

# テキストマイニングを用いたウイルス感受性昆虫と抵抗性昆虫の 比較トランスクリプトーム解析

岩田 賢人

東京農工大学大学院 農学府 生物制御科学専攻

## 要旨

昆虫病原ウイルスは害虫防除資材として用いられており、害虫のウイルスに対する抵抗性発達が問題となっている。昆虫のウイルス抵抗性メカニズムを解明するために、ウイルス感受性と抵抗性のチャノコカクモンハマキ（チョウ目ハマキガ科）を用いて、RNA-Seqにより比較トランスクリプトーム解析を行った。従来のGO解析等の手法では、RNA-Seqの結果からウイルス抵抗性系統に特徴的な機能を抽出することが出来なかった。これは、本研究で用いた昆虫のゲノム情報が限られており、解析に供試できる情報が充分でなかったためである。そこで本研究では、Blast検索により付与されたnrアクセッション番号のdescriptionに対して、Text Mining Studioを用いたテキストマイニングを実施した。その結果からウイルス感受性系統と抵抗性系統の情報を比較し、抵抗性系統で特徴的な遺伝子機能を抽出することが出来た。

## 諸言

節足動物である昆虫も人間と同様にウイルス病にかかり感染致死する。特に、バキュロウイルスは昆虫に特異的に感染し、非常に高い病原性を有することが知られている。バキュロウイルスとは、エンベロープを有する環状2本鎖DNAウイルスである。バキュロウイルスの感染機構は以下の通りである（岩永, 2014）（図1）。バキュロウイルスの感染は、ウイルス包埋体を昆虫の幼虫が摂食することにより始まる。ウイルス包埋体とは、タンパク質の結晶が包埋体由来ウイルス粒子を包埋した構造であり、包埋体が中腸と呼ばれる幼虫消化管内で溶解するとウイルス粒子が放出される。ウイルス粒子は、消化管の上皮細胞に侵入し、ウイルス遺伝子の発現、ゲノムDNAの複製により、娘ウイルスを産生する。その後、娘ウイルスは幼虫の体腔へ拡散し、全身の組織へ感染を拡大する。感染末期には感染細胞において次世代のウイルス包埋体が形成され、通常感染虫の液化により包埋体が植物体などを汚染することで次世代への感染源となる。

このバキュロウイルスは、その高い感染性から生物農薬として製剤化され、ブラジルのダイズ圃場におけるヤガ科害虫の防除や、欧州や米国を中心に20カ国以上の果樹園における果樹害虫コドリリングの防除に用いられている（仲井, 2013）。しかしながら、近年

ウイルス資材を散布した圃場において、バキュロウイルス抵抗性昆虫の発生が認められ欧州を中心に問題になっている (Fritsch *et al.*, 2006)。そこで、昆虫のバキュロウイルスに対する抵抗性のメカニズムを解明するために、当研究室において昆虫野外個体群に対して核多角体病ウイルス (バキュロウイルス科) を接種し、選抜実験を繰り返すことでウイルス抵抗性の昆虫系統を人為的に作出した。本研究に用いた昆虫は、チャノココクモンハマキ (チョウ目ハマキガ科) で、これは茶樹の主要な害虫であり、多くの化学合成農薬に対する抵抗性の獲得が報告されている (内山ら, 2013)。

本研究では、ウイルスによる選抜を行っていない感受性系統と、前述のウイルス抵抗性系統の比較トランスクリプトーム解析により、昆虫のバキュロウイルス抵抗性に関わる遺伝子機能の解明を目指した。まず、両系統の消化管由来の RNA を用いた RNA-Seq を実施し、その結果に対してモデル昆虫であるカイコの Gene Ontology 情報を用いた GO 解析を行った。しかしながら、カイコの GO アノテーションの付加率は 5%程度であり、感受性系統と抵抗性系統間における特徴的な差異を抽出することが出来なかった。そこで、Text Mining Studio を用いたテキストマイニングによる解析を実施した。

