · 전상진<sup>2</sup> · 박영두<sup>1\*</sup>

<sup>1</sup>경희대학교 원예생명공학과, <sup>2</sup>(주)캐로톱씨드 육종연구소

### Construction of cDNA Library and EST Analysis Related to Seed-hair Characteristics in Carrot

Gyu-Dong Oh<sup>1</sup>, Eun-Jo Shim<sup>1</sup>, Sang-Jin Jun<sup>2</sup>, and Young-Doo Park<sup>1\*</sup>

<sup>1</sup>Department of Horticultural Biotechnology, Kyung Hee University, Yongin 446-701, Korea

<sup>2</sup>Breeding Research Institute, Carrotop Seed Co., Anseong 456-882, Korea

**Abstract.** Carrot (*Daucus carota* L. var. *sativa*) is one of the most widely used crops in the world and is nutritionally important crop. However, seed-hair which is generated in epidermal cell of seeds causes the difficulty of the seedling process, because of the seed germination and absorption inhibitions. For these reasons, carrot seeds are commercialized after mechanical hair removal process. However, in this process, various damage and seed loss occur and breeding of hairless-seed carrot cultivar is needed to overcome these various weaknesses and additional seed production costs. In this study, cDNA libraries using 2 combinations, which were composed of short-hair seed CT-ATR 615 OP 666-13 & long-hair seed CT-ATR 615 OP 671-9, and short-hair seed CT-SMR 616 OP 659-1 & long-hair seed CT-SMR 616 OP 677-14, were constructed and EST sequences of each individuals were analyzed to reveal carrot seed-hair characteristics. Firstly, analyzed EST sequences were classified into FunCat functional categories. As a result, significant differences have been identified in metabolism category, protein folding and stabilization, protein binding, C-compound binding category from both of two combinations. Secondly, several candidate EST sequences related to seed trichome differentiation and cellulose biosynthetic process were selected based on GO data of EST sequences. These differences based on FunCat categories and candidate EST obtained by GO data analysis are thought to be involved in the formation of carrot seed hair. Finally, 741 SSR sites and 33 SNP sites were identified from analyzed EST sequences of two combinations. Then we designed SNP and SSR primer sets to develop molecular markers. These molecular markers will be used for classification of carrot cultivars and study seed-hair characteristic.

**Additional key words:** BlastX, *Daucus carota*, FunCat, SNP, SSR

## 서 언

당근(*Daucus carot* L. var. *sativa*)은 종자 수정의 증거로 표피세포에서 연장되는 형태로 종자모를 생성하는데 이는 당근 종자 파종 및 발아에 여러 가지 좋지 않은 영향을 준다 (Menon and Dave, 1989). 종자모는 발아억제 물질인 carotol 을 함유하고 있으며, 종자 발아 시 흡수를 방해하고 파종 시 종자모들의 엉킴으로 인해 취급상 어려움을 갖게 한다. 이

러한 문제점의 해결을 위해 기존 시판 당근종자들은 채종 후 기계적인 방법을 통해 종자모가 제거된 상태로 판매되고 있다. 그러나 물리적인 힘에 의한 제거과정에서 종자의 손실이 발생하며, 이에 따라 불균일한 발아율을 보이는 경향이 나타난다. 더불어 종자모 제거를 위한 추가적인 인력 및 경제적인 부담이 발생하고 있다. 이에 따라 종자모의 형성이 억제된 무모종자 당근 품종의 육종이 요구되고, 현재 진행과정 중에 있다.

\*Corresponding author: ydpark@khu.ac.kr

※ Received 11 July 2013; Revised 25 July 2013; Accepted 31 July 2013. 본 논문은 농촌진흥청 차세대 바이오그린21사업(과제번호: PJ008076 과 PJ009043)의 지원에 의해 이루어진 것임.

© 2013 Korean Society for Horticultural Science

한편 cDNA library의 임의적인 염기서열 분석은 해당 형질에 대한 유전학적 접근을 가능하게 하며, 이를 통해 확인된 expressed sequence tag(EST) 염기서열을 바탕으로 형질 관련 후보 유전자 선발 및 functional category의 분포를 확인할 수 있다(Adams et al., 1991). 또한 각각의 EST 염기서열의 BlastX 결과를 바탕으로 기능별로 분류하기 위해 사용하는 FunCat(functional catalogue)은 단백질 기능의 체계적인 분류를 위해 확립된 방법으로 28개의 main category와 세부적인 category로 구성되어 있다(Mewes et al., 2004; Ruepp et al., 2004). 이와 더불어 특정 형질과 관련된 후보 유전자의 선발을 위해 각각의 유전자와 해당 유전자를 통해 생성되는 단백질의 특성의 비교를 통해 작성된 gene ontology(GO) data를 이용한다(Ashburner et al., 2000).

cDNA library 작성을 통해 확인된 EST 염기서열을 이용한 simple sequence repeats(SSR) 및 single nucleotide polymorphism(SNP) 분자표지는 DNA 유래의 분자 표지로서 널리 이용되고 있다. EST-SSRs 및 EST-SNPs는 EST 염기서열을 이용함으로써 특정 형질의 발현과 관련된 분자 표지 개발에 용이할 뿐 아니라, 개발에 이용된 식물체의 관련 종에 적용이 가능하다는 장점을 가지고 있다(Gupta et al., 2003; Somers et al., 2003).

