

L_TSF1
ARATH
46 bp
 エlement①

CTP2
ARATH
228 bp
 エlement②

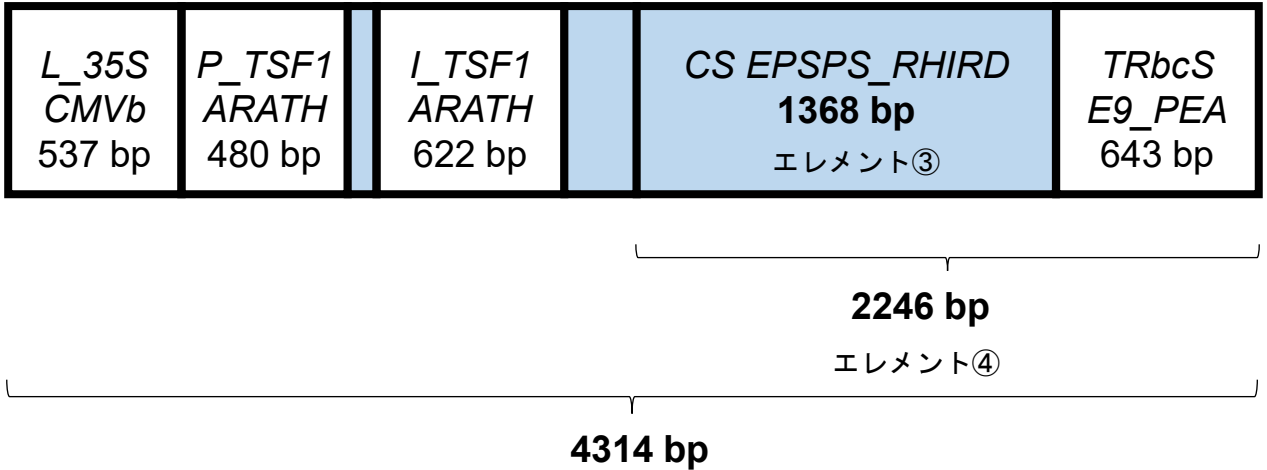


図1. 遺伝子組換えダイズRRS2に挿入されている外来遺伝子配列の概要

遺伝子組換えダイズRRS2系統は1番染色体に約4314 bpの外来遺伝子が挿入されている。
 本研究ではこの全長および各Element①~④について検討を行った。

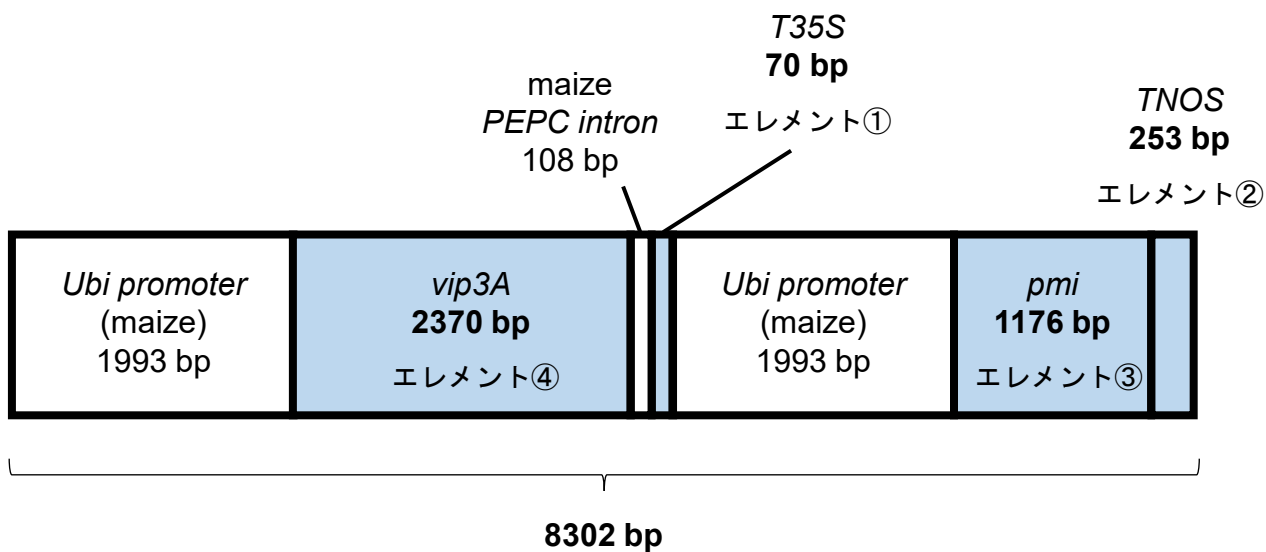


図2. 遺伝子組換えトウモロコシMIR162に挿入されている外来遺伝子配列の概要

遺伝子組換えトウモロコシMIR162系統は5番染色体に約8302 bpの外来遺伝子が挿入されている。本研究ではこの全長および各エレメント①～④について検討を行った。



図3. RRS2に挿入された外来遺伝子にマッピングしたショートリードシーケンスデータのIGV画像

各エレメント配列または全長にマッピングしたリードを再度全長配列にマッピングした。

Target (bp)	50×	40×	30×	25×	20×	15×	10×	5×
46	35.8 ± 4.4	28.6 ± 4.3	19.8 ± 5.1	16.5 ± 4.3	12.8 ± 3.9	9.3 ± 3.1	6.3 ± 2.9	2.2 ± 2.1
228	89.6 ± 2.4	71.8 ± 8.1	52.3 ± 8.2	42.3 ± 6.7	33.9 ± 7.2	25.4 ± 5.0	14.1 ± 3.3	7.4 ± 2.2
1,368	390.0 ± 11.1	316.2 ± 15.7	235.3 ± 17.8	203.0 ± 13.8	161.2 ± 10.6	123.8 ± 9.3	81.1 ± 9.6	41.3 ± 9.4
2,246	620.4 ± 14.7	502.3 ± 20.4	374.2 ± 18.5	251.4 ± 5.8	251.4 ± 5.8	192.1 ± 5.9	127.7 ± 11.9	65.3 ± 11.9
4,314	1,226 ± 22.8	1,002 ± 31.1	742.3 ± 28.1	623.9 ± 29.7	495.3 ± 25.3	373.1 ± 23.4	241.3 ± 19.2	119.0 ± 14.0