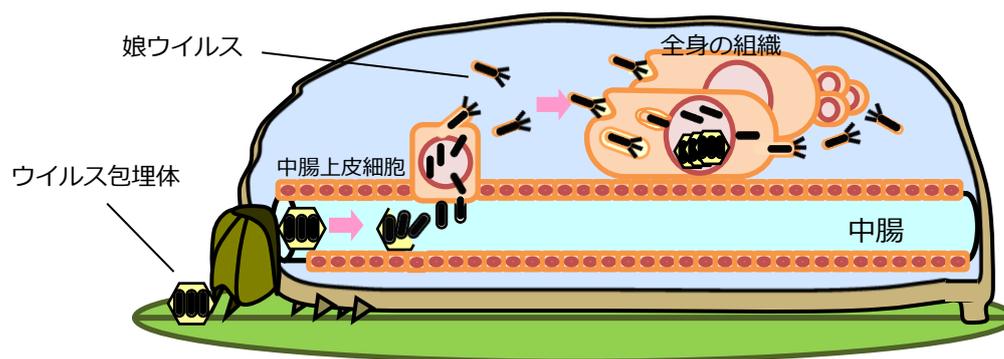


図 1. バキュロウイルスの感染過程。

## 方法

### RNA-Seq

ウイルス感受性系統と抵抗性系統のチャノココクモンハマキに、感受性系統の 95%以上の個体が致死する濃度のウイルスを飲下させた。飲下 1 日後に、解剖によって中腸 (消化管) 細胞を回収した。各系統 5 頭分の組織をプールし、それぞれ Total RNA を抽出した。その後、mRNA を精製し、cDNA ライブラリを作製、MiSeq Reagent Kits v3 (Illumina) を用いて MiSeq (Illumina) によりシーケンスを実施した。FastQC により得られたリードのクオリティチェックを行った。

### de novo assembly と nrBlast 検索

感受性系統由来のリードと抵抗性系統由来のリードをまとめて、Trinity による *de novo assembly* を実施した。得られた contig で、nrBlast 検索により bit score 80 以上のものに nr アクセション番号を付加した。

### 各系統由来のリードのマッピングと発現変動遺伝子の絞り込み

Bowtie2 を用いて、*de novo assembly* により得られた contig に対し各系統由来のリードをマッピングした。マッピングの結果をもとに、contig 群を各系統特異的 contig と共通発現 contig にグループ分けした。さらに、両系統でマッピングされた contig 群について、発現変動遺伝子を絞り込むために、TCC パッケージを用いて MA プロットを作製し、FDR < 0.05 で発現変動遺伝子を判定した。

### Text Mining Studio を用いた解析

マッピング及び MA プロットにより分けられた contig 群毎に、nrBlast 検索により付加された nr アクセション番号の description の解析を行った (図 2)。その結果から、感受性系統と抵抗性系統に特異的な contig 群を比較し、抵抗性系統で特異的な 遺伝子機能 (文章/単語) を抽出した。また、発現変動 contig 群も同様に系統間で比較し、抵抗性系統で特徴的に発現上昇している遺伝子機能を抽出した。

contig 情報の例

ID	nrBlast検索によるDescription
comp25700_c0_seq1	ubiquitin-specific protease [Danaus plexippus].
comp26815_c0_seq1	uncharacterized protein LOC105397321 [Plutella xylostella].
comp26864_c2_seq3	glutathione S-transferase epsilon 4 [Cnaphalocrocis medinalis].
comp11679_c0_seq1	PREDICTED: zinc metalloproteinase nas-15-like [Papilio xuthus].

前処理により種名等削除



Text Mining Studio による文章分類、単語頻度解析

図 2. テキストマイニングによる解析過程.

## 結果と考察

### de novo assembly と nrBlast 検索

感受性系統と抵抗性系統由来のリードを用いた *de novo assembly* の結果、60503 個の contig が作製された (表 1)。また、それらの contig に対して nrBlast 検索を行ったところ、23613 個の contig に bit score 80 以上で、nr アクセション番号が付与された (表 1)。

表 1. contig の作製と Blast 検索

	Contig数
総数	60503
blast検索の結果あり	23613

### 各系統由来のリードのマッピングと MA プロット

各系統由来のリードを *de novo* assembly により得られた contig にマッピングしたところ、contig は、感受性系統のリードのみマッピングされるもの、抵抗性系統のリードのみマッピングされるもの、両系統のリードがマッピングされるものに分類された (図 3)。また、両系統のリードがマッピングされた発現変動を示す MA プロット (図 4) を作製したところ、45 個の抵抗性系統発現上昇 contig と 44 個の感受性系統発現上昇 contig が抽出された。

