

ゲノム解読を基礎とする高等植物 ミトコンドリアゲノムの包括的研究

平成 26 年 4 月 24 日受付

寺 地 徹*

要 旨

本研究は、ダイコンとその近縁種、パンコムギとその近縁種という異なる 2 つの分類群の、それぞれ 10 種類程度のタイプの違う細胞質を対象に、ミトコンドリアゲノムの解読を進め、植物ミトコンドリアゲノムの構造と機能、ならびに変異と進化に関する包括的な研究を行い、このゲノムについての新しい概念を創出しようとするものである。そのため次世代シーケンサーを活用して、より多くの植物種のミトコンドリアゲノムを解読するとともに、パイオインフォマティクスの手法を積極的に取り入れ、全塩基配列ベースでゲノム構造の比較解析を行おうとする。今年度は、その第一歩として、ダイコンに雄性不稔を引き起こすことが知られている *Brassica maurorum* の mtDNA を次世代シーケンサーにより解読し、約 66.9Mb の塩基配列情報を得た。得られた 242 個のコンティグの中から、ミトコンドリアゲノムに由来する 32 個のコンティグを選抜し、PCR とダイレクトシーケンシングにより、隣接すると思われるコンティグ間の接続を調べた。その結果、32 個のコンティグすべての連続性が証明され、236,281bp からなる *B. maurorum* のマスターサークル分子が予想された。

キーワード：ミトコンドリアゲノム、次世代シーケンサー、*Brassica maurorum*、ダイコン、細胞質雄性不稔 (CMS)

1. はじめに

高等植物のミトコンドリアに固有の DNA (mtDNA) が存在することが、Suyama & Bonner (1966) によって初めて示されて以来、植物ミトコンドリアゲノムに関する研究は、基礎と応用の両面から数多く行われてきた。そのなかで 1980 年代の前半には、植物から純度の高い mtDNA を精製する方法が確立され、クローニングやサザンハイブリダイゼーションなどの分子生物学的実験により、トウモロコシ (Lonsdale *et al.* 1984) やブロッコリー (Palmer & Herbon, 1986) で、他の植物に先駆けてゲノム全体の制限酵素地図が作られた。現在でも多くの人々に信じられている、「高等植物のミトコンドリアゲノムはマスターサークルとサブサークルから構成され、それらの DNA 分子が動的平衡状態で存在する」という植物ミトコンドリアゲノムの一般的な概念は、これら初期の仕事で形成された

* 京都産業大学総合生命科学部

(図 1)。一方 mtDNA の全塩基配列決定 (ゲノム解読) は, 1997 年にシロイヌナズナで初めて行われ (Unseld *et al.*, 1997), 遺伝子の構成や RNA 編集を受ける塩基がすべて明らかにされた。それ以降, 現在までに 65 種あまりの植物でミトコンドリアゲノムの完全解読が行われている。注目すべきは, ゲノム解読研究の半数以上が, ここ数年以内に終了していることであり, いわゆる次世代シーケンサーの普及とともに, ゲノム解読が加速度的に進んでいる現状が窺われる。我々の研究室では, これまで, 他の研究者とともにパンコムギ (*Triticum aestivum*) の mtDNA の全塩基配列を決定し, ミトコンドリアゲノムの構造を初めて明らかにした (Ogihara *et al.*, 2005)。また, 最近では, 次世代シーケンサーを用いて, ダイコン (*Raphanus sativus*) の正常型とオグラ型の 2 種類のミトコンドリアゲノムの解読を終えている (Tanaka *et al.*, 2012)。今回, 本学の新規研究挑戦支援プログラムに採択された研究課題「ゲノム解読を基礎とする高等植物ミトコンドリアゲノムの包括的研究」は, この 2 つの研究をもとに, 内容をさらに深化・発展させようとするものである。