본 연구에서는 당근 종자모 형질관련 연구를 위해 cDNA library를 작성 EST 염기서열 분석을 통해 장모 및 단모 종자 표현형 당근 개체간 EST 분석과 종자모 형질 관련 후보 유전자 선발을 수행하였으며, SSR 및 SNP site 확인을 통해 단모당근 품종 육종에 이용하고자 한다.

## 재료 및 방법

### 식물재료

본 연구를 위해서 2008년도부터 2010년도에 걸쳐 (주)캐로 톱씨드(안성, 경기도)에서 육성된 2계통의 단모 및 장모 개체를 사용하였다. CT-ATR 615 계통 및 CT-SMR 616 계통은 동일한 인도 유래종인 imperator type의 Fusa red계 당근으로서 초기 추대성을 보이며 안정적인 추근성과 평균적인 초세를 나타내며, 근미비대 중만생, 중형의 뿌리 비대성, 중장형의 뿌리길이를 보이는 숙성 중만생의 특성을 가진 당근이다. 2008년도에 CT-ATR 615 OP 계통 및 CT-SMR 616 OP계통의 형매 교배를 통해 2009년도에 CT-ATR 615 OP 394와 CT-SMR 616 OP 389 계통을 육성하였으며, 그 중 단모종자 표현형을 보인 394-6개체와 389-1개체를 각각 자가수분 하여 2010년도에 CT-ATR 615 OP 666 계통 및 CT-SMR 616 OP 659 계통을 육성하였다. 육성된 계통 중

당근 종자모 형질관련 cDNA library 작성을 위한 실험재료는 총 두 조합으로 한 조합은 단모형 종자 표현형 CT-ATR 615 OP 666-13개체 및 대조군 CT-ATR 615 OP 671-9 개체를 선발하였으며, 또 다른 조합으로 단모형 종자 표현형 CT-SMR 616 OP 659-1 개체 및 대조군 CT-SMR 616 OP 677-14 개체를 선발하여 이용하였다(Hwang et al., 2010; Oh et al., 2010).

또한 개체별 종자 표현형 확인을 위해 Video Microscope System(Alphasystec Co., Korea)을 이용하여 종자모 길이 측정을 실시하였으며, 측정 시 개체별 5종자간 5개의 모를 측정하였다.

### cDNA Library 작성

당근 종자모 형성에 관련하여 단모종자 표현형 개체와 대조군 장모종자 표현형 개체를 이용하여 cDNA library를 작성하였다. 단모종자 표현형 CT-ATR 615 OP 666-13, CT-SMR 616 OP 659-1개체와 장모종자 표현형을 보인 CT-ATR 615 OP 671-9, CT-SMR 616 OP 677-14 개체를 각각 자가수분하여 자방이 팽창되기 시작할 시기 즉, 종자모가 생성되기 시작하는 시기에 각각의 종자 샘플을 채취하였다. 채취된 각각의 샘플로부터 total RNA를 분리하였으며, 전기영동을 통해 분리 상태와 농도를 확인하였다. 이후 분리된 total RNA에서 mRNA 농도를 확인하고 Oligo(dT) linker-primer와 reverse transcriptase를 이용하여 first-strand cDNA를 합성하였으며 DNA polymerase I을 이용해 second-strand 를 합성한 뒤 adaptor를 ligation하였다. 다음 단계로 colony-forming units titer  $0.2 \times 10^7$  cfu(total 20mL)의 값에서 pBluescript® SK- vector를 이용하여 excised phagemid type의 cDNA library를 작성하였다(Macrogen Inc., Korea)(Wang et al., 2006).

### EST 염기서열 분석

단모종자 표현형 개체와 대조군 장모종자 표현형 개체 2 조합을 이용해 작성된 cDNA library의 EST 염기서열 분석은 Macrogen Inc.(Korea)에 의뢰하여 ABI 3730(Applied Biosystems Co., USA)을 이용하여 실시하였다. 각각의 fragment 염기서열 분석 결과는 EST sequence assembly를 통해 contig와 singleton으로 작성하였으며 각각의 EST 염기서열은 NCBI의 Blast Program(V.2.2.10)을 이용하여 BlastX를 수행하였다. 확보된 각각의 EST 염기서열의 BlastX 결과는 단모종자 및 장모종자 개체별 발현양상의 비교분석을 위해 functional catalogue(FunCat, [http://mips.helmholtz-muenchen.de/proj/funcatDB/search\\_main\\_frame.html](http://mips.helmholtz-muenchen.de/proj/funcatDB/search_main_frame.html))를 이용하여 ESTs를 기능별 category로 구분하였다(Mewes et al., 2004; Ruepp et

al., 2004). 이와 더불어, 당근 종자모 형성에 관련된 후보 유전자 분석을 위해 각각의 EST의 BlastX 결과를 gene ontology (GO) data값을 바탕으로 비교분석 하였다(Ashburner et al., 2000).