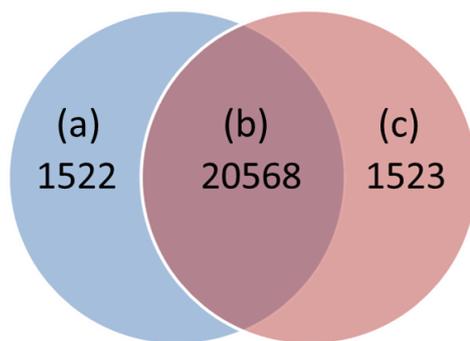


図 3. マッピングによる Contig の分類.

- (a) 感受性においてのみマッピングされた Contig 数.
- (b) 両系統でマッピングされた Contig 数.
- (c) 抵抗性においてのみマッピングされた Contig 数.

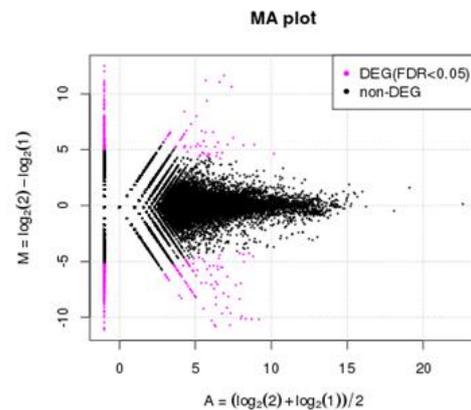


図 4. MA プロット.

赤点を発現変動 contig として抽出した.

### 系統特異的 contig の description 文章分類

マッピングによるグループ分けの内、感受性系統に特異的な contig 群 (図 3a) と抵抗性系統に特異的な contig 群 (図 3c) に着目し、nrBlast 検索により付与されたアクセッション番号の description について文章分類を行った。上位 10 位を比較したところ、感受性系統と抵抗性系統で類似した傾向が見られた (図 5)。これは、同様の機能を有するが、異なる遺伝子を両系統がそれぞれもつことを示唆する。また、意味をなさない分類も多く、テキストマイニングを行う前の description の前処理の方法を検討する必要があるだろう。ただし、一部抵抗性系統に特異的な分類として、「dehydrogenase」や「methyltransferase」

が抽出された (図 5)。FAD-glucose dehydrogenase は、昆虫の免疫系に関与することが知られており (McNeil et al., 2010)、抵抗性系統が特異的に有する dehydrogenase により、病原体に対する免疫系が活性化されている可能性が考えられる。また、methyltransferase の働きを阻害するとバキュロウイルスのタンパク質合成が抑制されることが報告されており (Bach et al., 1995)、抵抗性系統は methyltransferase に変異を有し、ウイルスのタンパク質合成を抑制しているのかもしれない。

