

Short Report

トランスクリプトームデータから新規の機能性分泌タンパク質群を探索する
フジツボを例として岡野桂樹¹, Wong Yue Him¹, 村口元¹, 原光二郎², 小黒-岡野美枝子^{1,3},¹ 秋田県立大学 生物資源科学部 応用生物科学科, ² 秋田県立大学 生物資源科学部 生物生産科学科,³ ヤマザキ学園大学 動物看護学部 動物看護学科

de novo トランスクリプトーム法はゲノム配列がわかっていない生物において、細胞や組織に発現する mRNA 配列の総体を Trinity などのアセンブラーを用いることで構築し、それぞれの mRNA の特徴づけ (アノテーション) を行い、かつそれぞれの mRNA の発現量を網羅的に調べる手法である。比較的安価に行うことができるため、非モデル生物において爆発的に使われるようになった。ただし、アノテーションのつかない mRNA 配列やデータベース中に相同性が見られない、いわゆる“新規配列群” (dark transcriptome) は、ほぼ完全に無視されている。我々はアカフジツボの幼生セメントタンパク質や殻形成に関わるタンパク質をコードする遺伝子 (新規遺伝子) 群を探索する上で、*de novo* トランスクリプトームデータを利用した。その中で、特定の分泌細胞群を同定し、その細胞群に特異的に大量に発現し、シグナル配列を有する新規遺伝子群の探索が、極めて有効であることを見出したので、ここに報告する。

キーワード: 海洋無脊椎動物, 新規機能性分泌タンパク質, 次世代シーケンス, トランスクリプトーム

トランスクリプトームとは、ある特定の細胞が発現する転写産物 (トランスクリプト) の総体 (オーム) を指す言葉である。次世代シーケンスを用いたトランスクリプトーム解析とは、細胞または組織などから抽出された mRNA を基に cDNA ライブラリーを作製し、その配列と発現量を調べる手法である。

同義語に「RNA-seq」がある。RNA-seq のほうがより広義で、RNA 分子の情報を網羅的に調べる手法全体をさす。RNA 亜集団を分取して行えば、タンパク質をコードする成熟 mRNA だけでなく、miRNA や ncRNA の解析も可能である。この事情に関しては優れた成書があるので、それらを参照していただきたい (二階堂, 2014; 鈴木, 2016)。

RNA-seq の典型的な解析例を図 1 に示す。これは、ウシグソヒトヨタケの子実体形成過程を RNA-seq 解析したものである (Muraguchi *et al.*, 2015)。転写産物のデータをゲノム上にマップすることで、どの遺伝子がどのくらいの程度発現するかを全ゲノムレベ

ルで調べることができる。さらに cDNA 合成時に、

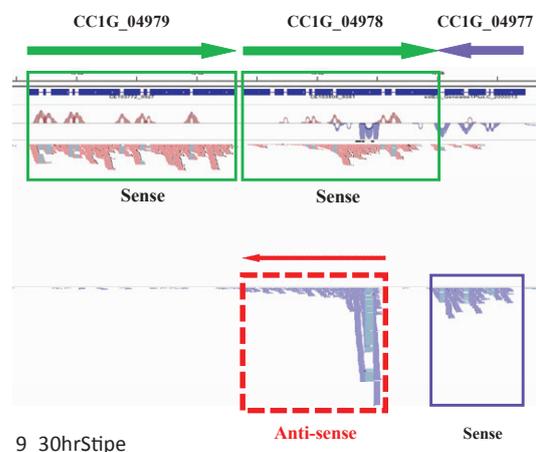


図 1 典型的な RNA-seq 解析の例: ウシグソヒトヨタケ
Strand-specific RNA-seq 法を利用すると、それぞれの遺伝子の Exon, Intron, 転写開始点などの情報とその発現量に加え、anti-sense の発現も同時に調べることができる (Muraguchi *et al.*, 2015 の図 9 を一部改変)。

sense 鎖と anti-sense 鎖を区別する手法 (Strand-specific RNA-seq) を用いることで、通常の遺伝子発現に加えて、anti-sense 鎖の発現も調べることができる (図 1, 赤の点線の囲み)。

品質の良いトータル RNA が得られれば、手軽にトランスクリプトーム情報が得られるため、あらゆる分野の研究者に爆発的に利用されるようになった。また、RNA-seq の普及は、モデル生物と非モデル生物の間の情報量の垣根を劇的に下げることにも大きく貢献している。非モデル生物を研究する多くの研究者にとって大きな朗報である。生物の多様性について、遺伝子レベルで研究することができる時代になってきている。

