

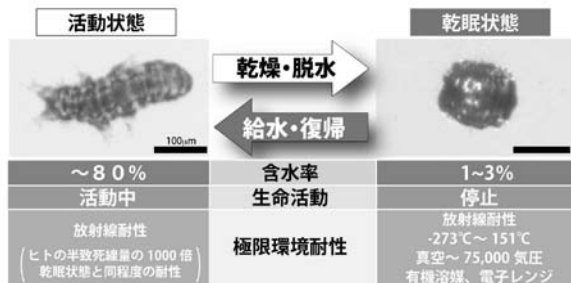
極限環境耐性動物のゲノム基盤の解析

●國枝 武和¹⁾ ◆片山 俊明²⁾ ◇豊田 敦³⁾

1) 東京大学大学院理学系研究科 2) 東京大学医科学研究所 3) 国立遺伝学研究所

<研究の目的と進め方>

クマムシは、極限的な環境耐性をもつことで知られる微小動物であり、分類学的には緩歩動物門という独自のグループを形成して、節足動物（昆虫など）と線形動物（線虫など）との間に位置するとされる。陸生種の多くは乾燥に曝されると脱水してほぼ完全に代謝を停止した乾眠と呼ばれる状態に移行し乾燥に耐える。乾眠状態では超高温・超低温・超高压・真空・放射線などに耐性を示し、これらのストレス曝露後も給水により速やかに生命活動を再開する。クマムシの持つ乾眠・耐性能力の解明は、産業応用（細胞や臓器の乾燥保存など）の観点から有用であるばかりでなく、進化によって生命が到達する極限の一つとして生物の進化多様性を理解する上でも重要である。しかしながら、乾燥耐性や極限環境耐性の分子機構には不明な点が多く、その解明が待たれている。本研究では近年効率的な培養が可能になったヨコヅナクマムシを用いて十分量の個体数を確保し、支援班の協力を得てゲノム配列の解読を行うとともに、乾眠時のクマムシの遺伝子発現プロファイルと比較することによって、乾燥耐性のゲノム基盤と関連する遺伝子セットを網羅的に明らかにすることを目的とする。本研究成果のゲノム配列とアノテーションは全てデータベース化し広く公開する。



クマムシの乾眠能力と極限環境耐性
写真は本研究に使用しているヨコヅナクマムシ(*Ramazzottius varieornatus*)

<研究開始時の研究計画>

1. ゲノム解析

ヨコヅナクマムシのゲノム概要配列の決定を目的に、ホールゲノムショットガン解析を行う。ゲノム配列の個体間変動の影響を避けるため、1つの卵に由来する純系統のヨコヅナクマムシ(YOKOZUNA-1系統)を確立し、解析にはすべてこの系統を用いる。ゲノム解析に必要な量のDNAを得るため、同系統を大量に飼育し成体約1万匹を準備する。餌がサンプルに混入するのを避けるため、2日ほど絶食させた後、大量の精製水で体表を洗浄しDNA調製用のサンプルとする。得られたゲノムDNAを用いて10x-15xのショットガン解析を行う。また、同時に Fosmid ライブラリも構築し、ペアエンドの配列を決定する。以上の解析で得られた配列情報から、コンティグとスキャフォールドを構築し、ゲノム概要配列を決定する。この配列をもとに、相同性検索による遺伝子推定、および ab initio 遺伝子予測を行うと同時に、