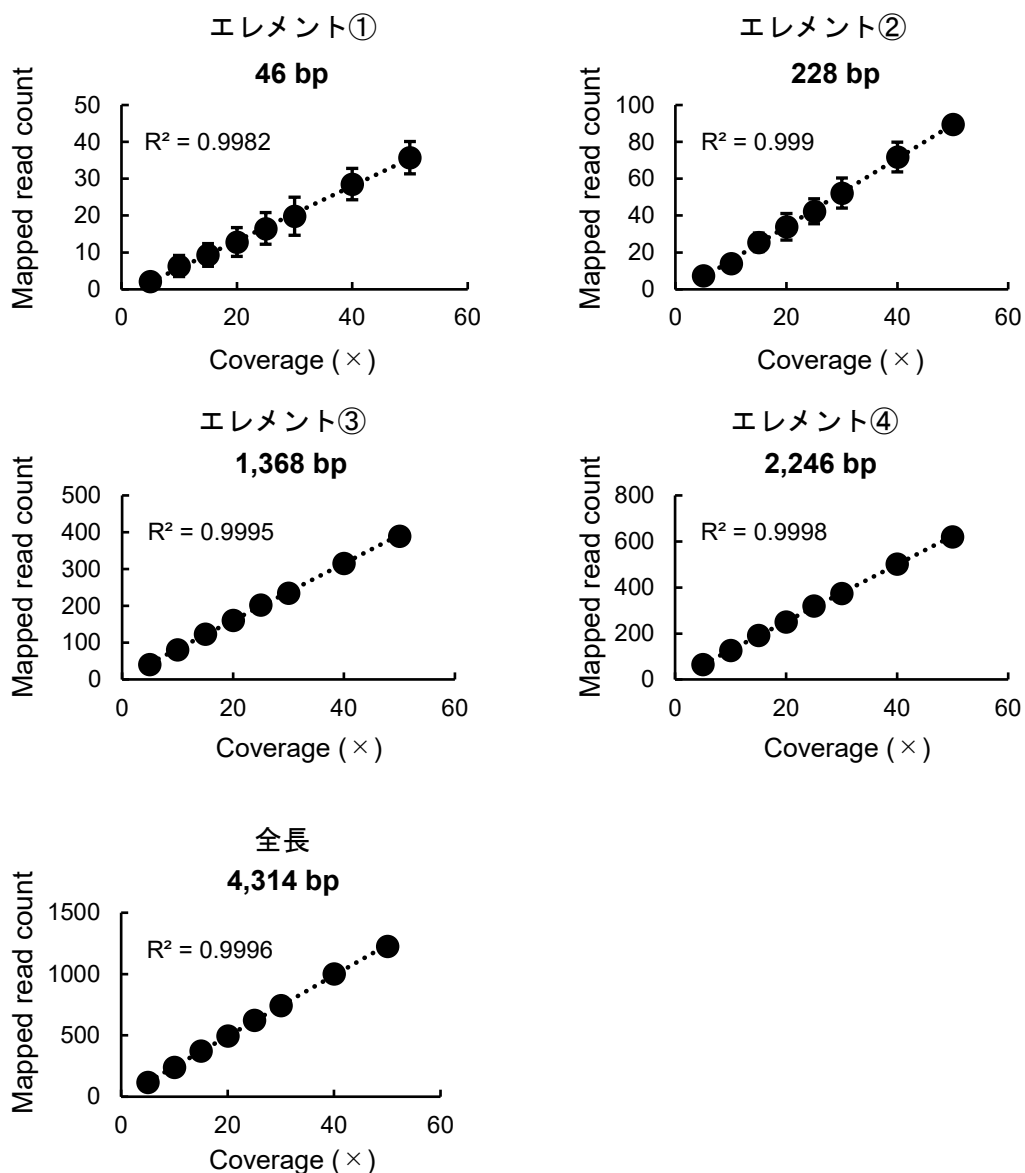


図4. Seqkitランダムサンプリングの併行精度 (RRS2)

各シークエンスカバレッジになるよう12回試行でランダムサンプリングを行い、リードデータを全長および各エレメントにマッピングした。マッピングしたリード数、その標準偏差、シークエンスカバレッジに対する相関を示す。

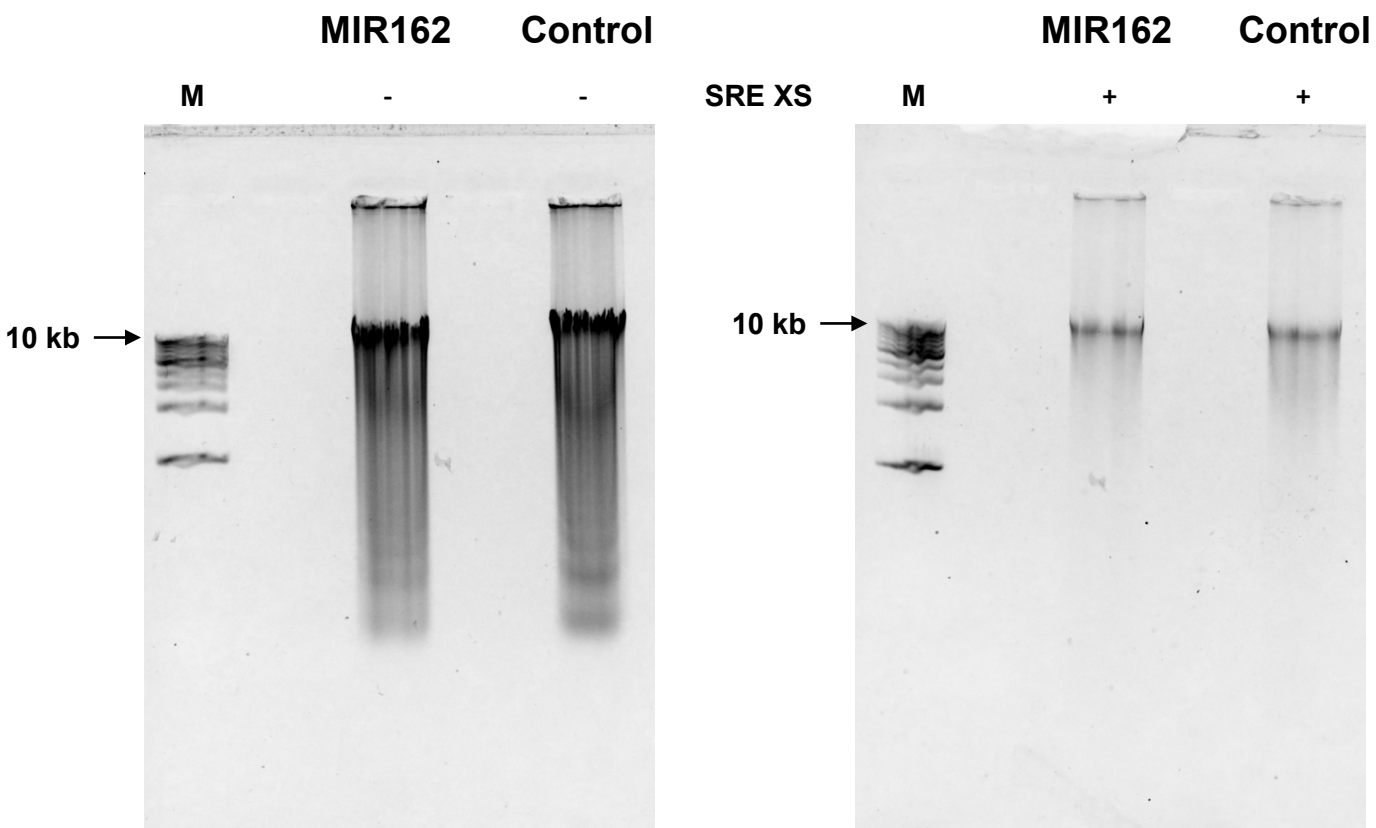


図5. トウモロコシから抽出したゲノムDNAの1%アガロース電気泳動図

コントロールトウモロコシおよびMIR162から抽出したゲノムDNAの電気泳動図を示す。

-はSRE XS (Circulomics社) 処理前、+は処理後を表す。

M: 10 kb Ladder Marker

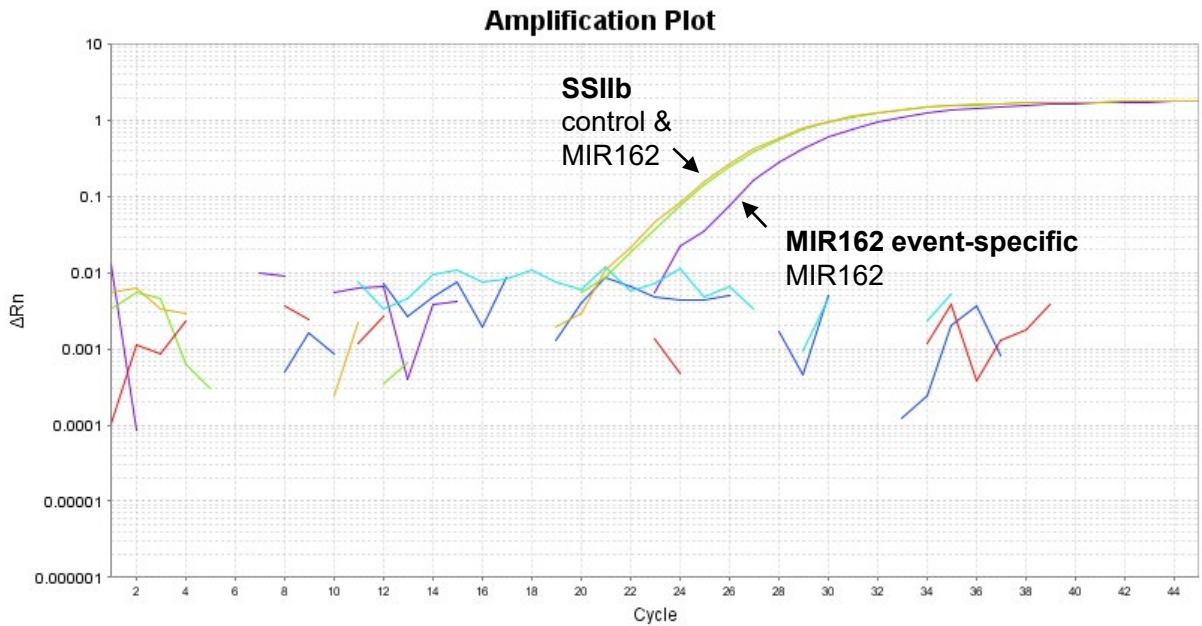
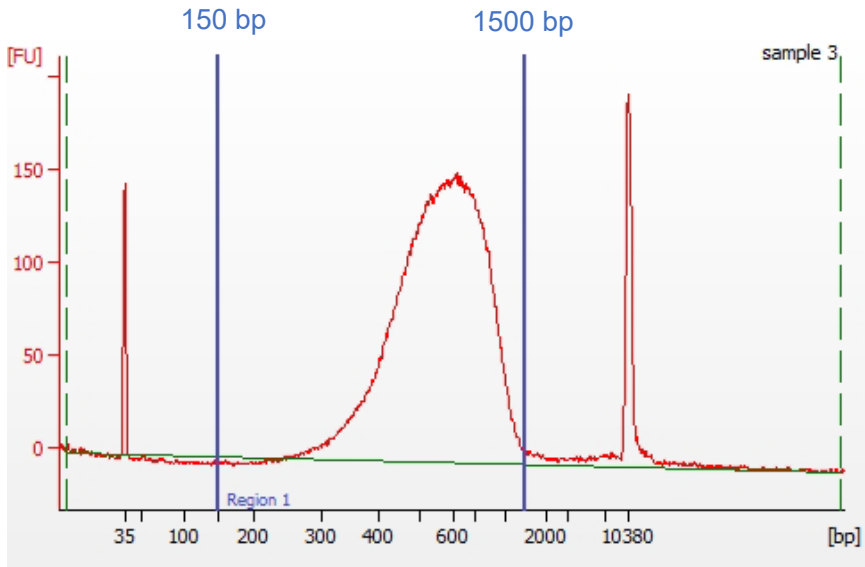