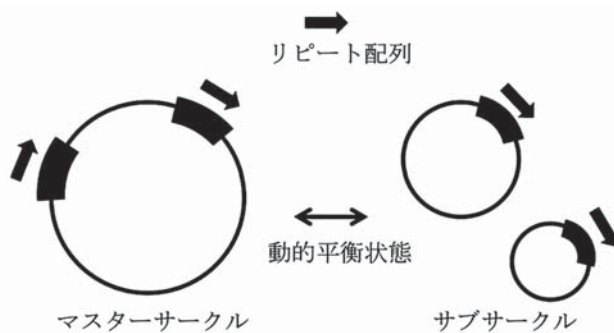


図 1. 高等植物のミトコンドリアゲノムの概念

より具体的に, 本研究では, ダイコンとその近縁種, パンコムギとその近縁種という異なる 2 つの分類群の, それぞれ 10 種類程度のタイプの違う細胞質を対象に, ミトコンドリアゲノムの解読を進め, 植物ミトコンドリアゲノムの構造と機能, ならびに変異と進化に関する包括的な研究を行い, このゲノムについての新しい概念を創出しようとするものである。そのため次世代シーケンサーを活用して, より多くの植物種のミトコンドリアゲノムを解読するとともに, バイオインフォマティクスの手法を積極的に取り入れ, 全塩基配列ベースでゲノム構造の比較解析を行おうとする。また, 分子コーミングの手法を応用してゲノム構造に関する物理的な情報を得て, ゲノム解読の結果と照合する予定である。最後に, ミトコンドリアゲノム上に多数見つかる新規オープンリーディングフレーム (ORF) にも着目し, 大腸菌とタバコ葉緑体を用いた発現解析を通じて, これら ORF の機能を明らかにすることをめざしている。

本研究課題は, 平成 25 年度, 科研費基盤研究 (B) へ申請し不採択になったものであるが, 規定により, 本学で実施されている前述のプログラムに採択され, 小規模ながら実験を 1 年前倒しで開始することができた。本報告では, 平成 25 年度に実施したいいくつかの実験の中から, ダイコンに雄性

不稔を引き起こすことで注目されている *Brassica maurorum* のミトコンドリアゲノム解読の現状を報告したい。

2. 材料及び方法

1) 植物材料

ミトコンドリアのゲノム解読には *B. maurorum* の細胞質を持つ、細胞質置換ダイコンを用いた (図 2)。この系統は、宇都宮大学農学部の方博士から譲渡されたもので、連続戻し交雑によって *B. maurorum* の細胞質を、ダイコンの一品種、“Four Season Leaf”へ導入したものである。宇都宮大学で生育させた複数の植物を京都産業大学へ移送し、以下の実験に用いた。



図 2. *B. maurorum* の細胞質を持つ細胞質置換ダイコン

2) 実験方法

植物の葉、茎から、定法により無傷ミトコンドリアを単離した。単離ミトコンドリアのライゼートから EtBr/塩化セシウム密度勾配遠心法により、mtDNA を精製した。mtDNA を REPLI-g Mini Kit (Qiagen) を用いて増幅した。増幅済みの精製 mtDNA を北海道システムサイエンス社に送付し、そこで品質チェックを受けたあと、ライブラリーの調製及び次世代シーケンサー、ロシュ 454 シーケンシングシステム (Roche), による塩基配列決定を委託した。

3. 結果と考察

植物材料を入手する都合から、細胞質置換ダイコンのミトコンドリアの単離は日を分けて 2 度行ったが、いずれの標品からも純度の高い mtDNA を精製することが可能であった (図 3)。しかし、DNA 量は標品 1 が $1.2 \mu\text{g}$ 、標品 2 が $2.5 \mu\text{g}$ であり、いずれも次世代シーケンシング用のライブラリー調製には量が不足すると思われた。そこで標品 1 を鋳型に、REPLI-g Mini Kit を用いて、mtDNA の全体を *in vitro* 増幅し、ライブラリーの作製に十分な量の DNA を確保した。

ロシュ 454XL+ 反応系による一次データでは、全部で 142,755 リードが得られ、リードの長さの平均は 584bp, 合計では約 83.3Mb (83,339,282bp) となった。ここからアッセンブルに適さない薄いリードを除去し、残りを再アッセンブルしたところ、約 67.5Mb (67,469,221) の塩基配列情報を含む 114,697 リードが残った。そのうち、何らかの形でアラインメントされたものは、113,685 リードであり、その合計は約 66.9Mb (66,857,855bp) であった。ここからコンティグとしてアッセンブルされたものは、112,938 リードであり、これらのリードは長さが 500bp 以上の 169 個のコンティグとして分配された。この 169 個のコンティグで、合計 528,396bp の塩基配列情報を提供する。ちなみに、コンティグの長さの平均は約 3kb (3,126bp) であり、もっとも大きいコンティグは約 31kb (30,844bp), N50 の値は 7,374bp であった。なお、100bp 以上の長さを持つものまで含めると、コンティグ数は 242 個であり、これらのコンティグで 548,609bp をカバーした。今、*B. maurorum* のミトコンドリアゲノム