次項で行うトランスクリプトーム解析の結果を併用して遺伝子モデルの構築とアノテーションを行う。

2. トランスクリプトーム解析

乾眠に移行するためには1-数時間をかけて乾燥する必要がある。この間に乾燥耐性の準備をしていると考えられる。このため、乾眠時に変動する遺伝子は乾燥耐性に関与する良い候補と考えられる。そこでまず、活動・乾眠・復帰の3状態のヨコヅナクマムシについて遺伝子発現プロファイルを解析し、これらを比較することで、乾眠時に変動する遺伝子を網羅的に同定する。遺伝子発現プロファイルは網羅的なEST解析によって行う。具体的には3状態それぞれについて完全長cDNAライブラリを構築し、各5-10万ペアリードを解読する。また、得られた配列データはゲノム概要配列の遺伝子アノテーションにも利用する。完全長cDNAライブラリの構築には1状態につき約1万匹のヨコヅナクマムシが必要と算定されたため、トランスクリプトーム解析のために計3万匹のYOKOZUNA-1系統を準備する。

・データベースと情報解析

本研究で得られたゲノム配列と遺伝子データ、アノテーションは統合してゲノムデータベースを構築する。ゲノム配列から推定されたそれぞれの遺伝子について、KEGGやGene Ontologyを利用した概略的な機能アノテーションを行うとともに、Pfamによるドメイン構成のアノテーションや、Rfamによる非コードRNA遺伝子のアノテーションを行い、それらの結果もゲノムデータベースに統合する。

<研究期間の成果>

1. ゲノム解析

・ホールゲノムショットガン解析

まず、単一個体に由来するヨコヅナクマムシの純系統(YOKOZUNA-1系統)を確立し、これを必要量に達するまで大量に飼育した。これを材料として、ゲノム解読に足る品質の長鎖ゲノムDNAを得た。クマムシの餌であるクロレラが試料に混入するのを避けるため、ゲノムDNAの調製直前に2日間の絶食を行っている。この時、同時に抗生物質を添加することで細菌の繁殖を抑制した。クマムシは独自の動物門を形成し近縁種のゲノム情報が利用できないことから、1リード長が長いサンガー法を用いてホールゲノムショットガン解析を行った。これまでに約77万リード、951Mbのショットガン配列を決定した。また、並行して Fosmid ライブラリを構築し、ペアエンドを約4万リード、55Mbの配列を決定した。得られた配列は総計1,006Mbとなり、予想ゲノムサイズ~60Mbの16xカバレッジとなった。これらの配列をもとにPCAPでアセンブルした結果、コンティグが4,191本(58.3Mb)、N50は183番目の90kbであった。GC含量は47.3%であり、顕著な偏りは見られなかった。スキャフォールド

は3,002本で、N50は7番目の3.6Mbとなった。総塩基長は58.8Mbとなり、予想ゲノムサイズの98%に相当した。得られたスキファフォールドのうち、トップ22個の総塩基長が55Mbと全体の94%を占めており、少数のスキファフォールドでゲノムの大部分をカバーしていることが分かった。また、イルミナナの高速シーケンサーによるゲノムの追加解読を行い、これまでにおよそ9Gbのショートリード配列を得た。現在は、高精度のゲノム配列の決定に向けて、ショートリードのプレアセンブルと、サンガー法の配列データと組み合わせたハイブリッド解析を行っている。(ゲノム配列の解読は、基盤ゲノム支援班との連携により行った。)

・遺伝子モデルの構築と解析

スキファフォールドの配列をもとにExonerateによる相同性ベースの遺伝子予測を行ったところ、およそ2,000前後の遺伝子が予測された。2,000前後では遺伝子数として少なすぎると考えられたため、計算機による、*ab initio* 予測を行った。予測プログラムとしてGlimmerとSNAPを比較した結果、SNAPによる遺伝子予測が既知遺伝子との整合性が高いことが分かったため、以降の解析にはSNAPによる遺伝子モデル約15,000個を用いた。これらのうち、他の生物種と相同性を示すものは2,000-3,000個程度であり、その他の多くはクマムシ固有の遺伝子である可能性がある(遺伝子予測は慶應大・荒川和晴博士との共同研究)。相同性を示した遺伝子についてKEGGを利用した概略的な機能アノテーションを行った。その中で、クマムシの耐性能力と関係するものについて部分的に解析した結果、DNA2本鎖切断の修復に関わる遺伝子の1つが他生物種に比べてゲノム上のコピー数が増加していることが見いだされた。これはクマムシの高い放射線耐性を支えるゲノム基盤の1つである可能性がある。この遺伝子は同一スキファフォールド上の3遺伝子座にコードされていたが、互いによく似た塩基配列をもち、特に2遺伝子座は非常によく似ており、1.5kbのほぼ中間に1塩基しか相違がない箇所もあった。この遺伝子座のアセンブルにはサンガー法による長いリード長が大きく寄与したと考えている。