図6. SSIbおよびMIR162系統特異的検知法結果

トウモロコシ内在性遺伝子SSIb検知法ではコントロールおよびMIR162の増幅が確認されたが、MIR162系統特異的検知法ではコントロールは増幅せず、MIR162のみで増幅が確認された。

Control



MIR162

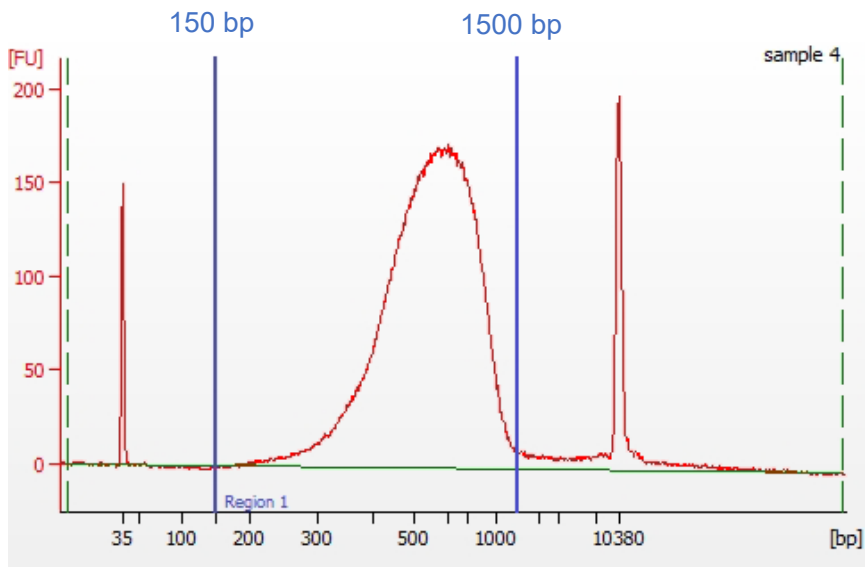


図7. コントロールトウモロコシおよびMIR162におけるWGS用ライブラリのQC結果

コントロールトウモロコシ（上図）およびMIR162（下図）のWGSライブラリはAgilent Bioanalyzer 2100による品質確認を行い、サイズやアダプターオリゴのコンタミネーションがないことを確認した後、シーケンスを行った。

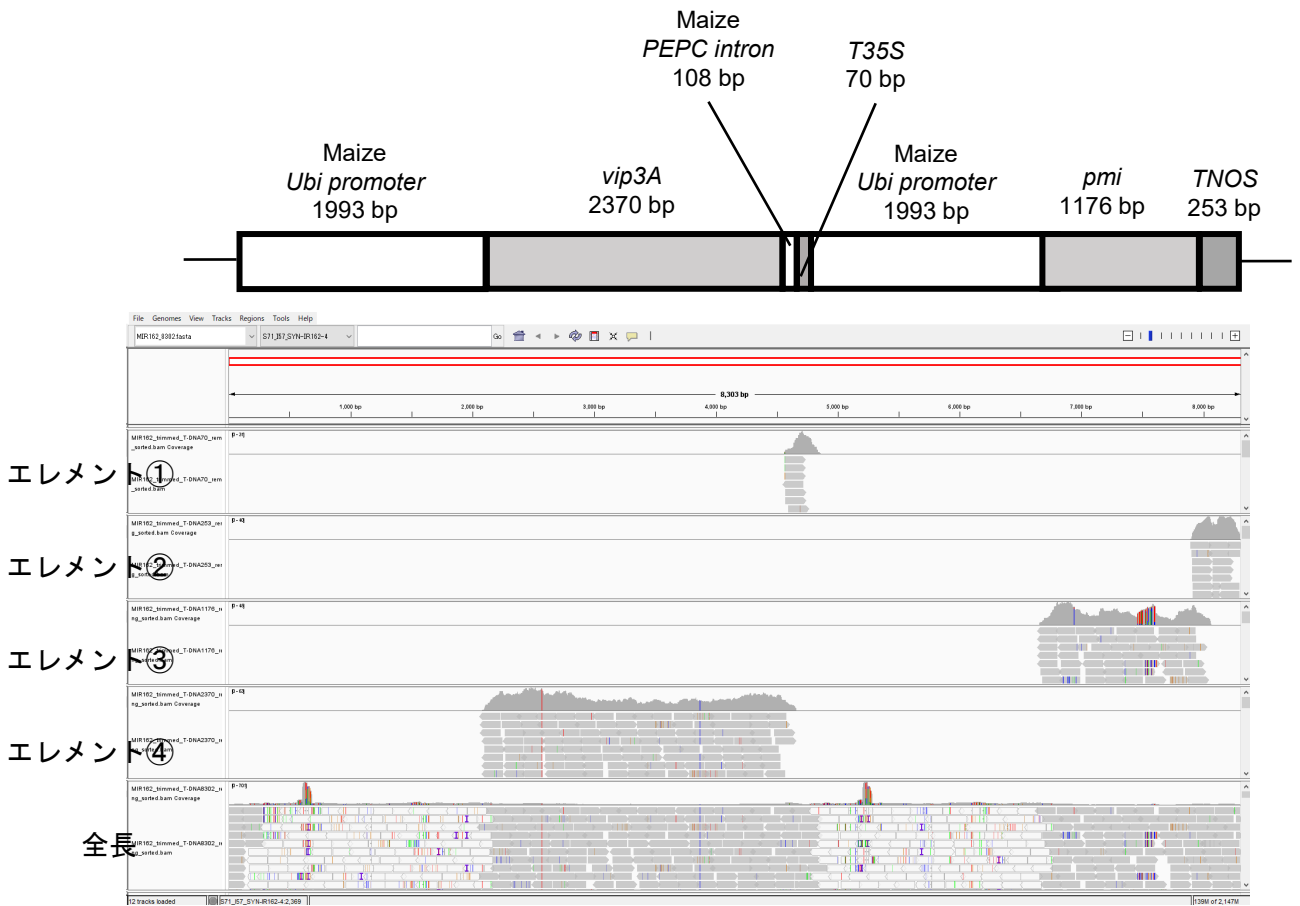


図8. MIR162に挿入された外来遺伝子にマッピングしたショートリードシーケンスデータのIGV画像

各エレメント配列または全長配列にマッピングしたリードを再度全長配列にマッピングした。白抜きになっているリードはMAPQが0であることを示す。

Target (bp)	50×	40×	30×	25×	20×	15×	10×	5×
70	28.3±1.5	22.2±3.4	16.6±2.8	13.5±3.1	10.2±3.6	7.7±3.1	4.8±2.6	2.9±1.9
253	74.1±4.8	60.8±6.9	46.6±5.8	39.3±5.1	30.7±4.6	22.5±5.2	14.7±4.7	6.8±3.0
1,176	220.8±7.4	182.2±14.8	139.3±14.9	118.7±10.4	93.4±11.1	69.8±8.2	46.9±8.3	24.8±4.5
2,370	542.7±16.5	430.8±29.1	327.9±25.0	281.1±11.0	222.7±13.6	163.5±14.1	106.2±16.1	49.8±16.8
8,302	3,915.3±112.0	3,144.2±209.2	2,378.1±209.4	1,996.6±113.3	1,567.7±117.8	1,176.3±104.3	787.3±99.0	394.3±97.9