の大きさを 300kb と仮定し、アラインメントされた約 66.9Mb のリードがすべて mtDNA に由来すると仮定すると、ゲノムのカバレッジは 200 倍以上となり、今回の実験で得られた情報量は、*B. maurorum* のミトコンドリアゲノム全体を解読するのに十分であると思われた。

それぞれのコンティグを構成するリード数や Blast 検索の結果から、32 個の mtDNA に由来するコンティグが選別された (表 1)。各コンティグのデータに付加されている、隣接コンティグの情報をもとに、適当なプライマーを設計し、当該コンティグと予想された隣接コンティグが本当に隣接しているか、コンティグ間の DNA 断片を実際に PCR 増幅した。増幅産物の塩基配列をサンガー法によるダイレクトシーケンシングで決定し、ミトコンドリアゲノム上におけるコンティグの連続性を確認した (表 2)。その結果を総合して、コンティグの並びの予想図を画いた (図 4)。この図で、例えばコンティグ No. 0065 から、No. 0002 を経て、間にある 28 個のコンティグを経由し、No. 0009 から No. 0027 を経て、再度 No. 0065 に戻るといふ、一筆書きの経路を辿ることができる。現時点で、まだ各コンティグの塩基配列の内容を詳細に解析できてはいないが、Blast 解析の結果をみる限りでは、これらのコンティグは植物ミトコンドリアに存在する既知の遺伝子をすべて含んでいる。したがって、合計 236,281bp からなる、このコンティグの並びが、*B. maurorum* のミトコンドリアゲノムのマスターサークルに相当するものと思われる。

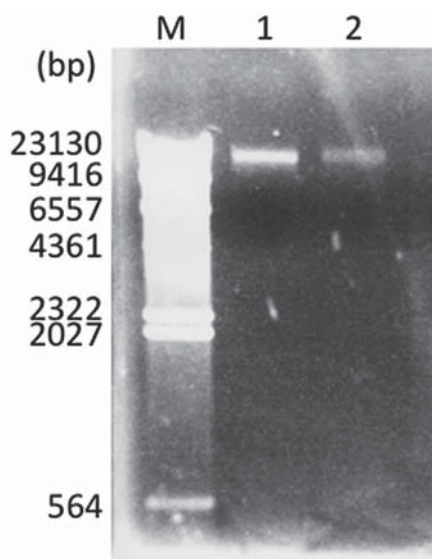


図 3. *B. maurorum* から精製した mtDNA の電気泳動像

M: 分子サイズマーカー (λ /HindIII), 1 及び 2: 精製 mtDNA (ミトコンドリアの単離を 2 回行ない、それぞれから mtDNA を精製して、標品 1, 2 とした)