ゲノム情報のない非モデル生物のトランスクリプトーム研究は、*de novo* トランスクリプトーム研究と言われる。ゲノムに換わるその生物のトランスクリプトーム (cDNA 配列の総体) 配列を Trinity などの *de novo assembler* プログラムを用いて得る方法である。得られた配列情報について、発現解析を行うとともに、Blast 検索を含めた Gene ontology (GO) 解析を行うことで、アノテーションを付加して使う。ホモロジーが期待できる酵素や Hox タンパク質、受容体、リボソーム関連タンパク質などをコードする遺伝子群の比較研究には絶大な効果を示す。

ただし、*de novo* トランスクリプトーム法については、以下の 2 点で不満もある。第一にアノテーションが哺乳類やいわゆるモデル生物由来で、非モデル生物への適用が必ずしも正確ではないため、アノテーションに基づく解析の多くについて、信頼性が十分ではないこと、第二にホモロジーが極端に低く、アノテーションがつかない未知 (新規) の遺伝子を完全に無視していることである。逆に言えば、*de novo* トランスクリプトーム研究には大きな可能性がある。すなわち、多数の新規の遺伝子配列がその中に存在し、解明を待っているのである。

我々はアカフジツボが作り出す素材に興味がある。フジツボの特徴といえば、殻を持ち、海水中で働く接着剤 (フジツボではセメントと称する) を持つことである (図 2)。これらはフジツボの仲間が独自に進化させた分子である可能性が高く、ホモロジー検索が有効ではない、いわゆるアノテーションのつか

ない新規遺伝子群の可能性が高い。

我々の目的は、フジツボの接着剤がどのような成分からなり、どのように作られるのか? 殻形成にはどのようなタンパク質が関与しているのかを明らかにすることである。特に分子レベルの研究のない幼生接着剤の研究を行ってきた。しかし、サンプルの入手はむずかしく、培養系はなく、幼生サイズが 1 mm 以下と小さいにも関わらず、極めて複雑な細胞構築を有するなど、研究材料として多くの問題点を抱えていた。その中で、学長プロジェクトを得て、本格的にトランスクリプトーム研究に着手した。ここでは、我々がフジツボ由来の新規遺伝子群を探索する過程とその過程で得たトランスクリプトーム解析の有効性に関する知見を記載する。

材料と方法

アカフジツボ、レイシガイ、寄生性フジツボは秋田県の沿岸で採集されたものを使用した。アカフジツボについては一部、岩手県沿岸で採集されたものも使用した。RNA 抽出には Nucleospin RNA kit (Takara) を使用した。次世代ライブラリーの作製には、TruSeq RNA Library Preparation Kit v2 (Illumina) または TruSeq Stranded mRNA Library Prep Kit

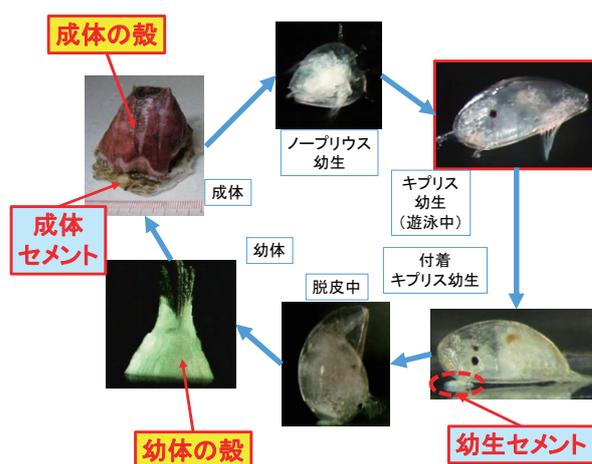


図 2 アカフジツボの生活環とつくりだす素材

フジツボはプランクトン生活を送る幼生期と岩などに固着して生きる成体期からなる。幼生期から固着期に移動する際に放出されるセメントを幼生セメント、固着を強固なものにするセメントを成体セメントと呼ぶ。殻は附着後形成を開始し (幼体の殻)、その後成長に従い、ずっと作られる (成体の殻)。

(Illumina) を使用した。シーケンスは秋田県立大学の HiSeq1000 (Illumina) を使用した。de novo トランスクリプトームの assembly には Trinity を、ホモロジー検索と GO 解析 (GO) には Blast2GO PRO を用いた。

