P02-013

경북의 힘으로 새로운 대한민국

Development of SSR molecular markers based on genotyping-by-sequencing in *Schisandra chinensis* cultivars

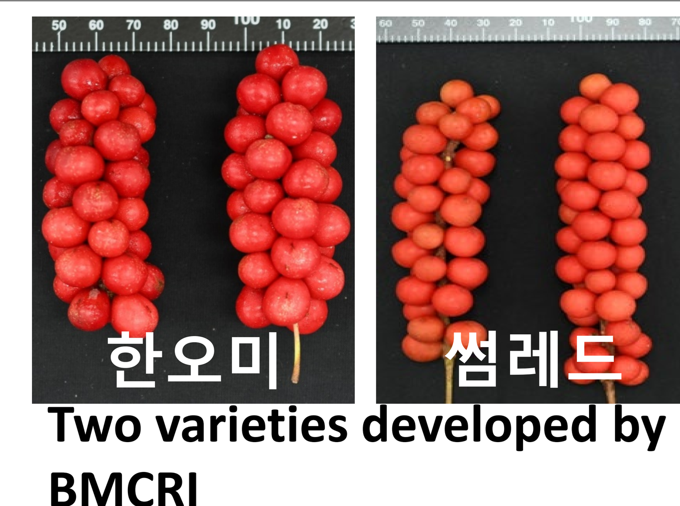
Sang Seok Lee[1)*], Beung Sung Kim[1)], In Kyu Song[1)], and Su Ryun Choi[2)]

[1)]Bonghwa Medicinal Crop Research Institute, Bonghwa-gun Gyeongsangbuk-do 36229, Republic of Korea.

[2)]XENOTYPE CO., LTD. Daejeon, 34912 Republic of Korea

Background

Schisandra chinensis is a climber species called 'Omiza' as fruits for medicinal food ingredient. Because of the agronomic importance of S. chinensis in the northern parts of Gyeongsanbuk-do, two varieties, 'Hanomi' and 'Ssumred' have been recently developed by Bonghwa Medicinal Crop Research Institute(BMCRI). In order to obtain molecular tools that distinguish these varieties from cultivars dispersed in South Korea, we developed single-sequence-repeat(SSR)-based molecular markers assisted with genotyping-by-sequencing technologies.



Two varieties developed by BMCRI

Methods & Results

Using genomic DNA of 'Hanomi' and 'Ssumred', the NGS sequencing data were generated and *De novo* assembly was conducted to detect polymorphic SSR regions. The SSR-targeted primer sets through *in silico* PCR analysis were designed, which expected PCR product sizes resided between 350 and 450 bp. We conducted a series of filtering procedure from the raw polymorphic SSR matrix and finally selected 122 SSR loci. Sequencing data of 81 *S. chinensis* accessions grown in BMCRI were obtained and principal coordinate analysis(PCoA) was followed using the product sizes of 88 filtered SSR markers. Additionally, in order to evaluate discriminating capability of the developed SSR molecular markers, 48 cultivars representing different cultivation regions throughout S. Korea were randomly collected and fluorescence tagging genotypings were implemented.

SSR Analysis

- Short reads (Resequencing data)
- Sequence pre-processing (SOAP *de novo*)
- *De novo* Assembly(MISA)
- SSR detection
- Primer design
- *In-silico* PCR
- SSR size matrix 작성
- 시료간 공통 SSR 선발

SSR-GBS Analysis

- Short reads(GBS data)
- Sequence demultiplexing
- SSR filtering
- Polymorphic SSR among samples
- Principal Coordinates Analysis (PCoA)

Table 1. Statistics of assembled contigs

Sample	Hash length (k-mer)	Num. of contigs	Length(bp) of contigs				
			Total length	MIN	MAX	AVG	N50
썸레드	69	5,659,524	2,024,666,317	200	10,812	357	362
한오미		6,338,400	2,406,509,556	200	10,875	379	390

