

令和5年度「植物品種保護に関する環境を総合的に整備する取組の実施」

(1)のソ トマト種の遺伝子マーカーを活用した品種識別の検証

報告書

(受託者)住 所 千葉県木更津市かずさ鎌足2-6-7

公益財団法人かずさDNA 研究所

(1)のセ トマト種の遺伝子マーカーを活用した品種識別の検証

【目的】

近年の育種技術の進展によって、毎年世界中で多くの新品種が開発されるようになったことに伴い、既存品種も大幅に増加しているため、従来の表現型のみの情報からでは、区別性を確認しなければならないと考えられる対照品種候補の数が増加し、特にヨーロッパの国々においては DUS 試験を効率的に実施することが難しくなっている。UPOV 作業部会においても、遺伝子情報の活用による植物種類ごとの品種の類似性の技術判定の実用化に向けた議論が進められており、EU や中国では品種登録への遺伝子情報の積極的な活用に向けた研究開発プロジェクトが進展している。その一環として、オランダが主体となってトマト種の DNA マーカーセットの妥当性確認試験が進められており、これに参加することで国際的に妥当性が確認された DNA マーカーセットを、関連する情報とともに入手することができ、品種識別にそのまま利用できるだけでなく、形質特異的なマーカーとして利用できる可能性もある。このようなプロジェクトの積み重ねによって、今後、UPOV においても遺伝子情報を品種登録等に積極的に活用する方向へと見直すことも想定され、我が国においても国際的な議論に資する必要な情報を提示できるようにする必要があることから、本試験に参加し、マーカーセットの検証を行う。

【実施方法】

(1) SNP 分析

1. 材料

品種: 国内種苗会社から提供されたトマト 12 品種

反復数: 1

2. 実験手法

提供をうけた種子をセルトレーに播種し、温室で育苗を行った。本葉展開時に1品種あたり 12 個体の幼苗をサンプリングし、これらの葉を等量混合して DNA を抽出した。SNP 解析は NuGEN 社が提供する Allegro Targeted Genotyping 法を用いて実施した。2022 年度の解析においてオランダ側より提供された 297SNPs に対して、プライマー設計と合成を実施した。なお、実験の反復数は1であるが、実験間における結果の安定性を確認するために、同一 DNA に対し3種類のタグを付与し、実験間の反復を確認した(表1)。

DNA の品質を精査したのち、Allegro Targeted Genotyping 法によりライブラリを作成した。。配列解析は NextSeq500 HighOut (150 cycles)を用いて、1ランで実施した。配列データを PRINSEQ (Schmieder and Edwards 2011) 及び fastx_clipper (http://hannonlab.cshl.edu/fastx_toolkit)の各プログラムを使用して品質で精査した。配列のマッピングはトマト参照配列(SL3.0)に対して Bowtie2 プログラム(ver2.5.2)により行い、Samtools を用いて変異を検出した。

表1. 実験手法の概要

手順	実施概要	反復数
DNA 抽出	1 品種あたり 12 個体をバルク	1
ライブラリ作成	1 品種あたり 3 種類のタグを付与	3
配列解析	NextSeq 500 75 塩基長 シングルエンド	1

【結果】

1 品種・1種類のタグあたり、平均で 27,024,869 本のリード配列を取得した(最小 22,212,354 本、最大 32,943,200 本、図1)。参照配列(SL3.0)に対するリード配列のマッピング率は平均で 98.3%(最小 98.1%、最大 98.5%)であり、全てのサンプルで問題なくリード配列が取得できていた。

プライマーを設計した 297 の SNPs のうち、遺伝子型データが取得できたのは 292SNPs だった。検出されなかった SNPs は下記のとおりである。

SL3.0ch04_1177529
SL3.0ch08_61867455
SL3.0ch08_62994811
SL3.0ch08_65375768
SL3.0ch12_1240802

うち、SL3.0ch08_65375768 は 2022 年度にオランダから提供された DNA を用いた解析でも SNPs が検出されなかったことから、プライマーの設計に問題がある可能性が考えられた。一方、他の 5SNPs は 2022 年度の解析では遺伝子型データが取得されており、本年度の供試サンプルが全て SL3.0 配列と同一の塩基を有していたことからバリエーションとして検出されなかった可能性が考えられた。

292 SNPs 上にマップされた SNPs に対して取得したリード数は 1 品種・1反復の平均で 257.28(最小・4.8、最大 274.2)だった(図1)。



図1. 検出された 292SNPs の 1 品種・1反復あたりのリード数。左からリード数が少ない順に 292 SNPs をソートして表示している。

遺伝子型が取得できた 292SNPs のうち、同一品種内における3反復間で遺伝子型データが異なっていたのは 27 だった(表2)。うち、19SNP は解析を行った 12 品種中、1 品種でのみ遺伝子型データが異なっていた。一方、6もしくは 7 品種で反復間の遺伝子型データが異なっていた SNPs がそれぞれ1つずつあった。

そこで、検出した SNPs を以下の条件でフィルタリングし、低品質の SNPs は<データなし>として扱う処理を行った。

フィルタリング条件: QUAL=999, DP20 以上、GQ30 以上

フィルタリングの結果、292SNPs のうち 2SNPs は 12 品種全てで<データなし>となった。また、品種内で異なる遺伝子型が検出された SNP の数が減り、合計で 13SNPs となった(表2)。