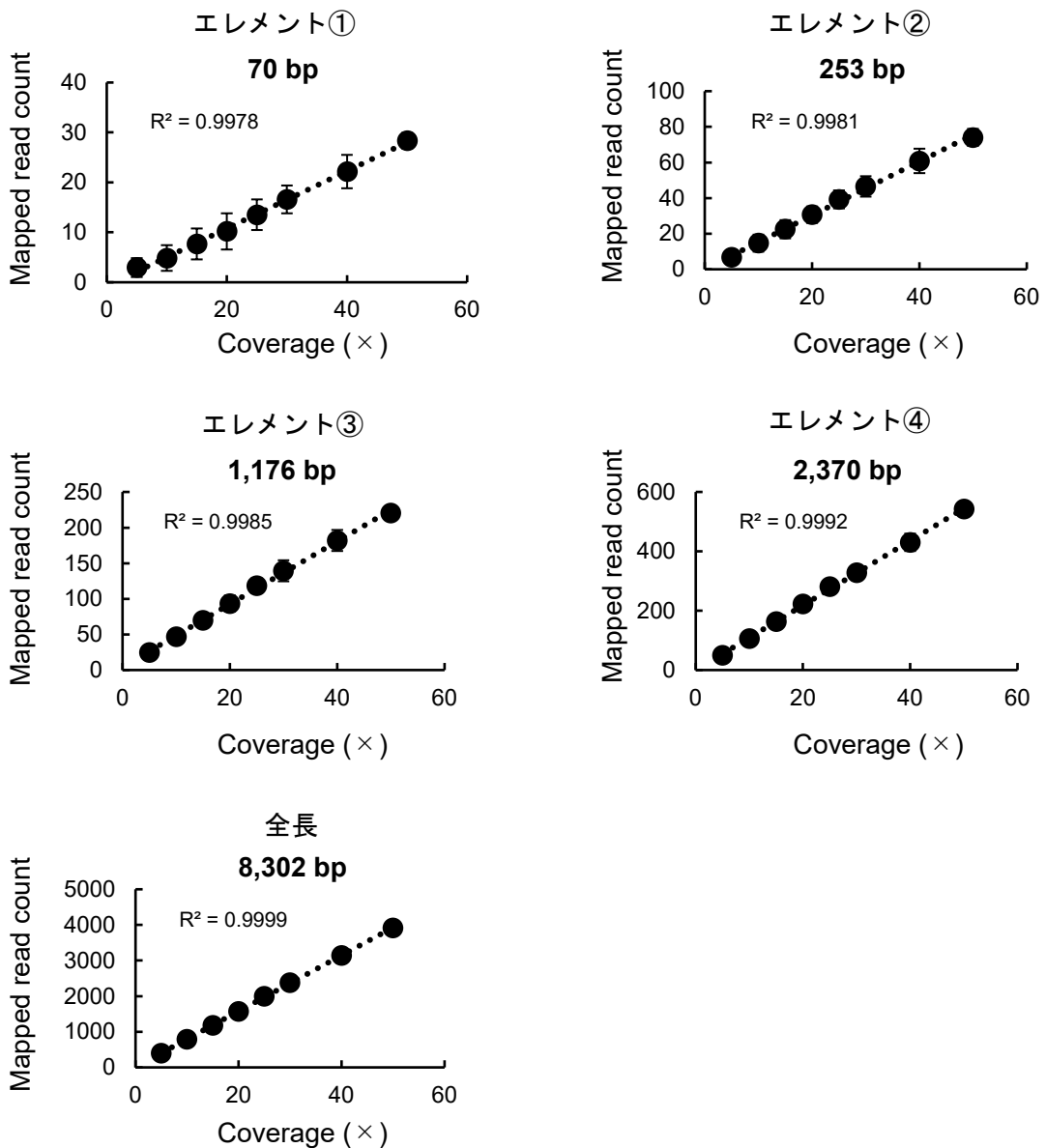


図9. Seqkitランダムサンプリングの併行精度 (MIR162)

各シークエンスカバレッジになるよう12回試行でランダムサンプリングを行い、リードデータを全長および各エレメントにマッピングした。マッピングしたリード数、その標準偏差、シークエンスカバレッジに対する相関を示す。

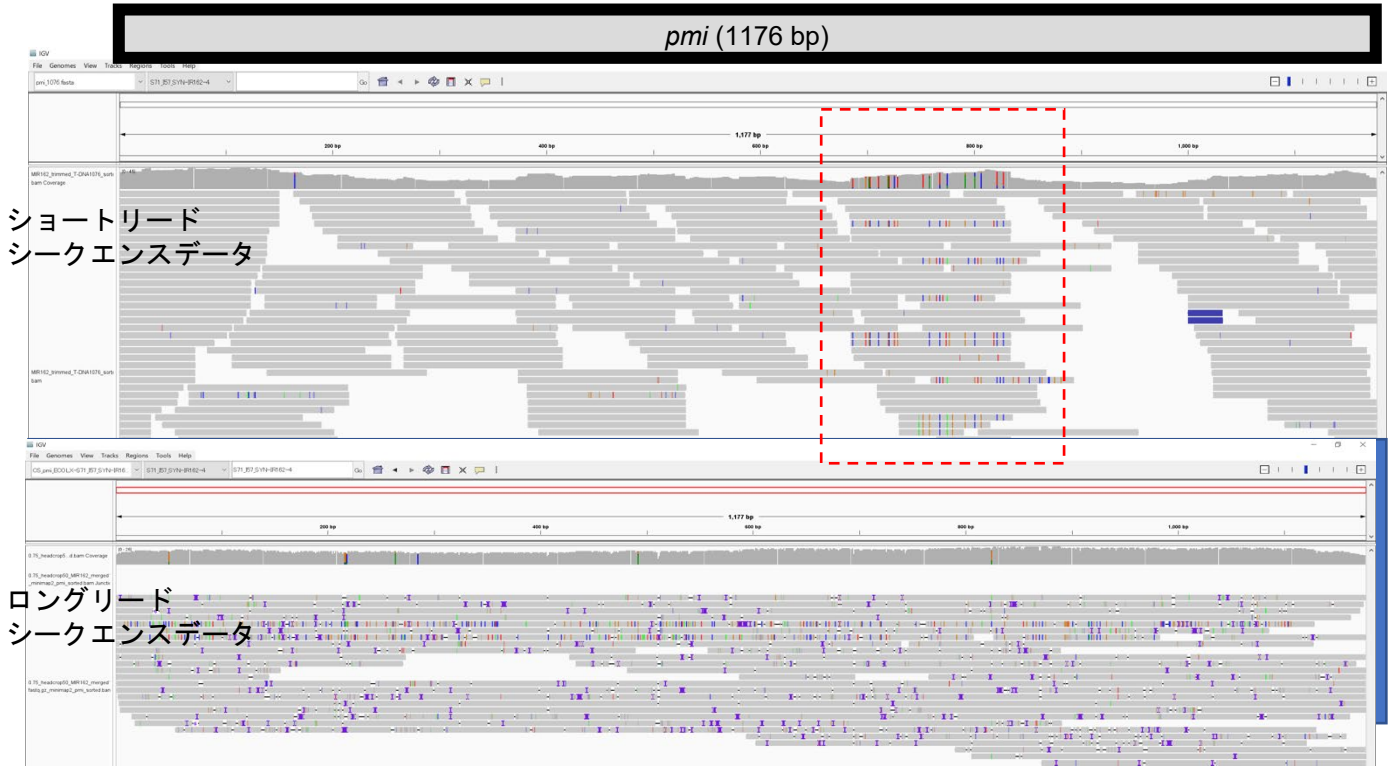


図10. *pmi*領域に各WGSリードデータをマッピングしたIGV画像 (MIR162)