### SNP와 SSR Site 분석

cDNA library작성을 통해 분석된 EST 염기서열상의 SNP site를 확인하기 위해 단모종자 표현형을 보인 CT-ATR 615 OP 666-13 및 CT-SMR 616 OP 659-1개체와 대조군 장모종자 표현형을 보인 CT-ATR 615 OP 671-9 및 CT-SMR 616 OP 677-14 개체 2 조합의 contig와 singleton EST 염기서열을 Sequencher ver. 4.10.1 program을 이용해 assembly한 비교분석을 통해 단모 및 장모 개체간 EST 염기서열상의 SNP site를 확인하였다(Ayeh, 2008; Kantety et al., 2002). 또한 WebSat site(<http://wsmartins.net/websat>)를 이용하여 각각의 개체 EST 염기서열상의 SSR site를 확인하였으며, mono, di, tri, tetra, penta, hexa motif에서 5 repeats를 최소값으로 설정하여 분석하였다(Varshney et al., 2005; Yu et al., 2004).

## 결과 및 고찰

### 계통 육성 및 표현형 조사

당근 종자모 형질 관련 cDNA library 작성을 위한 식물재료 선별을 위해 CT-ATR 615 OP 666 개체군과 CT-SMR 616 OP 659 개체군의 개체별 종자모 표현형을 조사하였다. 각각의 종자모 길이 측정 결과는 CT-ATR 615 OP 666-13개체가 CT-ATR 615 OP 671-9개체 종자모의 32.6%, CT-SMR 616 OP 659-1 개체가 CT-SMR 616 OP 677-14개체 종자모의 13.9% 길이의 종자모를 가지는 것으로 조사되었다. 그 결과를 바탕으로 단모종자 표현형 개체 CT-ATR 615 OP 666-13과 CT-SMR 616 OP 659-1 개체를 선별하였다. 또한

대조군으로서 장모종자 표현형 개체 CT-ATR 615 OP 671-9와 CT-SMR 616 OP 677-14개체를 선별하였다(Hwang et al., 2010; Oh et al., 2010).

### cDNA Library 작성

EST 비교 분석을 위하여 단모종자 표현형 개체와 장모종자 표현형 개체를 이용하여 당근 종자모 형질 관련 cDNA library를 작성하였다. 각각의 개체에서 분리한 total RNA의 농도는 CT-ATR 615 OP 666-13에서  $549.4\text{ng}\cdot\mu\text{L}^{-1}$ 로 확인되었으며, CT-ATR 615 OP 671-9에서  $930.5\text{ng}\cdot\mu\text{L}^{-1}$ , CT-SMR 616 OP 659-1에서  $2536\text{ng}\cdot\mu\text{L}^{-1}$ , CT-SMR 616 OP 677-14에서  $2172\text{ng}\cdot\mu\text{L}^{-1}$ 로 확인되었다. 또한 mRNA 분리 후 농도를 확인 결과 CT-ATR 615 OP 666-13에서  $95.5\text{ng}\cdot\mu\text{L}^{-1}$ 로 확인되었으며, CT-ATR 615 OP 671-9에서  $128.1\text{ng}\cdot\mu\text{L}^{-1}$ , CT-SMR 616 OP 659-1에서  $111.5\text{ng}\cdot\mu\text{L}^{-1}$ , CT-SMR 616 OP 677-14에서  $72.7\text{ng}\cdot\mu\text{L}^{-1}$ 로 확인되었다. 이를 이용해 작성된 cDNA library에서 제한효소 *EcoRI*과 *XhoI*를 이용하여 각각의 개체로부터 무작위로 선별된 8개씩의 colony에서 분리한 플라스미드 DNA에 대한 insertion check를 수행하여 확인 후 cDNA library 작성을 마무리 하였다. 마지막으로 작성된 cDNA library에 대한 sequencing을 위해 개체별  $96 \times 13$  plates의 cell stock을 작성하였다.

### EST 염기서열의 FunCat 분석 및 GO Data 비교 분석

단모종자 표현형 개체 및 대조군 장모종자 표현형 개체 2조합의 4개체에서 각각 작성된 cDNA library를 이용하여 개체 별 1248개의 EST 염기서열 분석을 수행하였다. 그 결과 CT-ATR 615 OP 666-13개체와 대조군 CT-ATR 615 OP 671-9개체에서 97 contig, 772 singleton data 총 874,420 base pair 및 107 contig, 786 singleton data 총 911,005 base pair의 결과를 각각 확인하였다. 또한 CT-SMR 616 OP 659-1개체와 대조군 CT-SMR 616 OP 677-14개체에서 98 contig,

**Table 1.** EST sequencing, cleaning process and assembly results of short-hair seed individuals CT-ATR 615 OP 666-13 and CT-SMR 616 OP 659-1, and long-hair seed individuals CT-ATR 615 OP 671-9 and CT-SMR 616 OP 677-14.