表 1. コンティグの塩基数 (bp) とコンティグを構成するリードの数

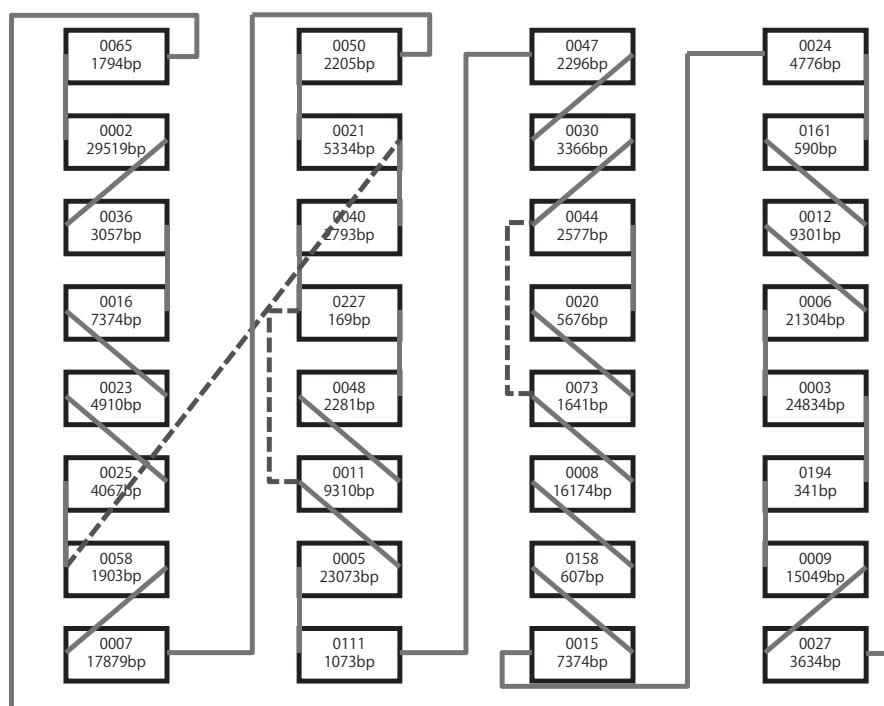
Serial No.	Contig No.	Length (bp)	Num. Reads
1	0002	29519	11751
2	0003	24834	12702
3	0005	23073	7017
4	0006	21304	9040
5	0007	17879	11881
6	0008	16174	6562
7	0009	15049	7267
8	0011	9310	5635
9	0012	9301	3567
10	0015	7374	3028
11	0016	7374	2173
12	0020	5676	3284
13	0021	5334	3003
14	0023	4910	2085
15	0024	4776	2292
16	0025	4067	2743
17	0027	3634	2478
18	0030	3366	1513
19	0036	3057	1109
20	0040	2793	1060
21	0044	2577	1741
22	0047	2296	890
23	0048	2281	1679
24	0050	2205	1266
25	0058	1903	1541
26	0065	1794	1143
27	0073	1641	1509
28	0111	1073	521
29	0158	607	405
30	0161	590	467
31	0194	341	340
32	0227	169	403

表 2. コンティグの隣接を証明するための PCR に用いたプライマー

No.	Name	Sequence (5'→3')	Contig		Result	Product size (bp)
			from	to		
1	(Bm) FSL_065-002_F	AGCATCGGGTAAATGGGATC	65	2	+	695
2	(Bm) FSL_065-002_R	TCCTTCACTCGACCAGAGCT	2	65		
3	(Bm) FSL_002-036_F	CCCTTTCCTCCTATGCTCAC	2	36	+	482
4	(Bm) FSL_002-036_R	TAAGCATGACTCGAGCACTT	36	2		
5	(Bm) FSL_036-016_F	CGTTAGAAGAAGATGAGCTG	36	16	+	474
6	(Bm) FSL_036-016_R	CCTTGTTGCTTGGTTGGATA	16	36		
7	(Bm) FSL_016-023_F	GAGCAGTATAAGCCCTTCTC	16	23	+	1225
8	(Bm) FSL_016-023_R	GATGGAGTCTACATCGTACG	23	16		
9	(Bm) FSL_023-025_F	CTATCTGCGTCGAGTGTGTC	23	25	+	1027
10	(Bm) FSL_023-025_R	CGTGCTTCCCTCACATTTTG	25	23		
11	(Bm) FSL_025-058_F	CCTCCTTCCAGCTCAGAATG	25	58	+	1102
12	(Bm) FSL_025-058_R	CTGAGATGTCCAGCGGATTC	58	25		
13	(Bm) FSL_058-007_F	CTCTACAATGCCCATAGAGG	58	7	+	1031
14	(Bm) FSL_058-007_R	ATGGAAGGAAGGCTAGAATC	7	58		
15	(Bm) FSL_007-050_F	GGTAGATAGCTCAGTAGTCG	7	50	一注	
16	(Bm) FSL_007-050_R	CTACTGCAGCCCATAAATG	50	7		
17	(Bm) FSL_050-021_F	TCACGACCTACAGGTAAGAC	50	21	+	1064
18	(Bm) FSL_050-021_R	CCGAACACCTGAAGTCCTTG	21	50		
19	(Bm) FSL_021-040_F	AGCGGTAAGGCATCCCAAGG	21	40	+	1140
20	(Bm) FSL_021-040_R	ATGACTTGTCCTGTGAAGC	40	21		
21	(Bm) FSL_040-227-048_F	TGTCGTATGAGAGGCTACAG	40	227	+	1056
22	(Bm) FSL_040-227-048_R	CCTACCTTACAAAGGGAACG	48	227		
23	(Bm) FSL_048-011_F	GCTCGTCTTATGGCATTAGC	48	11	+	1051
24	(Bm) FSL_048-011_R	ACTTCGCCTCTTTGTTGGAC	11	48		
25	(Bm) FSL_011-005_F	TCACTGGTATGACTCTGTGC	11	5	+	1082
26	(Bm) FSL_011-005_R	ACTTAGGCTTCTGCTTCTGG	5	11		
27	(Bm) FSL_005-111_F	TAGAGCAAGTTCCGTTCTGG	5	111	+	1058
28	(Bm) FSL_005-111_R	GAGAAGAAGGAACCGAGAAG	111	5		
29	(Bm) FSL_111-047_F	CTTCTCGGTTCCCTTCTTCTC	111	47	+	1134
30	(Bm) FSL_111-047_R	TACGCTTACCAAGCCTAGTC	47	111		
31	(Bm) FSL_047-030_F	ACCATAGCATGTTACACGAG	47	30	+	1076
32	(Bm) FSL_047-030_R	CGAGAACTTCCGGATCGAAG	30	47		
33	(Bm) FSL_030-044_F	TTCAAAGAACTGCGCTTAGC	30	44	+	1078
34	(Bm) FSL_030-044_R	TTGAACCTTGTCATGATCG	44	30		
35	(Bm) FSL_044-020_F	GATAGGACCAGTATCGGACG	44	20	+	1050
36	(Bm) FSL_044-020_R	TTCGCCCTCGTCAGAGAAAG	20	44		
37	(Bm) FSL_020-073_F	CGTAAAGGCAGATGTAGTAG	20	73	+	1030
38	(Bm) FSL_020-073_R	AGTCTCATCTTGTGTGTTGG	73	20		
39	(Bm) FSL_073-008_F	CGAATGTAACCTCCCGACTAC	73	8	+	1041
40	(Bm) FSL_073-008_R	CCGAAGAGGAAAGAAGAATC	8	73		
41	(Bm) FSL_008-158-015_F	GCAAGCCAAGCCGATGATAG	8	158	+	1189
42	(Bm) FSL_008-158-015_R	GCCATGGACTTGGATCTACC	15	158		
43	(Bm) FSL_015-024_F	GCACCAGAGGTGTACTAATC	15	24	+	1107
44	(Bm) FSL_015-024_R	AATAGGAATACTCGGACTCG	24	15		