ショートリードシーケンスデータ（上段）およびロングリードシーケンスデータ（下段）について、*pmi*領域にマッピングした結果を示す。赤点線枠はショートリードシーケンスデータで変異が蓄積していた領域。

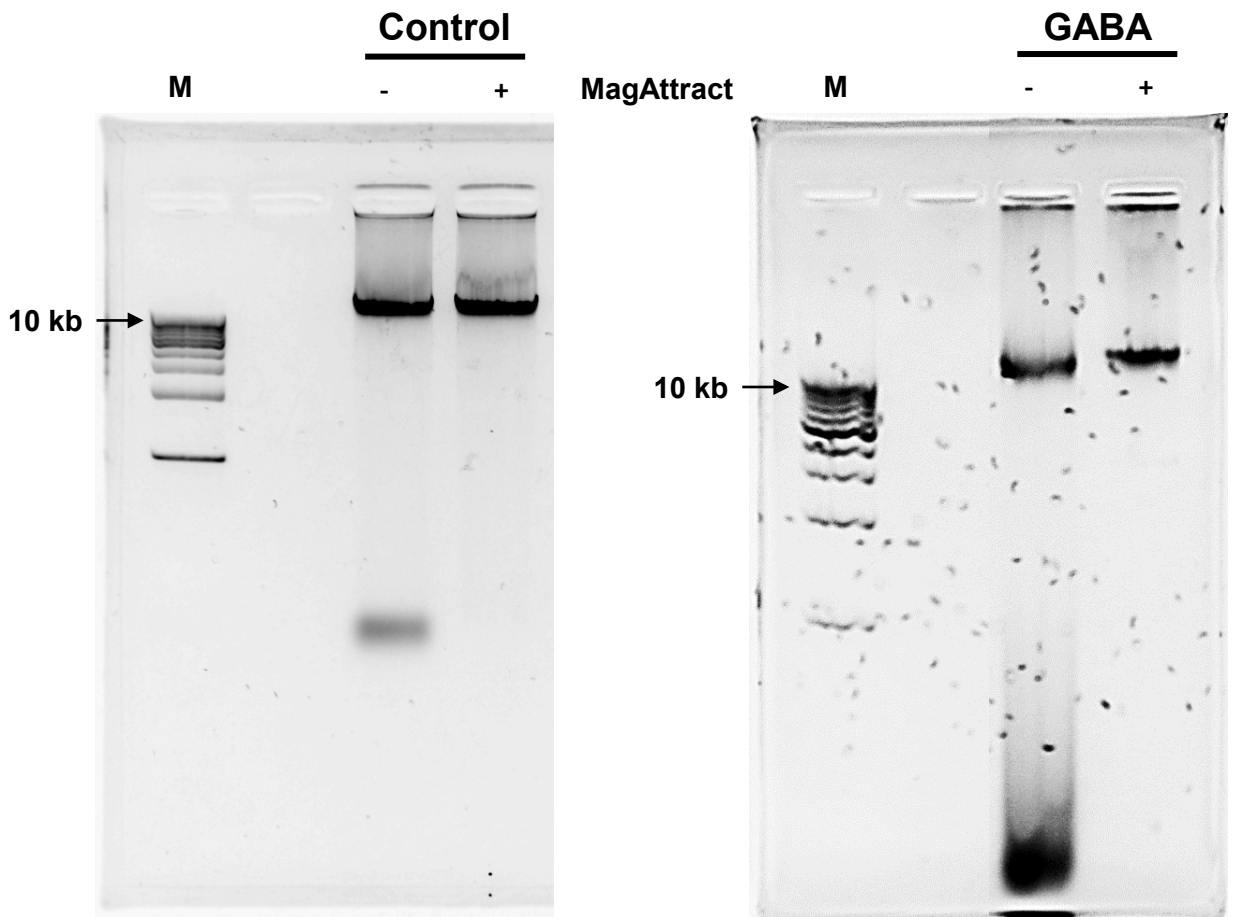


図11. トマト青果から抽出したゲノムDNAの1%アガロース電気泳動図

左図は野生型シリアンルージュ、右図はゲノム編集GABAトマトから抽出したゲノムDNAを表す。
 -はMagAttract (QIAGEN社) 処理前、+は処理後を表す。