Classification of EST sequencing step	CT-ATR 615 OP 666-13		CT-ATR 615 OP 671-9		CT-SMR 616 OP 659-1		CT-SMR 616 OP 677-14	
	No. of reads	Total bp	No. of reads	Total bp	No. of reads	Total bp	No. of reads	Total bp
Total EST sequences	1,244	874,420	1,231	911,005	1,240	797,105	1,236	954,718
After 3730 sequencing	1,238		1,204		1,236		1,233	
After seq cleaning process <sup>2</sup>	1,197		1,159		1,178		1,191	
Trashed reads <sup>2</sup>	41		45		58		42	
After clustering contigs	97	90,923	107	102,182	98	84,126	90	85,680
After clustering singletons	772	545,617	786	611,398	761	478,697	827	651,535

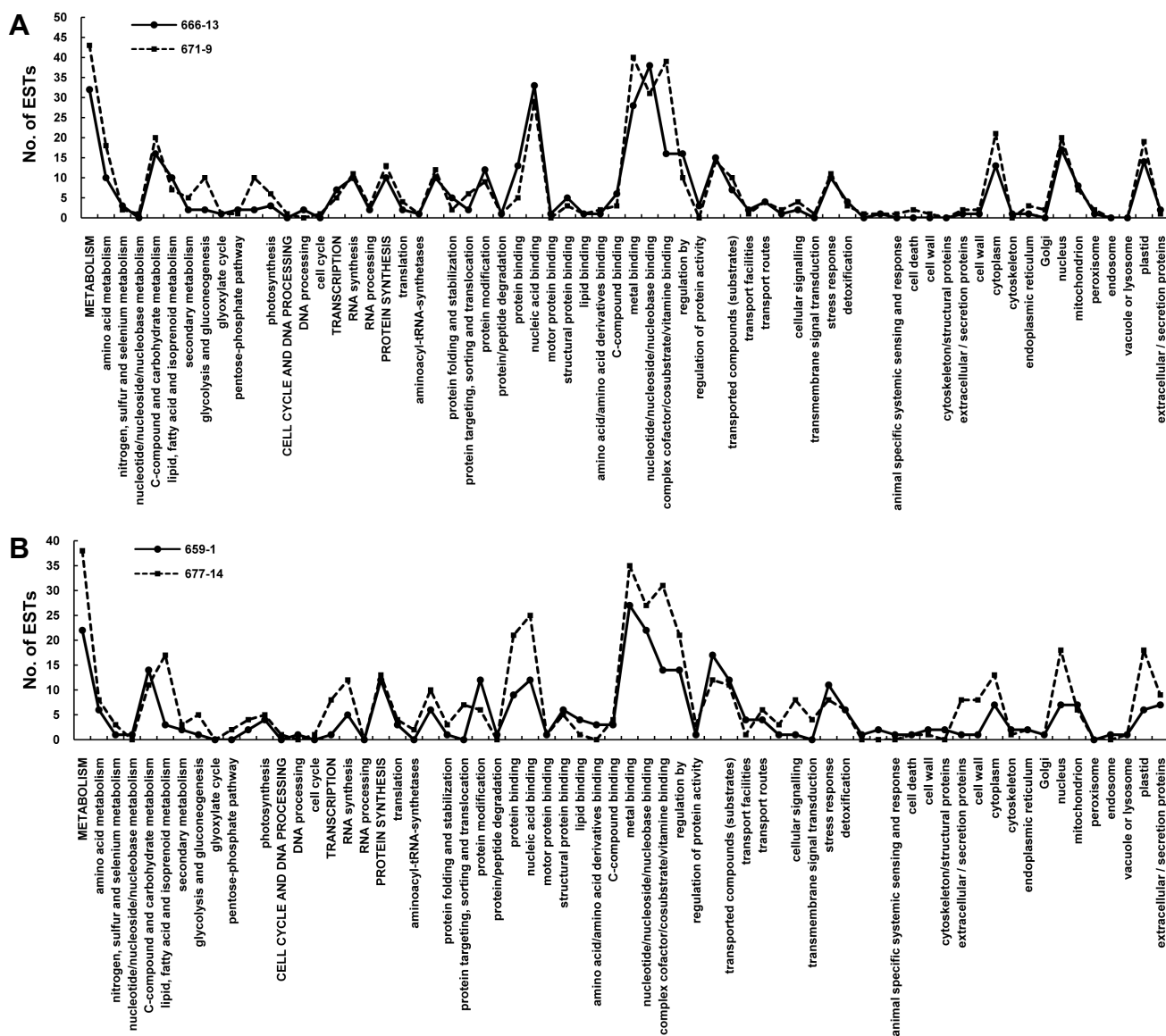
<sup>2</sup>Minimum length: 200 bp.

761 singleton data 총 797,105 base pair 및 90 contig, 827 singleton data 총 954,718 base pair의 결과를 각각 확인하였다(Table 1).

각각의 개체로부터 확보된 EST 염기서열의 BlastX 분석을 실시한 결과 CT-ATR 615 OP 666-13과 CT-ATR 615 OP 671-9 개체에서 각각 51%, 54%의 염기서열에서 유의적인 발현정보를 획득하였다. 또한 CT-SMR 616 OP 659-1와 CT-SMR 616 OP 677-14 개체에서 각각 42%, 49%의 염기서열에서 유의적인 발현정보를 획득하였다.

이들 유의적인 발현정보를 보인 EST 염기서열의 BlastX 결과를 각각의 두 조합에서 단모종자 표현형 개체와 대조군 장모종자 표현형 개체간 비교분석을 하였다. 먼저 첫 번째

조합(CT-ATR 615 OP 666-13와 CT-ATR 615 OP 671-9)에서 두 개체간 동일한 BlastX 결과를 보인 CT-ATR 615 OP 666-13 개체의 332 EST, CT-ATR 615 OP 671-9개체의 330 EST를 제외한 뒤 각각의 개체에서만 확인된 334개(CT-ATR 615 OP 666-13)와 365개(CT-ATR 615 OP 671-9)의 EST, 그리고 두 번째 조합(CT-SMR 616 OP 659-1와 CT-SMR 616 OP 677-14)에서 두 개체간 동일한 BlastX 결과를 보인 CT-SMR 616 OP 659-1개체의 261 EST, CT-SMR 616 OP 677-14 개체의 304 EST를 제외한 뒤 각각의 개체에서만 확인된 303개(CT-SMR 616 OP 659-1)와 351개(CT-SMR 616 OP 677-14)의 EST 염기서열의 BlastX 결과를 FunCat 기능적 category에 따라 비교하였다(Fig. 1).