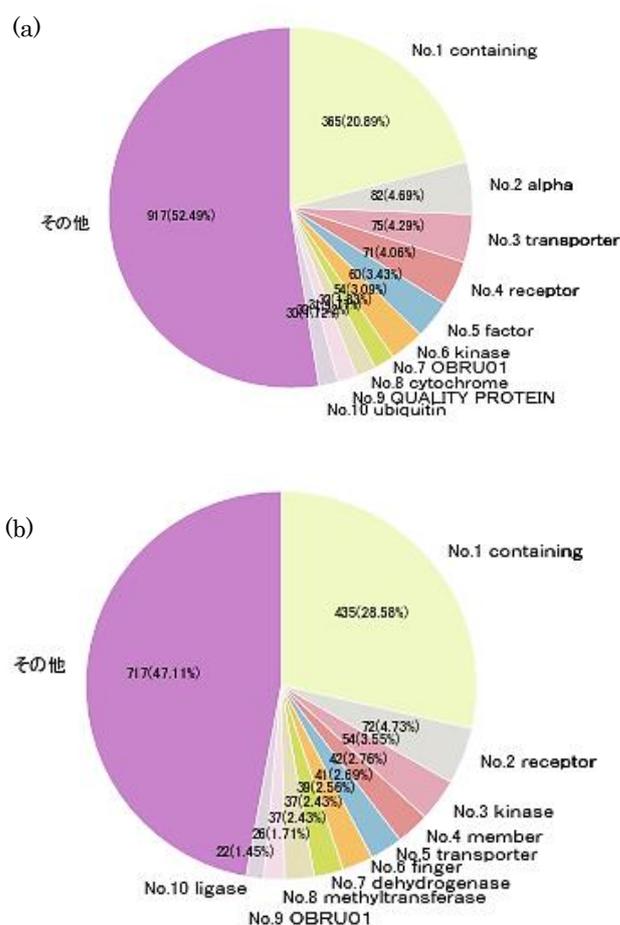


図 5. 系統特異的 Contig 群の文章分類。  
 (a) 感受性においてのみマッピング。  
 (b) 抵抗性においてのみマッピング。

### 系統ごとの系統発現上昇 contig の description 文章分類と全体に対する単語頻度解析

マッピングによるグループ分けの内、ウイルス感受性と抵抗性の両系統でマッピングされた contig 群 (図 3b) に着目し、MA プロット (図 4) から示された抵抗性系統発現上昇 contig と感受性系統発現上昇 contig について、nrBlast 検索により付与されたアクセシオン番号の description について文章分類を行った。上位 10 位を比較したところ、その分類は、抵抗性系統発現上昇 contig と感受性系統発現上昇 contig で大きく異なった。すなわち、ウイルスに対する応答の差異により、両系統の中腸 (消化管) 細胞における遺伝子発

現パターンが大きく異なることが明らかとなった。特に、感受性系統で発現上昇している分類には、「VP1054」や「polyhedral calyx envelope」といったバキュロウイルスがコードしている遺伝子機能が含まれおり、感受性系統においてウイルス感染がより進行していることが示唆された（図 6）。

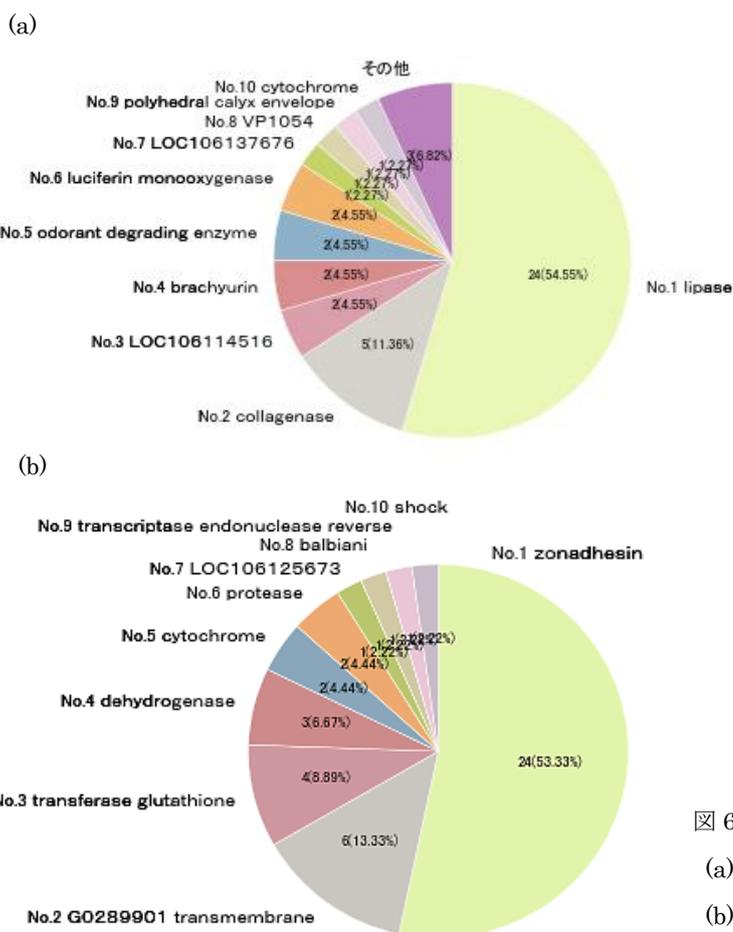


図 6. 系統ごとの系統発現上昇 Contig.  
 (a) 感受性系統で高発現.  
 (b) 抵抗性系統で高発現.