M: 10 kb Ladder Marker

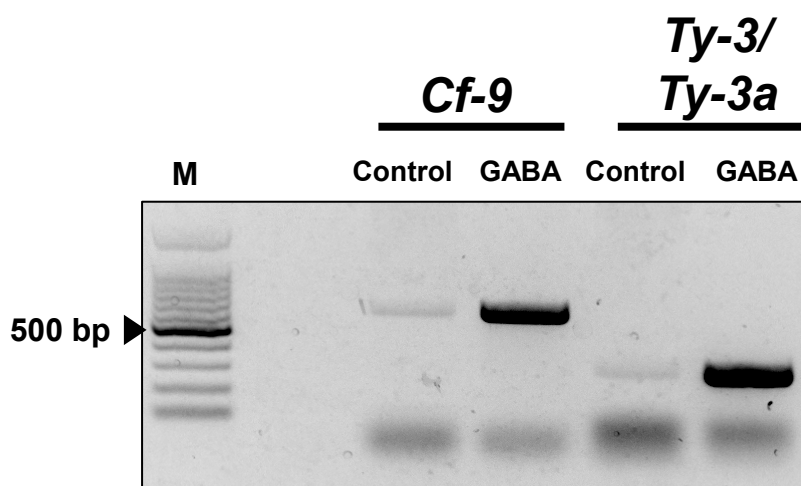


図12. DNAマーカーを用いたトマトの系統確認

DNAマーカーを用いて野生型シシリアンルージュの系統確認を行った。

Cf-9マーカーは陽性であれば650 bp、陰性であればバンドなし。

Ty-3/Ty-3aマーカーはTy-3陽性であれば450 bp、Ty-3a陽性であれば630 bp、陰性であれば320 bpにバンドが確認される。

M: 100 bp Ladder Marker

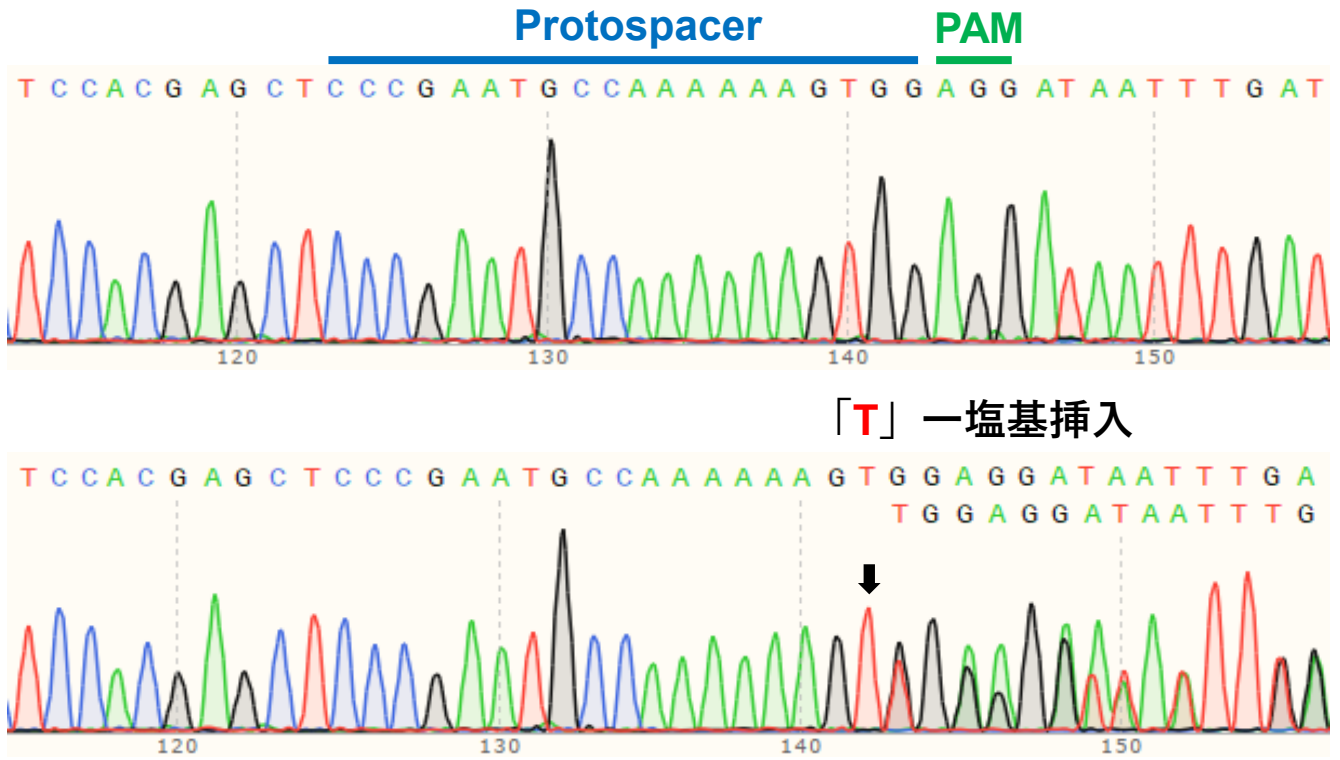
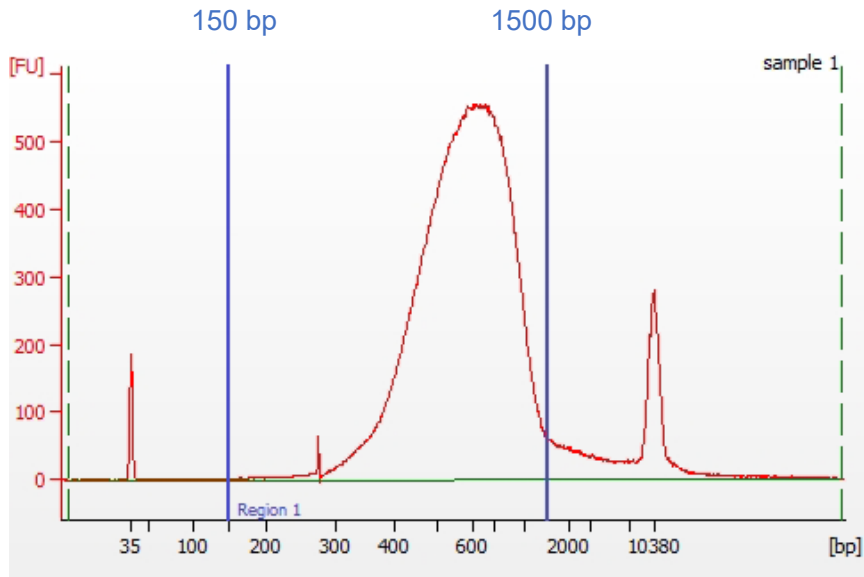


図13. 野生型トマトおよびゲノム編集GABAトマトにおけるGAD3配列の比較

野生型トマト（上段）およびゲノム編集GABAトマト（下段）におけるGAD3配列を比較した。図中のPAM（Protospacer Adjacent Motif）およびProtospacerは、シーケンス結果より推定されるCas9の結合配列、GAD3-gRNA配列を表す。

Control



GABA

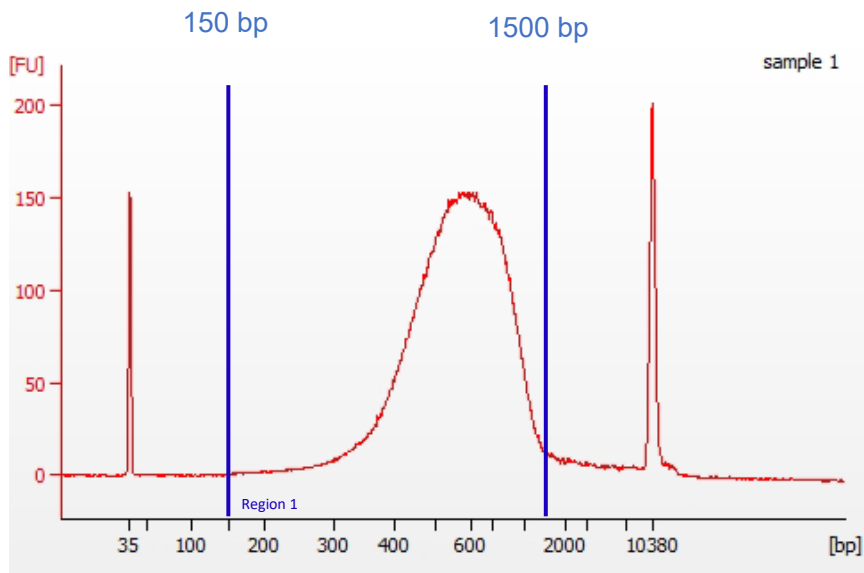
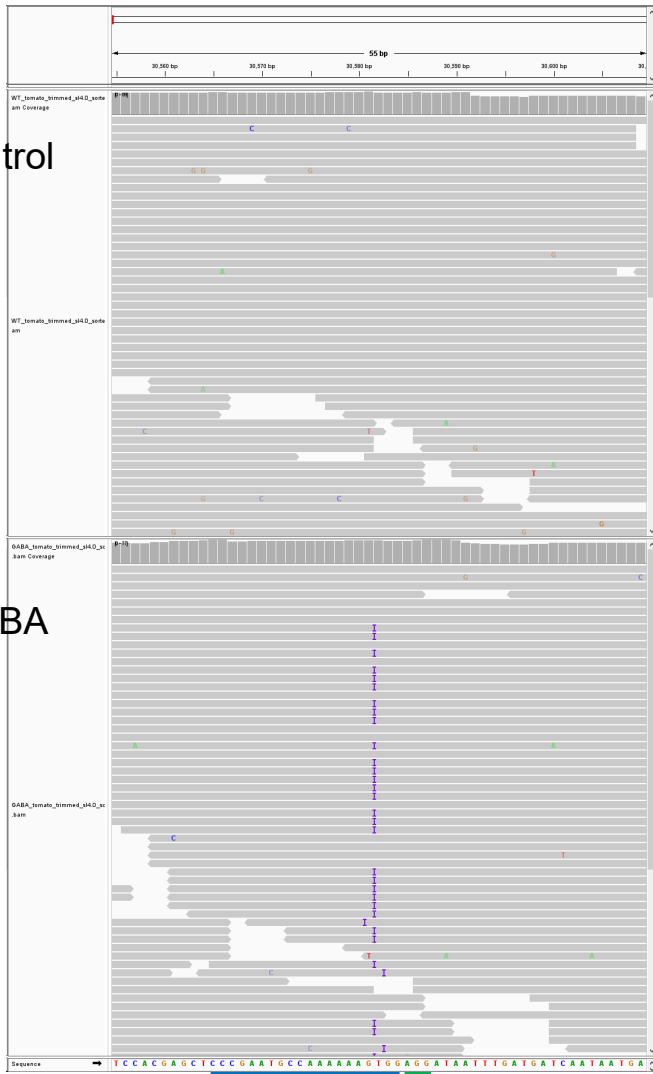


図14. 野生型トマトおよびGABAトマトのWGS用ライブラリのQC結果

野生型トマト（左図）およびGABAトマト（右図）のWGSライブラリはAgilent Bioanalyzer 2100による品質確認を行い、サイズやアダプターオリゴのコンタミネーションがないことを確認した後、シーケンスを実施した。