結果と考察

1. フジツボのトランスクリプトーム解析：新規機能性タンパク質をコードする遺伝子群の探索法

我々の基本的な戦略は単純である。体の外にある機能性構造体は、分泌性タンパク質と分泌性酵素が創り出す物質により構成されている。したがって、特殊な機能性構造体を作り出すには、特殊な分泌性タンパク質を生産する特殊な分泌細胞が必要である。また、それらの分泌細胞は特殊な機能性構造体タンパク質を生産する工場となっており、そのようなタンパク質が特異的に、かつ大量に発現しているはずである。そこで、特殊な分泌細胞を単離し、そのトランスクリプトームを調べ、細胞特異的に大量発現する遺伝子群を探索すればよい。

ただし、実際にはそうは簡単ではない。なぜなら、特殊な構造体を創り出す分泌細胞を厳密に同定するためには、その分泌細胞が分泌したものが、その構造体づくりに直接関与する証明が必要である。いわば、鶏と卵の関係にある。

もちろん、組織学的研究はフジツボには多数ある。そこで、分泌細胞を含む部分とそれ以外の部分にねらいをつけ、微細解剖でなるべく細かく組織を分けて単離し、RNA を抽出する。ついで、cDNA ライブラリーを作製して古典的なサンガーシーケンスで調べるか、トランスクリプトームを取得し、さまざまな状況証拠を加えて、まず候補遺伝子を明らかにする。さらに、タンパク質レベルでその証明を行うことで、分泌細胞と機能性構造体との関係を明らかにする。最後に、分泌細胞のトランスクリプトームに戻り、それを網羅的に解析することで、全体像を明らかにする。これが我々の戦略である。

すでに、我々は古典的な分子生物学的手法を用いて、アカフジツボの幼生接着剤がどのような細胞から分泌されるかについて、検討を重ねてきた (岡野、

2006; 岡野, 2007)。また、それらをコードする遺伝子の発現が、ノープリウス-キプリス変態の前後だけに局限され、鋭いピークを持つこと、すなわち、時期的な情報を得ていた。そこで、2 群のトランスクリプトームを作製した。一つはノープリウス-キプリス変態前後のノープリウス幼生 6 期(III ステージ)とキプリス変態直後の幼生 (図 2) のトランスクリプトームを中心とする幼生の発生段階のトランスクリプトームである。もう一方は、ノープリウス-キプリス変態直後のキプリス幼生を微細解剖して得たセメント腺複合体と殻、蔓脚それぞれのトランスクリプトームである。

微細解剖は狙いを絞る意味でとても良い方法である。しかし、RNA の質と量は必ずしもいつも満足できるわけではない。そこで、全長配列を幼生全体から得て、微細解剖サンプルを発現解析のための探索プローブの 1 つとして使用した。この方法論を用いて、現在 14 種類の候補を得て、そのうち 7 種類は免疫染色により、幼生接着剤タンパク質であることを確認することができた (Okano *et al.*, in prep.)。我々は、アカフジツボの幼生接着剤を構成するタンパク質のほぼ全容を明らかにできたと考えている。得られた遺伝子産物 (タンパク質) は酵素とプロテアーゼインヒビターを除き、12 種の遺伝子群がデータベース上で、他のタンパク質とホモロジーが検出できない新規タンパク質であったことを追記しておきたい。