機能プロテオミクスから、100℃に加熱しても可溶性を維持するクマムシのタンパク質が5種見いだされたが、これらはいずれも他生物種とは相同性が低いタンパク質であり、クマムシゲノムの配列情報を用いて初めて同定された。これらの遺伝子と相同性を示す遺伝子はクマムシのゲノム上に多数コードされていることが分かり、3つの遺伝子ファミリーを構成することが判明した。また、これ以外にも、他生物種とは相同性が低い、クマムシゲノムではファミリーを形成している遺伝子群が見いだされた。これらのクマムシ特有の遺伝子群は、極限環境耐性などクマムシ固有の性質に関与している可能性がある。

2. トランスクリプトーム解析

・mRNA-seq解析

活動・乾眠・復帰の各状態の遺伝子発現プロファイルについて、研究開始当初は3状態すべてについて完全長cDNAライブラリの構築と5-10万ペアエンドリードを計画していたが、次世代シーケンサーを用いたトランスクリプトーム解析(RNA-seq)の手法が確立されてきたこととクマムシゲノム解析が順調に進行していたことから、mRNA-seqによる手法に切り替えた。非常に高速に解析できるため、復帰状態についてタイムコースを2点に増やし、計4状態についてmRNA-seq解析を行った。現在までに、各状態について約3Gbの配列を得た。ショートリード配列だけを利用したde novoアセンブルのほか、前項で決定したゲノム配

列にリードをマップして遺伝子モデルの構築も行っており、両者を比較して精度の高いトランスクリプトームを決定する。発現量の変動の解析もあわせて進行中である。

・完全長cDNAライブラリ

クマムシの転写産物は耐性能力に関わる遺伝子の資源としても有用と考えられる。そこで、mRNA-seq法に加えて、転写産物の機能解析や機能によるスクリーニングに供するために、乾眠状態のクマムシ約2万匹を材料として完全長cDNAライブラリを構築した。現在、約1万クローンの末端配列を決定中であり、これらの情報を用いてアノテーションの検証と修正を行う予定である。(完全長cDNAライブラリを構築および末端配列の決定は基盤ゲノム支援班との連携により行った。)

・LEA 遺伝子

トランスクリプトーム解析において、クマムシ類から初めてLate Embryogenesis Abundant (LEA) 遺伝子を同定した。LEA遺伝子は植物の種子や、乾眠能力を持つ一部の動物で乾燥時に発現上昇する遺伝子で、乾燥耐性への関与が指摘されている。今回同定した遺伝子RvLEA-1は動物界で初めての分泌型LEAタンパク質をコードしており、リコンビナントタンパク質は高熱処理しても沈殿しないというLEAに特徴的な形質を示した。

3. データベース

本研究で解析した情報を利用できる形で統合するために、クマムシゲノムプロジェクトのWebサイトを開設し、各種データベースを構築した(<http://kumamushi.org/>)。ゲノムブラウザでは、スキファフォールドをベースとして、各コンティグやリードの位置関係を確認できるほか、予測遺伝子モデル(SNAPによる予測コーディング領域、rRNA、tRNA、リピート)との対応、KEGGを用いた機能アノテーション、SwissProt/UniRefなど他生物種との相同部位などもあわせて表示できる。解析を支援するため、本研究で解読した各種配列データに対するBLAST検索サーバも設置した。また、Pfamを用いた網羅的なモチーフ検索の結果もデータベースとして検索できるようになっている。モチーフ検索は、SNAPによる予測遺伝子モデルのほか、全ゲノム配列の6フレーム翻訳産物に対しても行った。その他、関連する写真・動画などの情報も同サイトで提供している。本研究成果の発表にあわせて、これらのデータベースも公開する予定である。