**Fig. 1.** Comparison of BlastX FunCat results of (A) short-hair seed individual CT-ATR 615 OP 666-13 and long-hair seed individual CT-ATR 615 OP 671-9 EST sequences, (B) short-hair seed individual CT-SMR 616 OP 659-1 and long-hair seed individual CT-SMR 616 OP 677-14 EST sequences depending on FunCat categories except same results of BlastX between two individuals.

비교한 결과를 바탕으로 단모종자 표현형 개체와 대조군 장모종자 표현형 개체에서 확인된 EST의 빈도가 상호간 50% 이상 차이 나는 category를 선별하였다. CT-ATR 615 OP 666-13 개체와 CT-ATR 615 OP 671-9 개체간 비교 결과 단모종자 표현형 개체에선 protein folding and stabilization, protein binding, C-compound binding category에서 높은 빈도의 EST를 확인하였으며, 장모종자 표현형 개체에선 amino acid metabolism, secondary metabolism, glycolysis and gluconeogenesis, tricarboxylic-acid pathway, photosynthesis, protein targeting, sorting and translocation, complex cofactor/cosubstrate/vitamin binding, cytoplasm category에서 높은 빈도의 EST를 확인하였다. 한편 CT-SMR 616 OP 659-1 개체와 CT-SMR 616 OP 677-14 개체간 비교 결과 단모종자 표현형 개체에선 protein modification, lipid binding, amino acid/amino acid derivatives binding category에서 높은 빈도

의 EST를 확인하였으며, 장모종자 표현형 개체에서 metabolism, lipid, fatty acid and isoprenoid metabolism, glycolysis and gluconeogenesis, pentose-phosphate pathway, tricarboxylic-acid pathway, protein targeting, sorting and translocation, complex cofactor/cosubstrate/vitamin binding, extracellular/secretion proteins, cell wall 등의 category에서 높은 빈도의 EST를 확인하였다(Table 2). 이러한 category에 속하는 EST가 당근 종자모 형질에 깊이 관여 할 가능성이 높을 것으로 예상되며, 특히 두 조합 모두 장모종자 표현형 개체에서 단모종자 표현형 개체에 비해 metabolism category에서 유의적인 EST 빈도의 차이를 보이는 것으로 확인되었다. 또한 complex cofactor/cosubstrate/vitamin binding category에서의 차이가 두 조합에서 두드러지게 나타나 해당 category에서 확인된 EST를 대상으로 종자모 형성과정에 영향을 미칠 수 있는 후보군을 선별하여 연구에 이용할 수 있으리라 기대된다.

**Table 2.** Categories showing difference of expression more than 50% between CT-ATR 615 OP 666-13 & CT-ATR 615 OP 671-9 and CT-SMR 616 OP 659-1 & CT-SMR 616 OP 677-14.

CT-ATR 615 OP 666-13 > CT-ATR 615 OP 671-9	CT-ATR 615 OP 666-13 < CT-ATR 615 OP 671-9
Protein folding and stabilization	Amino acid metabolism
Protein binding	Secondary metabolism
C-compound binding	Glycolysis and gluconeogenesis
	Tricarboxylic-acid pathway
	Photosynthesis
	Protein targeting, sorting and translocation
	Complex cofactor/cosubstrate/vitamine binding
	Cytoplasm
CT-SMR 616 OP 659-1 > CT-SMR 616 OP 677-14	CT-SMR 616 OP 659-1 < CT-SMR 616 OP 677-14
Protein modification	Metabolism
Lipid binding	Lipid, fatty acid and isoprenoid metabolism
Amino acid/amino acid derivatives binding	Glycolysis and gluconeogenesis
	Pentose-phosphate pathway
	Tricarboxylic-acid pathway
	Transcription
	RNA synthesis
	Protein targeting, sorting and translocation
	Protein binding
	Nucleic acid binding
	Complex cofactor/cosubstrate/vitamine binding
	Regulation by
	Cellular signaling
	Transmembrane signal transduction
	Extracellular / secretion proteins
	Cell wall
	Cytoplasm
	Nucleus
	Plastid

이러한 단모종자 개체 및 장모종자 개체간 비교분석 결과를 토대로 연구를 진행하는 한편, 개체 별 1248개의 EST 염기서열은 두 차례에 걸쳐 분석하였다. 1차에서 분석된 염기서열의 FunCat 결과와 2차로 분석된 염기서열을 통합적으로 분석하는 과정에서 보다 뚜렷한 차이점을 확인할 수 있었으며, 차후 추가적인 염기서열분석을 통해 보다 명확하고 일반화된 결과를 얻고자 한다.