Table 2. Statistics of polymorphic SSR between samples

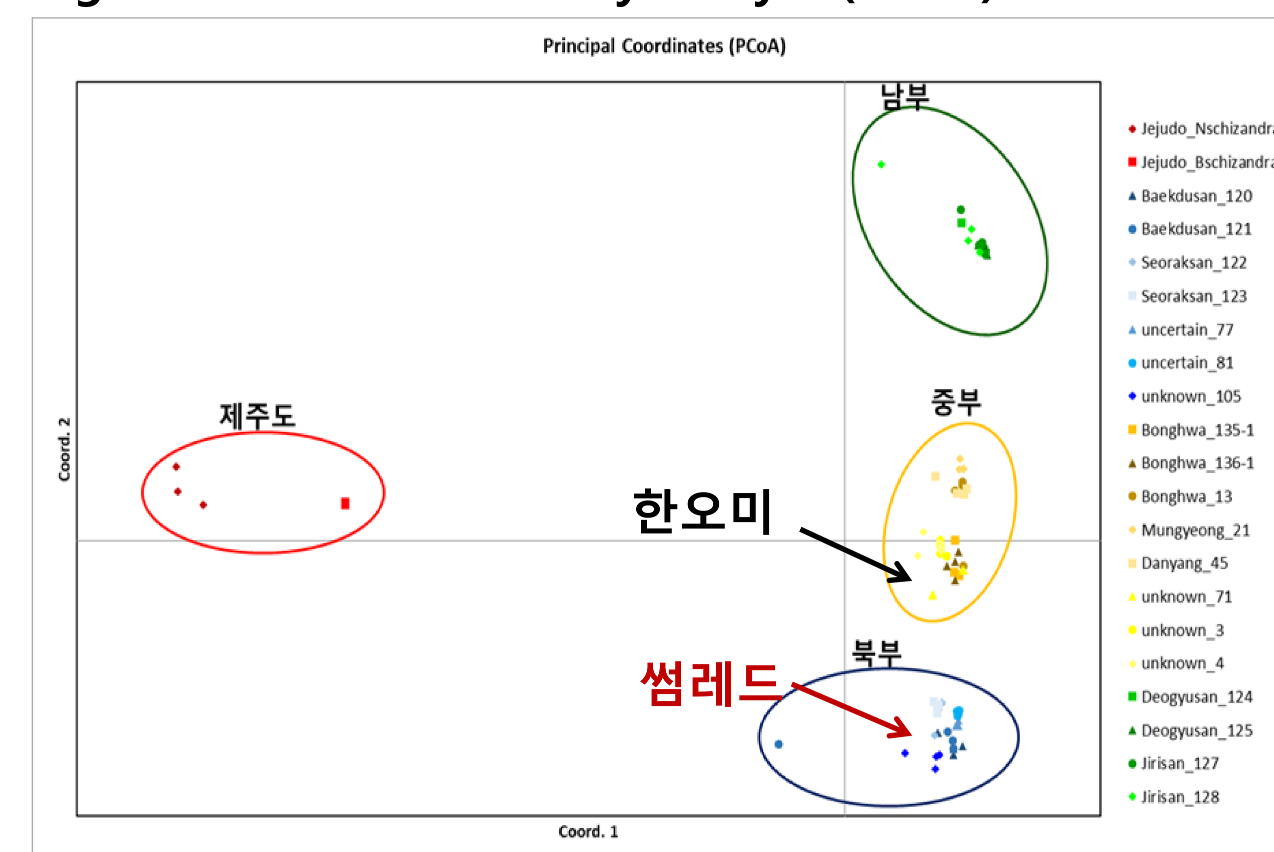
필터 단계	필터 항목	No. of SSRs
1	No. of SSR size matrix 선발*1	424,449
2	비교 샘플간의 motif size polymorphic좌 선발*2	27,664
3	SSR 주변서열 30bp내에 repeat 서열 없는 좌 선발*3	1,636
4	SSR type p2~p6 좌 선발*4	1,206
5	Expected SSR size가 1개인 좌 선발*5	484
6	Product size≤450bp인 좌를 선발*6	474
7	Expected SSR sequence가 motif의 반복이 아닌 경우 제거*7	132
8	Primer 필터*8	122

*1) **No. of SSR size matrix 선발**: *in silico* PCR을 수행하여 자기 자신의 contigs 서열에서 한 번 alignment하는 SSR primer set을 선발함.
*2) **비교 샘플간의 motif size polymorphic 좌 선발**: 비교 샘플간에 SSR motif size 차이를 보이는 SSR좌 개수.
*3) **SSR 주변서열 30bp내에 repeat 서열 없는 좌 선발**: 71-2R 샘플에서 찾은 SSR영역의 앞/뒤 주변서열 30bp 내에 repeat 서열이 존재할 때, 해당 좌를 제거함.
*4) **SSR type p2~p6 좌 선발**: SSR의 motif 길이가 2-6bp인 좌를 선발함.
*5) **Expected SSR size가 1개인 좌 선발**: In silico PCR을 통해 샘플 별로 예측한 SSR 크기가 1개인 좌를 선발함.
*6) **Product size≤450bp인 좌를 선발**: In silico PCR을 통해 샘플 별로 예측한 Product size가 450bp 이하인 좌를 선발함.
*7) **Expected SSR sequence가 motif의 반복이 아닌 경우 제거**: 샘플 간 expected SSR sequence의 차이가 motif 반복 수의 차이가 아닌 InDel등의 영향으로 일어난 경우를 제거함.
*8) **Primer 필터**: 실험과정에서 오류를 줄이기 위해 primer 간 alignment를 통해 서열이 유사한 경우를 제거함

Table 3. Genetic diversity analysis(PCoA)

SSR 마커 ID	71-2R 샘플에서 찾은 SSR 정보			Expected SSR size		Product_size	
		SSR	size	BH 71-2R	BH 105-2R	BH 71-2R	BH 105-2R
Xt-SSR05	71-2R_contig4480133	(TGCAAG)3	18	18	12	342, 350	344
Xt-SSR38	71-2R_contig5592418	(TCGGT)4	20	20	15	373	358, 368
Xt-SSR41	71-2R_contig5654696	(ACCCAA)2	12	12	18	416	422
Xt-SSR63	71-2R_contig6051486	(ACACG)4	20	20	35	369	384
Xt-SSR66	71-2R_contig6076102	(GAGAGT)3	18	18	12	406	400
Xt-SSR96	71-2R_contig6320708	(ATG)7	21	21	24	410	410, 413

Figure 1. Genetic diversity analysis(PCoA)



Conclusion

The PCoA showed that 'Hanomi' and 'Ssumred' clearly belong to different groups consisting of 81 accessions suggesting that the genetic backgrounds may differ. Totally six SSR marker sets were developed and two or three specific SSR marker sets in combination can be used to distinguish 'Hanomi' and 'Ssumred' from cultivars of S. chinensis. These results indicate that GBS-based SSR markers have accuracy and reliability applicable to cross-pollination species since the varieties of S. chinensis are known to have extremely high heterogeneity.

*(Corresponding author) E-mail: lsseok@korea.kr Tel: +82-54-673-8064