45	(Bm)	FSL_024-161-012_F	TTTCTCGAAGCCTCTTGTTTC	24	161	+	1027
46	(Bm)	FSL_024-161-012_R	CTTTCTCGGTCCGAAAAGTG	12	161		
47	(Bm)	FSL_012-006_F	AGAAACAGAGTCGATACGAC	12	6	+	1101
48	(Bm)	FSL_012-006_R	GACCAACCTATCACCTATCC	6	12		
49	(Bm)	FSL_006-003_F	CTAGCTTTTACCGCTATCGC	6	3	+	1119
50	(Bm)	FSL_006-003_R	TTCTTCCCAGCATGCTATGG	3	6		
51	(Bm)	FSL_003-194-009_F	TTTCCGTTGGGAGCTCTCTC	3	194	+	1052
52	(Bm)	FSL_003-194-009_R	GAACAGCGAACAACCAGAGC	9	194		
53	(Bm)	FSL_009-027_F	AATAGAGGTGCAAGGCTGAC	9	27	+	1039
54	(Bm)	FSL_009-027_R	GCAGCTCCATTTCGTTTGTTG	27	9		
55	(Bm)	FSL_027-065_F	TAGGCGAGTGCAGTCACTTC	27	65	+	1111
56	(Bm)	FSL_027-065_R	GCTGGACTCTATTATCTCCG	65	27		
57	(Bm)	FSL_007-050_F _{new}	GACGAGACGTTAATGCGAAC	7	50	+	828
58	(Bm)	FSL_007-050_R _{new}	CTTTGATCCGGAGGAAAGAG	50	7		

図 4. 次世代シーケンシングにより得られた 32 個のコンティグの位置関係
 実線で繋がれたコンティグの環状経路が *B. mauroram* のミトコンドリアゲノムのマスターサークルに相当すると思われる。点線は、マスターサークルに収まらないコンティグの接続の可能性を示している。