FunCat category를 통한 분석과 함께 EST 염기서열의 GO 값 중 종자모 형질에 관련되어 있다고 생각되는 seed trichome differentiation tree 및 cellulose biosynthetic process tree에 속하는 EST를 선발하였다. Seed trichome differentiation tree는 크게 trichome differentiation과 seed development process로 나뉘어 상호간 연관관계에 대해 작성된 tree이며, 특히 종자모 형성에 중요한 epidermal cell differentiation에 관련되어 EST 선발에 효과적인 것으로 생각된다. Cellulose biosynthetic process 역시 종자모 형성에 중요한 과정인 cellulose의 합성 및 축적에 관련된 후보 EST 선발에 효과적이다(Gou et al., 2007). CT-ATR 615 OP 666-13와 CT-SMR 616 OP 659-1 단모형 종자 표현형 개체에서 BlastX 결과를 바탕으로 동일하게 확인된 EST 38개 중 해당 GO 값을 가지는 4개의 EST를 선발하였으며, CT-ATR 615 OP 671-9와 CT-SMR 616 OP 677-14 장모형 종자 표현형 개체에서 BlastX 결과를 바탕으로 동일하게 확인된 EST 39개 중에서 해당 GO 값을 가지는 10개의 EST를 선발하였다(Table 3). 확인된 14개의 EST 중 당근에서 확인된 단백질인 식물 세포벽의 구조에

관여하여 epidermal cell의 신장에 영향을 주는 extensin 단백질과 cellulose 합성에 관여하는 sucrose synthase 단백질을 우선적으로 선발하였으며 선발된 EST 염기서열을 이용한 단모종자 표현형 개체 및 장모종자 표현형 개체 각각에서의 유전자 발현양상으로 비교함으로써 당근 종자모 형성 관련 유전자를 탐색하고자 한다.

### SNP 및 SSR Site 확인

확인된 EST 염기서열 분석 자료를 이용하여 SNP 및 SSR site를 확인하여 당근 종자모 형질 관련 분자마커 개발에 이용하고자 하였다. 두 조합의 단모종자 표현형 개체와 대조군 장모종자 표현형 개체로부터의 EST 염기서열을 각각 assembly하여 상동성이 높은 EST 염기서열간 SNP site를 분석하였다. 단모종자 표현형 CT-ATR 615 OP 666-13 개체와 대조군 장모종자 표현형 CT-ATR 615 OP 671-9 개체간 비교를 통해 14개의 SNP site를 확인하였다. 또한 단모종자 표현형 CT-SMR 616 OP 659-1 개체와 대조군 장모종자 표현형 CT-SMR 616 OP 677-14 개체간 비교를 통해 19개의 SNP site를 확인하였다. 확인된 SNP site의 당근 종자모 형질 관련 분자마커 개발을 위해 high resolution melting(HRM) 실험 및 allele-specific PCR 실험을 수행하기 위해 primer를 design하였으며(Myakishev et al., 2001; Oh et al., 2013; Ye et al., 2001) 작성된 primer를 이용하여 분자마커 개발을 위한 연구를 수행 중에 있다.

한편 단모종자 표현형 2개체, 대조군 장모종자 표현형 2

**Table 3.** Candidate EST sequences related to seed hair formation based on gene ontology (GO) tree of seed trichome differentiation and cellulose biosynthetic process.

EST	ID	BlastX result
Short-hair seed phenotype		
394-13_03-B01-R.ab1	Q8VYW9	Aminotransferase 1
394-13_01-D11-R.ab1	Q9FQD2	Isocitrate lyase
394-13_06-F02-R.ab1	Q3KN68	Isoflavone reductase-like protein
394-13_01-E11-R.ab1	Q9LLQ6	Seed maturation protein PM34
Long-hair seed phenotype		
671-9_11-D05-R.ab1	Q306J8	Acetoacetyl-CoA thiolase
671-9_10-C03-R.ab1	A7QEB3	Chromosome chr1 scaffold_84
671-9_12-C12-R.ab1	P42495	Cinnamyl-alcohol dehydrogenase 1
671-9_04-E12-R.ab1	O80433	Citrate synthase, mitochondrial
671-9_11-G04-R.ab1	P06599	Extensin
CL37Contig1	Q38JJ4	Fructose-bisphosphate aldolase
671-9_10-B08-R.ab1	P17783	Malate dehydrogenase, mitochondrial
671-9_12-E02-R.ab1	Q93X23	Myrcene synthase, chloroplastic
671-9_10-E04-R.ab1	P49035	Sucrose synthase isoform 1
CL36Contig1	Q38JI4	Triosephosphate isomerase