表2. 品種間で異なる遺伝子型が検出された SNP の数

遺伝子型データが異なる品種数	フィルタリングなし	フィルタリングあり
0	267	277

1	19	9
2	6	2
3	0	0
4	0	2
5	0	0
6	1	0
7	1	0
データなし	-	2
遺伝子型データが異なった SNP 合計数	27	13

*フィルタリング条件:QUAL=999, DP20 以上、GQ30 以上

異なる遺伝子型を検出した SNP 数を品種ごとに確認したところ、平均で 3.7 の SNP が同一品種内で異なる遺伝子型を検出していた。品質でフィルタリングを行ったあとは異なる遺伝子型を検出した平均 SNP 数は 1.75 となった。12 品種の中では 23Tomato_02 が最も多くの遺伝子型の違いが認められた。23Tomato_08 はフィルタリング有無にかかわらず全ての SNPs で同一の遺伝子型を検出した。品種内における遺伝子型検出の安定性はプローブ設計など SNP マーカーの品質だけでなく、品種内多型が影響している可能性も示唆された。

表3. 各品種において異なる遺伝子型を検出した SNP 数

品種 ID	フィルタリングなし	フィルタリングあり*	
	品種内で異なる遺伝子型を検出した SNP 数	遺伝子型データなし**	
23Tomato_01	2	0	3
23Tomato_02	7	2	4
23Tomato_03	5	0	3
23Tomato_04	3	3	2
23Tomato_05	5	2	2
23Tomato_06	3	3	2
23Tomato_07	6	3	2
23Tomato_08	0	0	2
23Tomato_09	2	1	3
23Tomato_10	4	3	3
23Tomato_11	3	2	3
23Tomato_12	4	2	2
平均	3.7	1.75	2.58

*フィルタリング条件:QUAL=999, DP20 以上、GQ30 以上 senn

**フィルタリング後、全てのサンプルで遺伝子型データがない SNPs は 2 つ

(2) データベースの構築

得られた遺伝子型データを Tomato Genotype Database に格納した(図2、図3)。

URL は下記の通りであるが、現在非公開である。

<https://sites.google.com/kazusa.or.jp/tomatogenotype2023>

問い合わせ先メールアドレスとして plantgarden@kazusa.or.jp を設定した。アクセスのリクエストの連絡があった場合、アクセス許可の発行が可能な場合にのみ該当するメールアドレスでのアクセスを可とする



図2. データベースのトップページ

国内12品種の遺伝子型データ

トップページの画像もしくは文字をクリックすることで
12品種×3反復の遺伝子型データにアクセスできる

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB													
1	CHROM	POS	REF	ALT	23Tomato01_1	23Tomato01_2	23Tomato01_3	23Tomato02_1	23Tomato02_2	23Tomato02_3	23Tomato03_1	23Tomato03_2	23Tomato03_3	23Tomato04_1	23Tomato04_2	23Tomato04_3	23Tomato05_1	23Tomato05_2	23Tomato05_3	23Tomato06_1	23Tomato06_2	23Tomato06_3	23Tomato07_1	23Tomato07_2	23Tomato07_3	23Tomato08_1	23Tomato08_2	23Tomato08_3	23Tomato09_1	23Tomato09_2	23Tomato09_3	23Tomato10_1	23Tomato10_2	23Tomato10_3	23Tomato11_1	23Tomato11_2	23Tomato11_3	23Tomato12_1	23Tomato12_2	23Tomato12_3	
2	SL3.0ch01	5122021	A	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
3	SL3.0ch01	79617233	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
4	SL3.0ch01	79710376	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
5	SL3.0ch01	79672391	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
6	SL3.0ch01	79935275	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
7	SL3.0ch01	80000763	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
8	SL3.0ch01	83934629	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T		
9	SL3.0ch01	83946577	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
10	SL3.0ch01	83952645	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
11	SL3.0ch01	83959274	T	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
12	SL3.0ch01	83999269	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
13	SL3.0ch01	86416566	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
14	SL3.0ch01	89692751	A	G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	A/G	
15	SL3.0ch01	89704074	C	T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T	C/T
16	SL3.0ch01	96109247	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
17	SL3.0ch01	97333593	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	

図3.
トップページから
の遺伝子型データ
アクセス

【考察】

解析を実施した 297SNPs のうち、遺伝子型データが得られたのは品質によるフィルタリング無しで 292、フィルタリングありで 290 だった。また、同一品種・同一 DNA を用いた3反復実験において品種内で異なる遺伝子型を検出した SNPs はフィルタリング前で 27、フィルタリング後で 13 だった。フィルタリング条件は QUAL=999 (Samtools で検出される QUAL 値の最高値)、マップされたリード数 (DP) が 20 以上、GQ (Phred Score) 30 以上であり、バリエントコールにおいて標準的に用いられる条件である。にもかかわらず、フィルタリング後の結果で品種内で異なる遺伝子型が検出されたことから、高精度な遺伝子型データを取得する場合にはより厳しい条件でフィルタリングを実施する必要性を示唆した。

引用文献

- Danecek et al. (2011) *Bioinformatics* 27:2156-2158.
Kitashiba et al. (2014) *DNA Res.* 21(5): 481–490.
Kumar et al. (2008) *Brief Bioinform* 9:299-306.
Langmead and Salzberg (2012) *Nat Methods* 9:357-359.
Schmieder and Edwards (2011) *Bioinformatics* 27:863-864.
Shirasawa et al. (2016) *DNA Res* 23:145-153.
Van Berloo, R. (1999) *Journal of Heredity* 90:328–329.