4. 終わりに

ここ数年、さかんに行われている、次世代シーケンシングを用いた植物ミトコンドリアゲノムの解読は、このゲノムについて多くの知見をもたらしたのは事実であるが、植物ミトコンドリアゲノムはデータの蓄積が進めば進むほど、構造と機能、変異及び進化など、あらゆる面で不思議な事象が浮かび上がってくる。例えば、植物ミトコンドリアのゲノム上には、長短各種のタンデムあるいは逆位リピート配列が存在するが、マスターサークル上でアクティブに組換えを起こしてサブサークルを形成する配列は限られており、組換えにおけるリピート配列の活性が何によって規定されるのか不明である。また、植物ミトコンドリアゲノムでは、逆位や転座などの構造変異（進化的変化）により、新たなオープンリーディングフレーム（ORF）がしばしば形成されるが、個々の ORF が転写・翻訳されて機能を持つのか、多くの場合調査されていない。

さらに植物ミトコンドリアゲノムには、マスターサークルとサブサークルには反映されない、サブリモンと呼ばれる分子種が存在している。通常サブリモンは極めて低いコピー数で維持されているが、*msh1* などの核遺伝子が機能を失うと急激にコピー数を増大させ、ごく短期間にマスターサークルと置き換わってしまう（この現象は Substoichiometric shifting, SSS と呼称される）。サブリモンの存在に何か積極的な意味があるのか、またこの分子がゲノムの進化にどのような寄与をしているかなど、SSS に関しても不明な点が多い。

このように高等植物のミトコンドリアゲノムには、未だ解明すべき問題が多く残されている。これらのうちのいくつかは、ミトコンドリアのゲノム構造を近縁種間で詳細に比較することにより解決できると思われる。幸い本研究課題は、平成 26 年度の科研費基盤研究（B）に採択され、より大規模な実験を実施することが可能となった。この研究を通じて、植物ミトコンドリアゲノムに関する謎が、ひとつでも多く明らかになることを期待したい。

5. 参考文献

- Lonsdale, D. M., Hodge, T. P., and Fauron, C. M. R. 1984. The physical map and organisation of the mitochondrial genome from the fertile cytoplasm of maize. *Nucl. Acids Res.* **12** (24): 9249-9261.
- Ogihara, Y., *et al.* 2005. Structural dynamics of cereal mitochondrial genomes as revealed by complete nucleotide sequencing of the wheat mitochondrial genome. *Nucl. Acids Res.* **33** (19): 6235-6250.
- Palmer, J. D., and Herbon, L. A. 1986. Tricircular mitochondrial genomes of *Brassica* and *Raphanus*: reversal of repeat configurations by inversion. *Nucl. Acids Res.* **14** (24): 9755-9764.
- Suyama, Y., and Bonner, W. D. 1966. DNA from plant mitochondria. *Plant Physiol.* **41** (3): 383-388.
- Tanaka, Y., Tsuda, M., Yasumoto, K., Yamagishi, H., and Terachi, T. 2012. A complete mitochondrial genome sequence of Ogura-type male-sterile cytoplasm and its comparative analysis with that of normal cytoplasm in radish (*Raphanus sativus* L.). *BMC Genomics* **13**: 352.
- Unsold, M., Marienfeld, J. R., Brandt, P., and Brennicke, A. 1997. The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nature Genet.* **15** (1): 57-61.

6. 謝辞

本研究に用いた材料は、宇都宮大学農学部、房 相佑博士から分譲していただいた。ここに深謝の意を表す。また実験データは、京都産業大学工学研究科修士1年、岡部真弥氏によるものである。この研究への尽力に感謝する。

Comprehensive studies on the mitochondrial genome of higher plant based on the complete genome sequences

Toru TERACHI

Abstract

This project aims at giving a new idea on the structure, function, diversity and evolution of mitochondrial genome of higher plant, by comparing genome sequences among ten different types of mitochondrial genome from each of wheat and radish taxa. For this purpose, the next-generation sequencer will be efficiently used to obtain a large amount of sequence data, and bioinformatic approach will be adopted to compare whole mitochondrial genomes among the taxa. As the first step toward the project goal, we have obtained about 66.9Mb sequence data from mtDNA of *Brassica maurorum* whose cytoplasm is known to induce male-sterility to radish. The 32 contigs with mitochondrial sequence were selected from 242 contigs assembled. Connection between two neighboring contigs was examined by PCR amplification of an overlapping fragment and its direct sequencing, and a master circle molecule consisting of 236,281bp sequence was predicted for *B. maurorum* mitochondrial genome.

Keywords : mitochondrial genome, next-generation sequencer, *Brassica maurorum*, Radish, Cytoplasmic male sterility (CMS)