4. その他・予想外の成果

・大量飼育法の改良

クマムシの飼育は労働集約的な作業であり、大規模解析においても材料の供給が最大の律速段階となる。そこで、本研究期間においても継続的に飼育方法の改良を押し進め、飼育系の大幅なクリーン化と省力化を実現した。具体的には、次亜塩素酸ナトリウムを用いた卵表面の殺菌工程を追加したほか、クマムシ個体群をまとめて扱えるように方法を改良し飼育作業を効率化した。これにより、サンプルの最大供給能力は当初の約2-3倍に向上した。

・アルビノ変異体の単離

本解析に用いているヨコヅナクマムシは通常茶色い色素を蓄積しており、外部から体内の構造を観察するのは困難である。大規模解析にむけてクマムシを大量に飼育する過程で、偶然白いクマムシ個体を発見した。単離して飼育したところ、この形質は子孫にまで継承されており、アルビノ変異体が得られたと考えられた。この変異体では、体内の観察が格段に容易であり、蛍光染色

にも有用であることが分かった。また、後段で述べるような変異体解析のパイロットケースとしても活用できると考えている。

<国内外での成果の位置づけ>

本研究は、乾眠能力・極限環境耐性を持つ動物として初のゲノム解析であり、クマムシに限らず、動物の持ちうる耐性能力の研究としてプレゼンスは大きい。クマムシは高い耐性能力を持ちながら、他生物と比較して分子機構はほとんど解明されていなかった。本研究は、本邦由来のヨコヅナクマムシを極限環境耐性動物のモデルとして利用する基盤を提供するものであり、今後の解析の進行に大きなインパクトを与える成果である。

本研究の成果について、2つの国際シンポジウムで発表したところ、大変大きい反響を得た。海外の研究室でも、YOKOZUNA-1系統ではないが、同じ学名のクマムシを研究対象に加えたところも既に現れている。

<達成できなかったこと、予想外の困難、その理由>

・予想外の困難

本計画に必要な個体数は、事前の予備実験から算定していたが、ゲノムDNAの大量調製の過程で、予備実験に用いた手法では十分な品質の試料を得ることができないことが分かった。このために、最初に約3ヶ月かけて準備した1万匹のクマムシ個体からは試料を得ることができず、その後も1,000匹単位のクマムシを用いて調製法の検討をする必要が生じた。最終的に、当初想定3倍程度の個体数を消費し、その準備に予想外の時間がかかることになった。この障害の影響は、大量飼育法の改良により幾分緩和することができた。

<今後の課題、展望>

今後は、大量のショートリード配列を利用してゲノム配列の高精度化を進めるとともに、mRNA-seqのデータを核とした遺伝子モデルの再構築とアノテーション、発現変動の解析を行う。本研究で必要となるデータはこれまでの解析でほぼ得ることができたと考えており、今後はこれらのデータをまとめあげることが課題である。

本研究で解読したゲノムとトランスクリプトームの結果は、並行して解析が進んでいるプロテオーム解析やメタボローム解析など、各種オミクス解析の解釈を強力に支援する基盤になると考えられる。特にゲノムは生命現象の基礎であり、他のオミクスで得られた情報をゲノムの上に統合していくことで、クマムシの耐性能力の分子基盤を多面的にとらえることができるようになるだけでなく、こうした基盤の進化的起源の解析にも結びつくことが期待される。