この過程で我々の構想を具体化した機能性構造

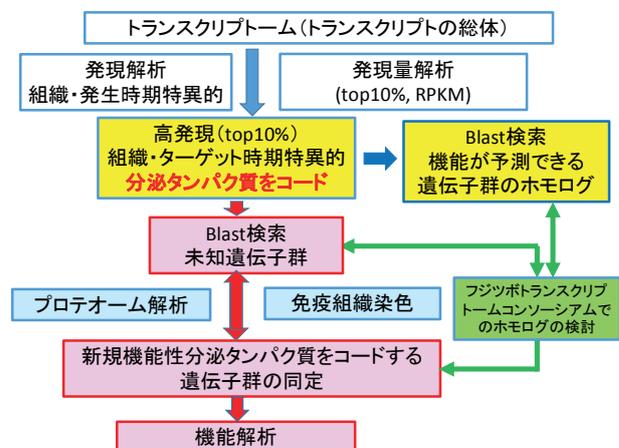


図3 新規機能性分泌タンパク質群をコードする遺伝子群の探索に関する概念図

体を構成する新規分泌タンパク質遺伝子の探索法の概要を図3に示す。

適切な時期に、標的となる分泌細胞を含む分画からトランスクリプトームを作製し、1) 高発現遺伝子群 (top10%) を選抜する。ついで、2) すでに得られていた幼生セメントと同じ組織、発現時期を示す遺伝子群 (時期、組織・細胞特異的遺伝子) を選抜する。さらに、3) N末端部位にシグナルペプチド配列を有する分泌タンパク質群をコードする遺伝子群を選抜する。この一次スクリーニングで得られた遺伝子群について、アノテーションの有無を調べ、特にアノテーションがつかない遺伝子群を選抜する。これらが主要な標的である。これらの新規遺伝子について、定量PCRを用いた発現パターンによる特異性の確認と、プロテオームおよびそれぞれの遺伝子産物の特異抗体を用いた免疫染色による局在データを加味して、新規機能性分泌タンパク質の配列を持つ遺伝子群を決定するのである。

出発点のトランスクリプトームは標的となる分泌細胞 (組織) を含む集団である。その対照に標的を含まない組織、および、標的細胞が分化していない発生段階を選んでおく。ただし実際には、適切な時期、標的となる分泌細胞を探す上でも、トランスクリプトームはきわめて有効であり、相互依存の関係である点を強調しておきたい。そこで最初から、標的の分泌細胞がはっきりしなくても問題はない。微細解剖で可能性のある部分をなるべく細かく分画する (再現性を得るために必ず写真を撮っておく) ことが肝要である。この方法はアカフジツボの殻形成に関わるタンパク質の同定 (Wong *et al.*, in prep) でも、有効性が確認できている。

なお、レイシガイの卵嚢形成に関わるタンパク質の同定の場合、レイシガイの卵嚢をつくりつつある個体から標的細胞 (組織) がほぼ純粋に採取できたため、組織特異的な遺伝子で分泌タンパク質をコードする遺伝子群の上位 20 個の遺伝子を解析するだけで十分な結果が得られた (Wong *et al.*, in prep)。この例は分泌細胞が特定の分泌タンパク質の発現に特化していることを示す良い例である。

2. フジツボトランスクリプトームコンソーシアム

アカフジツボのトランスクリプトームにおいては、アノテーションがつかない遺伝子 (新規遺伝子) は約 50 %を占める。我々はそれを宇宙における dark matter を模して、dark transcriptome と呼んでいる。分泌タンパク質をコードする組織・時期特異的大量遺伝子にしぼっても、新規遺伝子は数多く存在する。我々は、dark transcriptome 中に存在する遺伝子の多くは、フジツボ特有の進化の過程で生み出されてきたものであると仮定している。幼生セメント遺伝子群も、殻形成に関与する遺伝子群も、我々にとっては「フジツボとはどのような生物か」を明らかにするための dark transcriptome 研究の一環であるとも言える。

トランスクリプトーム情報をうまく活用する鍵は、ホモロジー検索とアノテーションである。dark transcriptome を解析する上では、明らかにフジツボの分類群内で、情報を蓄積する必要がある。

節足動物の分類は、劇的に変貌をとげているが、最近の分類 (Regier *et al.*, 2010) に従うと、フジツボの仲間 (蔓脚下綱) はエビやカニ (軟甲綱) から分かれた鞘甲亜綱の一員である。