次に、抵抗性系統発現上昇 contig (44 個の contig) と感受性系統発現上昇 contig (45 個の contig) を合わせた、全体 (89 個の contig) に対する単語頻度解析を行い、各系統で特異的に発現が上昇している遺伝子機能 (単語) の抽出を行った。その結果、「collagenase」をはじめ文章分類で上位を占めた多くの機能について、感受性系統でのみ発現上昇が見られることが示された。ただし、「lipase」に関しては、感受性系統において頻度が高いが、一部抵抗性系統発現上昇 contig にも含まれることが分かった (図 7a)。カイコにおいて、ある種の lipase は、バキュロウイルスの感染の有無に関わらず発現しており、ウイルス粒子を不活化する作用を有することが知られている (Ponnuvel et al., 2003)。このことから、抵抗性系統では、感受性系統が多く発現する lipase の代わりに、ウイルス失活作用を持つ

ような lipase を消化管内に多く発現し、感染を抑制している可能性が考えられる。さらに、抵抗性系統発現上昇 contig からは、「protease」や「hydrolase」といった消化酵素を意味する単語が抽出されており（図 7b）、これらが消化管内においてウイルス粒子を失活させているかもしれない。これらの他にも、系統ごとに特徴的な単語を抽出することができた。しかしながら collagenase など、ウイルス感染に対する応答と、これらの単語（機能）の関係性について先行研究が見当たらない遺伝子機能も存在した。