개체의 EST 염기서열을 이용하여 SSR site를 확인하였다. 단모종자 표현형 CT-ATR 615 OP 666-13 개체의 EST 염기서열에서 144 SSR site, 장모종자 표현형 CT-ATR OP 671-9 개체의 EST sequence에서 233 SSR site를 확인하였다. 또한 단모종자 표현형 CT-SMR 616 OP 659-1 개체의 EST 염기서열에서 196 SSR site, 장모종자 표현형 CT-SMR 616 OP 677-14 개체의 EST 염기서열에서 198 SSR site를 확인하였다. 확인된 각각의 SSR site를 200-300bp 크기의 primer로 작성하였으며, 이는 당근 종자모 관련 분자마커 개발에 이용할 것이다. EST 염기서열 상의 SSR site의 repeat type을 분석한 결과, 전체 741개의 SSR site 중 di-nucleotide type이 68.9%로 가장 높은 빈도로 분포하고 있었으며, tri-nucleotide type은 27.6%, tetra-nucleotide type은 28.6%, hexa-nucleotide type은 0.4%의 빈도로 분포하고 있음을 확인할 수 있었다(Fig. 2). 또한 가장 높은 빈도로 분포하고 있는 di-nucleotide type의 SSR motif를 분석한 결과, 전체 517개의 di-nucleotide motif중 AT motif가 35.4%로 가장 높은 분포율을 보였으며 CT motif 31.7%, AG motif 21.1%로 뒤를 이었다(Fig. 3). 이를 대부분의 식물에서의 SSR site 분포 정도와 비교해 보았을 때, 당근에서의 SSR site 분포 또한 유사함을 확인할 수 있었다(Powell et al., 1996; Wang et al., 1994). 이와 같이 확보한 SSR를 이용한 당근 종자모 형

성 관련 분자마커 개발은 진행 중에 있다. 한편 SSR site에 비해 적은 SNP site가 확인된 이유는 확보된 개체별 염기서열을 장모 및 단모 개체간 조합으로 다시 assembly하는 과정에서 서로 동일한 염기서열간 contig의 구성이 적었기 때문에 나타난 것으로 생각된다. 앞서 기술했듯이 추가적인 EST 염기서열의 분석을 수행하게 된다면 장모 및 단모 개체간 EST 염기서열에서 보다 많은 contig 조합을 확인할 수 있을 것이며, 그 과정에서 다량의 SNP site를 추가로 선별할 수 있을 것으로 생각된다. 결론적으로 본 연구에서 확인된 SNP site와 SSR site들은 종자모 형성에 관련된 분자마커 개발에 이용하는 것 뿐만 아니라 당근의 계통 분류 및 여러 형질 관련 분자마커 연구에 활용할 수 있으리라 판단된다.

## 초 록

당근(*Daucus carota* L. var. *sativa*)은 세계적으로 널리 이용되는 작물 중이며, 영양학적으로도 중요한 작물이다. 하지만, 종자 표피세포에서 생성되는 종자모는 발아를 억제하고 흡수를 저해하여 육묘에 어려움을 야기한다. 이러한 어려움을 타파하기 위해 당근 종자는 기계적인 제모작업을 거쳐 상품화 되고 있다. 이 과정에서 생산상의 여러 가지 단점들이 발생하며, 이를 보완하기 위해 단모종자 당근 품종의 육종이 필요하다. 따라서 본 연구는 단모종자 표현형 CT-ATR 615 OP 666-13개체와 장모종자 표현형 CT-ATR 615 OP 671-9개체 및 단모종자 표현형 CT-SMR 616 OP 659-1 개체와 장모종자 표현형 CT-SMR 616 OP 677-14개체 등 두 조합의 종자 cDNA library를 작성 후 EST 염기서열의 비교분석을 통해 당근 종자모 형질관련 연구에 이용하고자 하였다. 첫째로 EST 염기서열의 BlastX 결과를 바탕으로 각각의 EST를 FunCat 기능별 category로 분류하였다. 그 결과 Metabolism category와 protein folding 및 stabilization, protein binding, C-compound binding category에서 2조합 모두 동일한 유의적인 차이를 확인하였다. 두 번째로 EST 염기서열의 GO data를 바탕으로 seed trichome differentiation 및 cellulose biosynthetic process에 관련된 EST를 선별하였다. 이러한 FunCat category에서의 차이점과 GO data 분석을 통해 확인된 후보 EST 들이 당근 종자모 형성에 많은 영향을 미치는 것으로 생각된다. 마지막으로 분석된 개체 별 EST 염기서열을 바탕으로 33개의 SNP site, 741개의 SSR site를 확인하였다. 확인된 SNP 및 SSR site는 당근 종자모 형성에 관련된 분자마커 개발에 이용할 수 있음은 물론 당근의 여러 형질에 대한 연구에 활용 가능할 것으로 기대된다.

추가 주요어 : BlastX, *Daucus carota*, FunCat, SNP, SSR

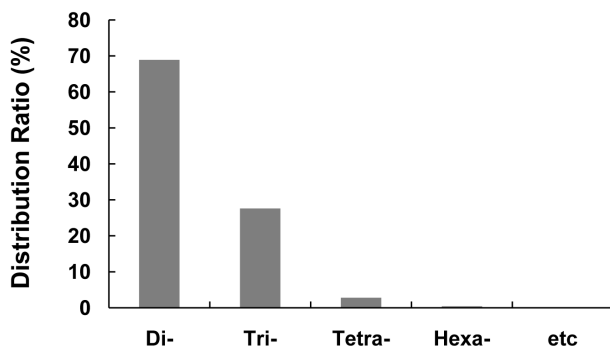


Fig. 2. Frequency distribution of repeat type of SSR sites.