本研究によりクマムシのゲノムには、他の生物種には無い固有の遺伝子が多数存在し、その一部はクマムシゲノム内でパラログとしてファミリーを形成していることが分かった。加熱処理で沈殿しないタンパク質群もこうしたクマムシ特有の遺伝子ファミリーに属していることが判明したことから、耐性能力に関わる遺伝子としてクマムシ特有の遺伝子ファミリーは良い候補である可能性がある。今後はこうした遺伝子の機能解析が有望な手法と考えられる。

クマムシの極限環境耐性のように、解析の手がかりの少ない生命現象には、変異体を用いた遺伝学が有力なツールとなりうる。しかしながら、単為生殖を行うヨコヅナクマムシのような種においては、変異体を単離できたとしても連鎖解析による原因遺伝子の同定は不可能である。近年、高速シーケンサーの発展により、リファレンスゲノムがあれば、変異体と野生型の全ゲノム比較が

実現可能なものになってきた。今回、ヨコヅナクマムシの純系統(YOKOZUNA-1)についてゲノム概要配列が決定できたことから、クマムシについても同様の戦略を適用することが考えられる。本研究では、偶然ではあるが、YOKOZUNA-1系統に由来する初の変異体としてアルビノ変異体を単離することができた。クマムシの変異体解析のパイロットケースとして利用できるのではないかと考えている。

クマムシの乾眠は、生命活動という動的なシステムが乾眠という静的状態と可逆的にスイッチするという現象であり、特に乾眠からの復帰は生命活動のない物体から生命活動が開始する過程でもある。こうしたメカニズムの解析から、システムとしての生命の理解や、さらには人工生命の創出へと研究を展開していければと期待している。

<研究期間の全成果公表リスト>

1) 学会発表

1. Kunieda, T. *et al.*: The tardigrade genome of an anhydrobiotic extremotolerant species, *Ramazzottius varieornatus*. 11th International Tardigrade Symposium, 5 Aug 2009, Tubingen.
2. Katayama, T. *et al.*: Draft genome sequence assembly and preliminary annotations of *Ramazzottius varieornatus* genome. 11th International Tardigrade Symposium, 5 Aug 2009, Tubingen.
3. Horikawa, D. *et al.*: Life history of the tardigrade *Ramazzottius varieornatus* under artificial culture conditions. 11th International Tardigrade Symposium, 5 Aug 2009, Tubingen.
4. Kunieda, T. *et al.*: The tardigrade genome of an anhydrobiotic extremotolerant species, *Ramazzottius varieornatus*. 3rd International Symposium on the Environmental Physiology of Ectotherms and Plants, 24 Aug 2009, Tsukuba.
5. 國枝武和: 極限環境耐性動物クマムシのゲノム解析 第10回極限環境微生物学会、2009年10月29日、東京
6. 桑原宏和: 極限環境耐性動物ヨコヅナクマムシにおけるアルビノ変異株の単離 第10回極限環境微生物学会、2009年10月29日、東京
7. 山口理美: 乾燥耐性を持つクマムシ由来の抗凝集性タンパク質の解析 第10回極限環境微生物学会、2009年10月28日、東京
8. Kunieda, T. *et al.*: Genome analysis and functional proteomics of anhydrobiotic extremotolerant tardigrade, *Ramazzottius cf. varieornatus*. 第32回日本分子生物学会、2009年12月10日、横浜
9. Katayama, T. *et al.*: Construction of the genome database and draft genome sequence analysis of tardigrada. 第32回日本分子生物学会、2009年12月11日、横浜
10. Kuwahara, H. *et al.*: Identification of an albino mutant strain in brown anhydrobiotic tardigrade, *Ramazzottius cf. varieornatus*. 第32回日本分子生物学会、2009年12月11日、横浜
11. Yamaguchi, A. *et al.*: The analysis of anti-aggregation proteins from tardigrades with desiccation tolerance. 第32回日本分子生物学会、2009年12月9日、横浜

2) データベース/ソフトウェア

1. クマムシゲノムデータベース (公開準備中)
<http://kumamushi.org/>