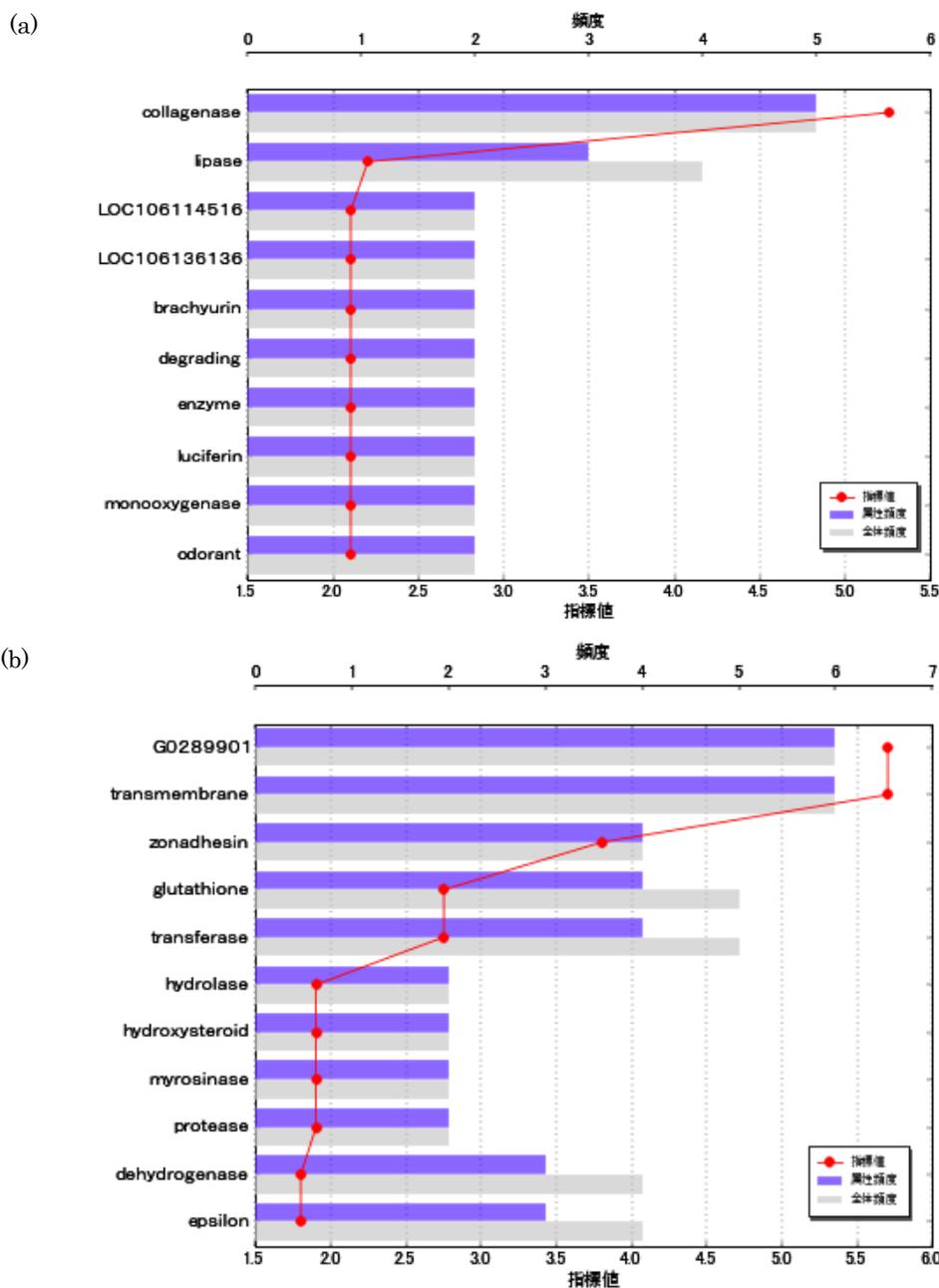


図 7. 各系統で発現上昇した contig における両系統を合わせた全体に対する単語頻度解析。  
 (a) 感受性系統で高発現. (b) 抵抗性系統で高発現.

## まとめ

Text Mining Studio を用いた解析により、昆虫のバキュロウイルス抵抗性系統における特徴的な遺伝子機能を初めて抽出することができた。本研究によって、バキュロウイルス感染に関与する可能性のある、新たな遺伝子機能が抽出された。この知見は、今後ウイルス感染と宿主昆虫の相互作用をより詳細に調査する手がかりとなるだろう。また、ウイルス抵抗性に関与する遺伝子を特定するためには、RNA-Seq の反復実験を行い、同様なテキストマイニングを用いた解析を実施することで、実際にウイルス抵抗性と感受性に関与する様な、文章分類や高頻度の単語の榨りこみを行う必要がある。さらに、今後は抽出された遺伝子機能を元に、生物学的な手法によって抵抗性に関与する遺伝子の特定を目指す。

## 謝辞

本研究を行うにあたりご指導を賜りました、東京農工大学 石井一夫 特任教授、仲井まどか准教授に感謝いたします。

## 引用

- 岩永将司. (2014). バキュロウイルスの特性と増殖, 第 6 章ウイルス病. *最新昆虫病理学*, pp134-145.
- 内山徹, 小澤朗人, & 劉主. (2013). 静岡県のチャ園に生息するチャノコカクモンハマキの薬剤感受性とジアシルヒドラジン系 IGR 剤に対する薬剤抵抗性. *日本応用動物昆虫学会誌*, 57(2), 85-93.
- 仲井まどか. (2013). 昆虫ウイルスを利用した害虫防除資材の現状と展望. *化学と生物*, 51(10), 693-701.
- Bach, C., Cramer, A., & Scholtissek, C. (1995). Effect of methyltransferase inhibitors on the regulation of baculovirus protein synthesis. *Journal of general virology*, 76(4), 1025-1032.
- Eberle, K. E., & Jehle, J. A. (2006). Field resistance of codling moth against *Cydia pomonella* granulovirus (CpGV) is autosomal and incompletely dominant inherited. *Journal of invertebrate pathology*, 93(3), 201-206.
- McNeil, J., Cox-Foster, D., Slavicek, J., & Hoover, K. (2010). Contributions of immune responses to developmental resistance in *Lymantria dispar* challenged with baculovirus. *Journal of insect physiology*, 56(9), 1167-1177.
- Ponnuvel, K. M., Nakazawa, H., Furukawa, S., Asaoka, A., Ishibashi, J., Tanaka, H., & Yamakawa, M. (2003). A lipase isolated from the silkworm *Bombyx mori* shows antiviral activity against nucleopolyhedrovirus. *Journal of Virology*, 77(19), 10725-10729.