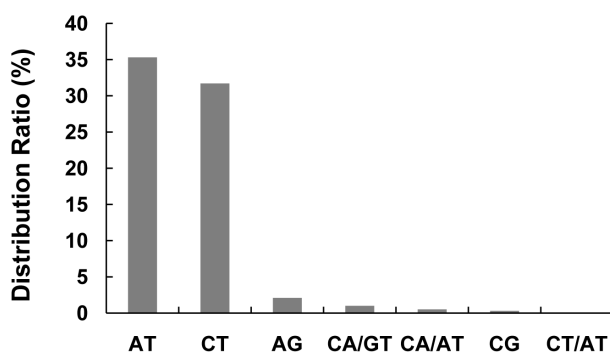


Fig. 3. Di-nucleotide frequency distribution of SSR sites.

## 인용문헌

- Adams, M.D., J.M. Kelley, J.D. Gocayne, M. Dubnick, M.H. Polymeropoulos, H. Xiao, C.R. Merril, A. Wu, B. Olbe, R.F. Moreno, A.R. Kerlavage, W.R. McCombie, and J.C. Venter. 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252:1651-1656.
- Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, and G. Sherlock. 2000. Gene ontology: Tool for the unification of biology. *Nature Genet.* 25:25-29.
- Ayeh, K.O. 2008. Expressed sequence tags (ESTs) and single nucleotide polymorphisms (SNPs): Emerging molecular marker tools for improving agronomic traits in plant biotechnology. *African J. Biotechnol.* 7:331-341.
- Gou, J.Y., L.J. Wang, S.P. Chen, W.L. Hu, and X.Y. Chen. 2007. Gene expression and metabolite profiles of cotton fiber during cell elongation and secondary cell wall synthesis. *Cell Res.* 17:422-434.
- Gupta, P.K., S. Rustgi, S. Sharma, R. Singh, N. Kumar, and H.S. Balyan. 2003. Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Mol. Gen. Genomics* 270:315-323.
- Hwang, E.M., G.D. Oh, E.J. Shim, S.J. Jeon, and Y.D. Park. 2010. Analysis of seed hair formation related genes by EST profiling in carrot (*Daucus carota* var. *sativa*). *Kor. J. Hort. Sci. Technol.* 28:1039-1050.
- Kantety, R.V., M. La Rota, D.E. Matthews, and M.E. Sorrells. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Mol. Biol.* 48:501-510.
- Menon, A.R.S. and Y. Dave. 1989. Micromorphology of hairs and spines on ovary and fruit of *Daucus carota* L. var. *sativa*. *Bot. Mag. Tokyo* 102:503-509.
- Mewes, H.W., C. Amid, R. Arnold, D. Frishman, U. Güldener, G. Mannhaupt, M. Münsterkötter, P. Pagel, N. Strack, V. Stümpflen, J. Warfsmann, and A. Ruepp. 2004. MIPS: Analysis and annotation of proteins from whole genomes. *Nucleic Acids Res.* 32:D41-D44.
- Myakishev, M.V., Y. Khripin, S. Hu, and D.H. Hamer. 2001. High-throughput SNP genotyping by allele-specific PCR with universal energy-transfer-labeled primers. *Genome Res.* 11:163-169.
- Oh, G.D., E.J. Shim, S.J. Jun, and Y.D. Park. 2013. Development of SNP molecular markers related to seed-hair characteristic based on EST sequences in carrot. *Kor. J. Hort. Sci. Technol.* 31:80-88.
- Oh, G.D., E.M. Hwang, E.J. Shim, S.J. Jeon, and Y.D. Park. 2010. EST profiling for seed-hair characteristic and development of EST-SSR and SNP markers in carrot. *Kor. J. Hort. Sci. Technol.* 28:1025-1038.
- Powell, W., G.C. Machray, and J. Provan. 1996. Polymorphism revealed by simple sequence repeats. *Trends Plant Sci.* 1: 215-222.
- Ruepp, A., A. Zollner, D. Maier, K. Albermann, J. Hani, M. Molrejs, I. Tetko, U. Güldener, G. Mannhaupt, M. Münsterkötter, and H.W. Mewes. 2004. The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes. *Nucleic Acids Res.* 32:5539-5545.
- Somers, D.J., R. Kirkpatrick, M. Moniwa, and A. Walsh. 2003. Mining single-nucleotide polymorphisms from hexaploid wheat ESTs. *Genome* 46:431-437.
- Varshney, R.K., R. Sigmund, A. Börner, V. Korzun, N. Stein, M.E. Sorrells, P. Langridge, and A. Graner. 2005. Interspecific transferability and comparative mapping of barley EST-SSR markers in wheat, rye and rice. *Plant Sci.* 168:195-202.
- Wang, Y.C., C.P. Yang, G.F. Liu, J. Jiang, and J.H. Wu. 2006. Generation and analysis of expressed sequence tags from a cDNA library of *Tamarix androssowii*. *Plant Sci.* 170:28-36.
- Wang, Z., J.L. Weber, G. Zhong, and S.D. Tanksley. 1994. Survey of plant short tandem DNA repeats. *Theor. Appl. Genet.* 88:1-6.
- Ye, S., S. Dhillon, X. Ke, A.R. Collins, and I.N.M. Day. 2001. An efficient procedure for genotyping single nucleotide polymorphisms. *Nucleic Acids Res.* 29:e88.
- Yu, J.K., M. La Rota, R.V. Kantety, and M.E. Sorrells. 2004. EST derived SSR markers for comparative mapping in wheat and rice. *Mol. Gen. Genomics* 271:742-751.