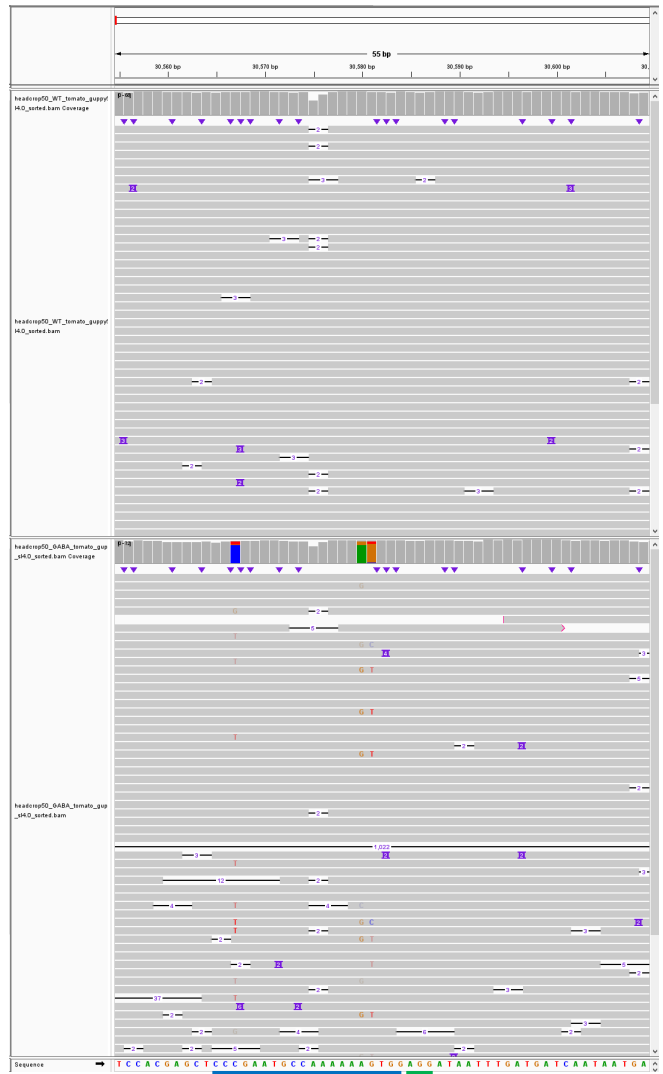
Control

GABA



Protospacer PAM

ショートリード
シークエンスデータ



Protospacer PAM

ロングリード
シークエンスデータ

図15. ショートリードシークエンス、ロングリードシークエンスにおけるGAD3標的部位のマッピングリードの比較

全ゲノムシークエンスデータにおけるGAD3 標的部位のIGV画像。左図がショートリードシークエンスリードデータ、右図がロングリードシークエンスリードデータ。いずれも上段が野生型トマト由来のリードを表し、下段がGABAトマト由来のリードを表す。図中のPAM（Protospacer Adjacent Motif）およびProtospacerは、シークエンス結果より推定されるCas9の結合配列およびGAD3-gRNA配列を表し、紫のマークはInsertionを表す。

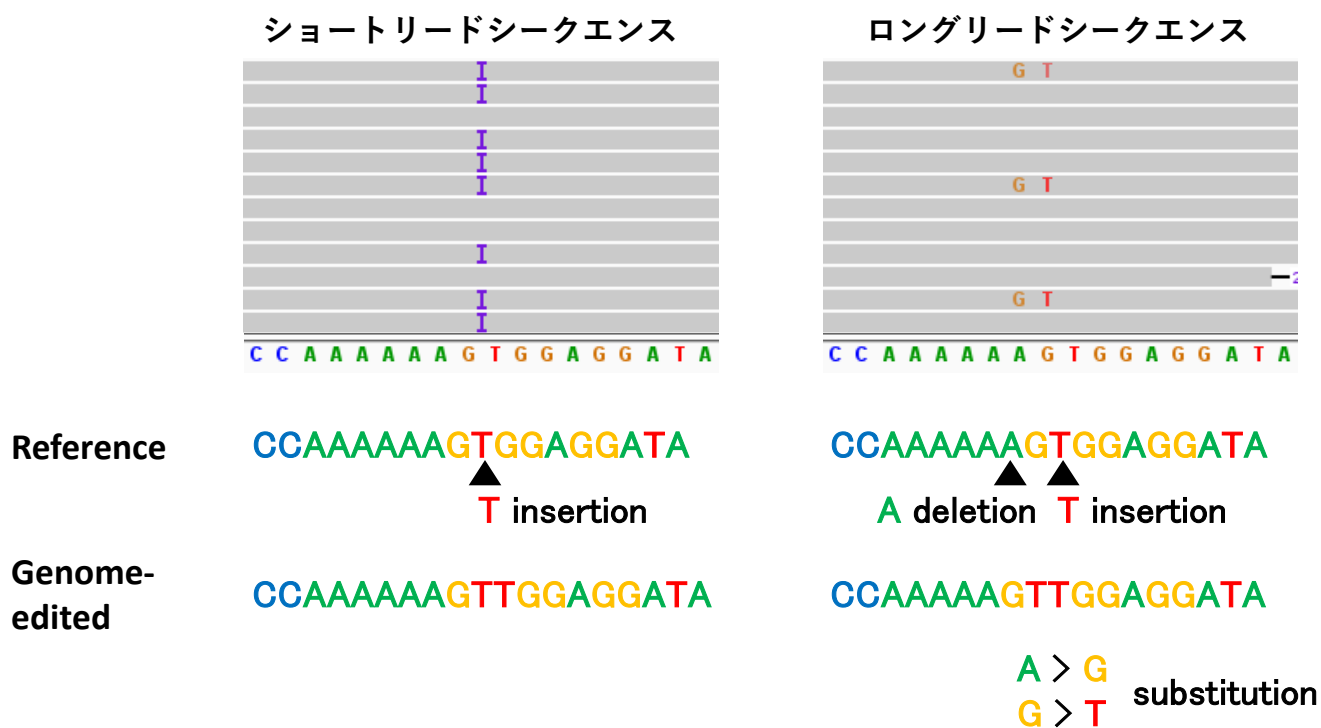


図16. ゲノム編集トマトにおけるGAD3標的部位のショートリードシーケンスとロングリードシーケンスのマッピングリードの比較

ショートリードシーケンス（左図）ではゲノム編集によるTの一塩基挿入を正しく「insertion」と表示するが、ロングリードシーケンス（右図）ではAのホモポリマー配列のエラーによりAが一塩基「deletion」し、結果A > GおよびG > Tの「substitution」と表示されている。

表1. RRS2およびコントロールサイズのWGSにより得られたリードデータの統計値と解析に使用したソフトウェア

		ショートリードシーケンス		ロングリードシーケンス	
		Control	RRS2	Control	RRS2
ミ ニ コ ト 後	Total reads	474.8M	454.7M	9.89 M	3.9 M
	Total bases	70.0 G	66.9 G	26.4 G	52.1 G
	N50	150 b	150 b	4.26 kb	25.0 kb
	Coverage	63.6 x	60.8 x	23.5 x	47.2 x
Trimming	Trim Galore (ver. 0.6.7)		NanoFilt (ver. 2.8.0)		
Mapping	BWA-MEM (ver. 0.7.17-r1188)		Minimap2 (ver. 2.21-r1071)		
Assembly	Velvet (ver. 1.2.10) SPAdes (ver. 3.15.5)		Flye (ver. 2.9)		

表2. RRS2に挿入された外来遺伝子全長および各エレメントにマッピングしたリードの統計値とアセンブリ結果

Target (bp)	ショートリードシーケンス						ロングリードシーケンス				
	Mapped reads	Mapped bases (b)	Average length (b)	>Q30 (%)	Contigs	Accuracy (%)	Mapped reads	Mapped bases (b)	Contigs	Assembly length(b)	Accuracy (%)
46	46	6,882	149.6	94.55	1	100	0	0	-	-	-
228	106	15,654	147.7	91.54	1	100	41	1,380,543	1	90,947	100
1,368	475	70,498	148.4	93	1	100	44	1,408,707	1	121,913	100
2,246	762	112,971	148.3	94.03	1	100	49	1,498,178	1	121,893	99.9
4,314	1,481	218,718	147.4	93.85	1	100	58	1,662,854	1	91,213	99.9

-: 未検討

表3. Velvetによるアセンブリ結果 (RRS2)