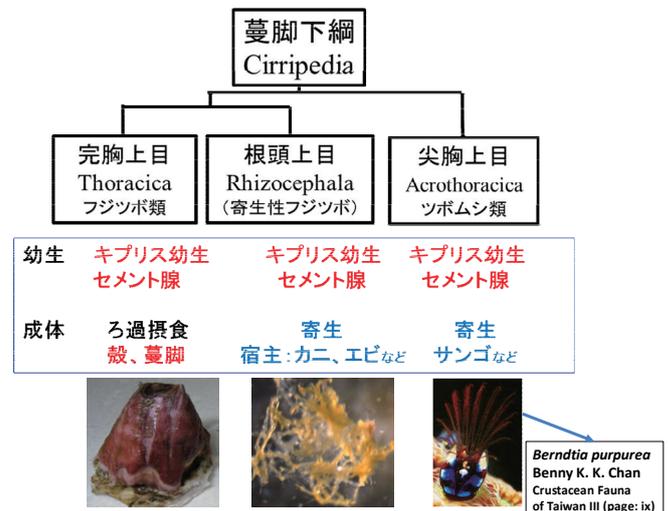


図4 フジツボの仲間 (蔓脚下綱) の分類と幼生形態、成体形態の多様性

図4は蔓脚下綱における幼生形態の共通性と成体形態、および生活スタイルの多様性を示している (参考: Hoeg, 2009)。富士山型の炭酸カルシウム結晶を

持つ殻をもつ典型的なフジツボは、完胸上目の特に無柄目に属す(図4左)。しかし、他の2つのグループの成体の構造は驚くほど多様で、それぞれ異なった宿主と成体の形態を有する寄生性生物である。蔓脚下綱に共通するのは、フジツボの仲間だけにみられる付着に特化したキプリス幼生と付着のためのセメント腺をもつことである。また、キプリス幼生の体づくりはノープリウス-キプリス変態の前後で行われることも共通している。

そこで、台湾中央研究院の Benny K. K. Chan 博士らと共同で国際フジツボトランスクリプトームコンソーシアムをつくり、幅広いフジツボ分類群におけるキプリス幼生の体づくり(幼生セメントを含む)に関する研究を行いつつある(Wong *et al.*, in prep)。同様に殻をつくる完胸上目における殻づくり遺伝子群のホモロジー検索によるドメインの検索も実行中である。

謝辞

本研究は、秋田県立大学平成26~28年度学長重点プロジェクト研究費の支援を受けて行われた。次世代シーケンスは秋田県立大学バイオテクノロジーセンター所有のHiSeq1000を利用した。

文献

- Blast2GO PRO (<https://www.blast2go.com/blast2go-pro>)
- Hoeg, J. T., (2009) Evolution of Morphology, Ontogeny and Life Cycles within the Crustacea Thecostraca. *Arthropod Systematics & Phylogeny* 67 (2) : 199 – 217.
- 二階堂愛 (2014) 実験医学別冊「次世代シーケンス解析スタンダード NGS のポテンシャルを活かし切る WET & DRY」羊土社
- 岡野桂樹 (2006) 「キプリス幼生の付着機構 2—幼生セメント分泌のしくみと幼生セメントの構造」“フジツボ類の最新学”の第9章, 恒星社厚生閣(日本付着生物学会編) ISBN:4-7699-1033-9 pp. 168-189)
- 岡野桂樹 (2007) “フジツボの幼生セメント” *化学と生物* 45(12): 823-825.
- Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, *et al.* (2015) Strand-Specific RNA-Seq Analyses

of Fruiting Body Development in *Coprinopsis cinerea*

PLOS ONE, 10(10): e0141586.doi:10.1371/

journal.pone.0141586

Regier, J.C., Shultz, J. W., Zwick, A., Hussey, A., Ball, B., Wetze, R., Martin, J. W. and Cunningham, C. W. (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature*, 463: 1079-1083. doi:10.1038/nature08742

鈴木謙 (2016) 実験医学別冊「RNA-Seq 実験ハンドブック」羊土社

Trinity (<https://github.com/trinityrnaseq/trinityrnaseq/wiki>)

〔平成29年6月30日受付〕
〔平成29年7月11日受理〕

Use of *de novo* transcriptome data to identify novel genes encoding functional secretory proteins in barnacle, *Megabalanus rosa*

Keiju Okano¹, Wong Yue Him¹, Koujiro Hara², Hajime Muraguchi¹, and Mieko Oguro-Okano^{1,3}

¹ Department of Biotechnology, Faculty of Bioresource Sciences, Akita Prefectural University,

² Department of Biological Production, Faculty of Bioresource Sciences, Akita Prefectural University,

³ Department of Animal Health Technology, Faculty of Animal Health Technology, Yamazaki Gakuen University

De novo transcriptome assembly of Illumina next-generation sequencing reads by Trinity offers the comprehensive transcript sequences in non-model organisms without building their genomes. Further gene ontology studies of annotated genes in combination with their expression profiling have enormously extended our knowledge of the molecular biological processes in non-model organisms. However, there are a huge number of non-annotated genes, so-called novel genes in certain transcriptomes termed dark transcriptome, which is usually ignored and cannot be accounted for in normal analysis. Barnacles express many functional secretory proteins to undergo larval cementation, adult cementation, juvenile shell formation, and adult shell formation, each of which is a building block of the unique barnacle lifestyle. These unique structures should be constructed by many unique novel proteins. Here, using the larval cement of *Megabalanus rosa*, Akafujituboa as a model, we have established a simple procedure to extract the genes responsible for encoding functional secretory proteins for life-specific events of barnacles.

Keywords: de novo transcriptome, novel genes, secretory protein, *Megabalanus rosa*