	Target (bp)	50×	40×	30×	25×	20×	15×	10×	5×
	46	12/12	12/12	12/12	12/12	12/12	12/12	10/12	5/12
	228	12/12	12/12	12/12	12/12	12/12	12/12	9/11	9/12
Contig 複数で全体を カバー	1368	12/12	11/12	12/12	11/12	11/12	12/12	9/12	1/12
	2246	12/12	12/12	12/12	12/12	11/11	11/11	10/12	1/12
	4314	12/12	12/12	11/12	10/12	10/12	8/12	7/12	0/12
.....									
	46	12/12	12/12	12/12	12/12	12/12	12/12	10/12	5/12
	228	12/12	12/12	12/12	12/12	12/12	12/12	9/11	9/12
Contig 1本で全体を カバー	1368	12/12	10/12	10/12	8/12	6/12	3/12	4/12	0/12
	2246	10/12	10/12	9/12	8/12	3/11	1/11	2/12	0/12
	4314	11/12	9/12	7/12	5/12	2/12	0/12	1/12	0/12

表4. SPAdesによるアセンブリ結果 (RRS2)

	Target (bp)	50×	40×	30×	25×	20×	15×	10×	5×
	46	12/12	12/12	9/12	6/12	3/12	1/12	0/12	0/12
	228	12/12	12/12	12/12	11/12	9/12	5/12	0/11	0/12
Contig 複数で全体を カバー	1368	12/12	12/12	12/12	12/12	12/12	12/12	4/12	0/12
	2246	12/12	12/12	12/12	12/12	11/11	11/11	6/12	0/12
	4314	12/12	12/12	12/12	12/12	12/12	12/12	5/12	0/12
.....									
	46	12/12	12/12	9/12	6/12	3/12	1/12	0/12	0/12
	228	12/12	12/12	12/12	11/12	9/12	5/12	0/11	0/12
Contig 1本で全体を カバー	1368	12/12	12/12	12/12	12/12	12/12	11/12	4/12	0/12
	2246	12/12	12/12	12/12	12/12	11/11	10/11	4/12	0/12
	4314	12/12	12/12	12/12	12/12	11/12	10/12	4/12	0/12

表5. ロングリードシーケンスデータのランダムサンプリング、およびマッピングリード数とFlyeによるアセンブリ結果 (RRS2)

	Target (bp)	40×	30×	25×
Mapped read count (Average ± SD)	46	0	0	-
	228	31.3 ± 2.2	24.2 ± 2.5	-
	1368	33.8 ± 2.7	26.2 ± 3.1	-
	2246	37.3 ± 2.7	28.9 ± 2.7	24.1 ± 2.3
	4314	45.1 ± 2.4	34.8 ± 2.6	
リード長 N50 (kb) (Average ± SD)	46			-
	228	42.0 ± 2.9	41.6 ± 3.2	-
	1368	41.2 ± 2.4	40.7 ± 3.5	-
	2246	40.9 ± 2.3	40.2 ± 3.6	41.3 ± 3.1
	4314	39.3 ± 2.1	38.6 ± 3.0	
残存配列の再現 (Contig1本で全体をカバー)	46	-	-	-
	228	10/12	11/12	-
	1368	10/12	9/12	-
	2246	10/12	9/12	8/12
	4314	11/12	10/12	-

-: 未検討

表6. MIR162およびコントロールトウモロコシのWGSにより得られたリードデータの統計値と解析に使用したソフトウェア

	ショートリードシーケンス		ロングリードシーケンス		
	Control	MIR162	Control	MIR162	
ト リ ミ ン グ 後	Total reads	405.9 M	447.1 M	34.5 M	57.2 M
	Total bases	118.6 G	131.7 G	72.7 G	97.7 G
	N50	150 b	150 b	3.93 kb	2.78 kb
	Coverage	49.4 x	54.9 x	30.3 x	40.7 x
Trimming	Trim Galore (ver. 0.6.7)		NanoFilt (ver. 2.8.0)		
Mapping	BWA-MEM (ver. 0.7.17-r1188)		Minimap2 (ver. 2.21-r1071)		
Assembly	Velvet (ver. 1.2.10) SPAdes (ver. 3.15.5)		Flye (ver. 2.9)		

表7. MIR162に挿入された外来遺伝子全長および各エレメントにマッピングしたリードの統計値とアセンブリ結果

Target (bp)	ショートリードシーケンス						ロングリードシーケンス				
	Mapped reads	Mapped bases (b)	Average length (b)	>Q30 (%)	Contigs	Accuracy (%)	Mapped reads	Mapped bases (b)	Assembly length (b)	Contigs	Accuracy (%)
70	31	4,553	146.9	94.33	1	100	15	58,373	11,834	1	100
253	83	12,206	147.1	90.35	1	100	24	110,259	14,919	1	100
1,176	243	35,650	146.7	89.73	1	100	42	143,444	18,379	1	100
2,370	594	87,886	148	88.57	3	-	51	129,286	4,593	1	100
8,302	4,316	636,362	147.4	89.62	11	-	274	1,001,196	-	3	-

-: 未検討

表8. Velvetによるアセンブリ結果 (MIR162)

	Target (bp)	50×	40×	30×	25×	20×	15×	10×	5×
Contig 複数で全体を カバー	70	11/11	12/12	12/12	12/12	12/12	11/12	5/11	3/12
	253	-	-	-	-	-	-	-	-
	1176	-	-	-	-	-	-	-	-
	2370	-	-	-	-	-	-	-	-
	8302	-	-	-	-	-	-	-	-
Contig 1本で全体を カバー	70	11/11	12/12	12/12	12/12	12/12	11/12	5/11	3/12
	253	-	-	-	-	-	-	-	-
	1176	-	-	-	-	-	-	-	-
	2370	-	-	-	-	-	-	-	-
	8302	-	-	-	-	-	-	-	-

-: 未検討

表9. SPAdesによるアセンブリ結果 (MIR162)

	Target (bp)	50×	40×	30×	25×	20×	15×	10×	5×
Contig 複数で全体を カバー	70	11/11	11/12	7/12	2/12	1/12	0/12	0/11	0/12
	253	12/12	12/12	11/11	12/12	9/12	3/12	1/12	0/12
	1176	12/12	12/12	12/12	12/12	10/12	1/12	0/12	0/12
	2370	12/12	12/12	12/12	11/11	12/12	12/12	1/11	0/12
	8302	0/12	0/12	0/12	0/12	0/12	0/12	0/12	0/11
Contig 1本で全体を カバー	70	11/11	11/12	7/12	2/12	1/12	0/12	0/11	0/12
	253	12/12	12/12	11/11	12/12	9/12	3/12	1/12	0/12
	1176	0/12	3/12	7/12	7/12	8/12	1/12	0/12	0/12
	2370	12/12	12/12	12/12	11/11	10/12	9/12	0/11	0/12
	8302	0/12	0/12	0/12	0/12	0/12	0/12	0/12	0/11

表10. ロングリードシーケンスデータのランダムサンプリング、およびマッピングリード数とFlyeによるアセンブリ結果 (MIR162)

	Target (bp)	30×	20×
Mapped read count (Average ± SD)	70	11.8 ± 1.5	7.3 ± 1.8
	253	18.5 ± 1.2	12.2 ± 2.9
	1176	31.9 ± 1.4	21.2 ± 2.9
	2370	37.6 ± 2.7	25.3 ± 3.4
リード長 N50 (kb) (Average ± SD)	70	8.1 ± 0.4	11.9 ± 5.6
	253	8.4 ± 1.3	8.0 ± 1.5
	1176	6.5 ± 1.3	6.4 ± 2.0
	2370	4.6 ± 1.3	4.6 ± 1.8
残存配列の再現 (Contig 1本で全体をカバー)	70	2/12	1/12
	253	11/12	4/12
	1176	12/12	5/12
	2370	12/12	9/12

表11. GABAトマトおよび野生型トマトのWGSにより得られたリードデータの統計値と解析に使用したソフトウェア

		ショートリードシーケンス		ロングリードシーケンス	
		Control	GABA	Control	GABA
ト リ ミ ン グ 後	Total reads	191.4 M	258.0 M	3 M	5.78 M
	Total bases	55.9 G	75.3 G	58.5 G	52.4 G
	N50	150 b	150 b	35.5 kb	13.8 kb
	Coverage	58.9 x	79.2x	61.6 x	55.1 x
Trimming	Trim Galore (ver. 0.6.7)		NanoFilt (ver. 2.8.0)		
Mapping	BWA-MEM (ver. 0.7.17-r1188)		Minimap2 (ver. 2.21-r1071)		