

大量 DNA シーケンシングと生命システム比較解明への応用

●小原 雄治

情報・システム研究機構 国立遺伝学研究所生物遺伝資源情報総合センター

<研究の目的と進め方>

ゲノム解析の進展は遺伝子からゲノムへというパラダイムシフトをもたらした。遺伝子の機能は、相互作用する相手（ゲノム環境）によって変わりうるのであり、その組み合わせの数は莫大になる。これがゲノムが高次かつ多様な生命現象をつくりだすことの基本である。このことは、ゲノムを単位として研究することにより、生命現象の素過程を統合し生物を形作り働かせる仕組み（生命システム）や生物個体、環境との相互作用により進化・多様化を生み出す仕組み（生物システム）の解明にアプローチできることを示している。

本研究では *C.elegans* 線虫という基軸生物のシステムを徹底的に明らかにし、近縁線虫のゲノム機能・発現と比較することにより、それら生命システムの構造、進化・多様化を解明することを目的とする。先端レベルの研究に基づく研究支援という基盤ゲノムの役割を明示するためにも、正面突破の研究を進めたい。

<研究開始時の研究計画>

1) 線虫 *C.elegans* の解析

・線虫 *C.elegans* ゲノムについて発現パターン（mRNA、タンパク質）・機能パターンの徹底的解析を続け、発現パターンのクラスタリング解析などから、発生過程の遺伝子カスケード抽出をおこない、この結果をもとに発現制御領域の体系的同定をおこなう。さらに制御領域への結合タンパク質の同定を進め、関与する遺伝子群の全貌、さらには遺伝子カスケードの全貌解明をめざす。いくつかの現象、例えば初期胚における翻訳制御・局在化のメカニズム全貌の解明や、そのことによる細胞運命決定機構の解明もおこなう。

・これらの解析のために、発生パターン（細胞系譜や細胞配置など）及び発現パターンの自動解析ソフトウェアを開発・改良する。さらに、以上の結果をもとに発生のコンピュータモデル化/シミュレーションをおこなう。

2) 比較システム解析

・近縁線虫のうち、まず *Diploscapter coronatus* (100Mb程度) のドラフトゲノム配列を支援班において決定し、*C.elegans* とのオルソログ遺伝子を同定し、DNAチップや *in situ* ハイブリダイゼーション法、RNAi などにより、それらの発現パターン、機能パターンを比較する。

・D02との連携により完全長cDNAライブラリーを作成し、トランスクリプトーム解析をおこない、*C.elegans* と比較する。

・以上により比較システム解析の基礎データを得る。

<研究期間の成果>

1) 線虫 *C.elegans* の解析

A) NEXTDBの拡張と *C.elegans* トランスクリプトーム

2004年度までに136,466のcDNAクローン（λ-ZAPおよびオリゴキャップ）についてEST解析を行い11,904のユニークなcDNA種を得ていた。本研究では、V-capping法により発生時

期ごとの胚および雄をサンプルとしてcDNAライブラリーを作成し、得られた150,041クローンについてEST解析を行い、新たに4,909のユニークなcDNA種を得た。これらEST配列はゲノム上に割当てることによって、既知遺伝子や既予測遺伝子と対応付けた。

分類同定されたcDNA種について whole mount *in-situ* hybridization 法による発現解析を継続し、最終的に11,309種について、発現ステージおよび位置を同定する画像を取得した。これら画像について、手動によるアノテーション付けを行い、検索可能な形式に整理した。また、発現パターンより選択したcDNA種についてRNAi解析を継続し、920種について表現型画像を取得した。

上記の結果はNEXTDBとして統合され、WormBaseをはじめとする外部リソースとのリンクを加えた形で公開されている。

また、本データベースを発現パターンから検索するためのエンジンとしてSearchEXを開発した。

さらに、次世代型シーケンサ SOLEXA を用いた全トランスクリプトーム解析を試みた。具体的には、胚、幼虫、成虫期の雌雄同体と、雄リッチな混合時期の計4種のmRNA試料をランダムヘキサマでcDNA化し、SOLEXAによりシーケンスし、各試料600-1000万、計3500万リードを得た。これらをゲノム配列にマッピングすることで、エキソン構造の推定と同時に、マップ可能リード数を指標に発現レベルのステージ間/性別間比較を行った。(図1) ここで、リードを有効に活用するため、エキソン構造を考慮し長いギャップを許容してマッピングを行うプログラム、および、スプライスリーダー配列をトリミングしてマッピングを行うプログラムを開発した。その結果、2900万リードをゲノム上にマップすることができた。また、全トランスクリプトームのアセンブルをVELVET (Zerbino, EBI) を用いて試行した。得られたコンティグをゲノム配列にマッピングし、その結果をWormBaseに登録された遺伝子モデルと比較することで、その修正を試みた。その結果、60遺伝子について新たなエキソンの存在が示唆され、92遺伝子についてWormBaseモデルと異なるスプライシングパ

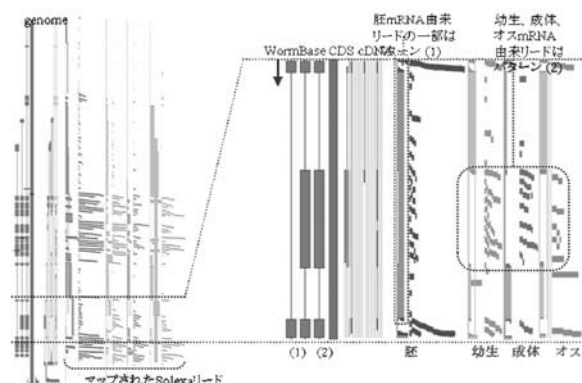


図1 ショートリードによるトランスクリプトーム解析例：tag-204

ターンの存在が示唆された。

これらの結果は、ショートリード/コンティグ専用ビューアを開発し NEXTDB に導入することで、同データベースへ統合された。

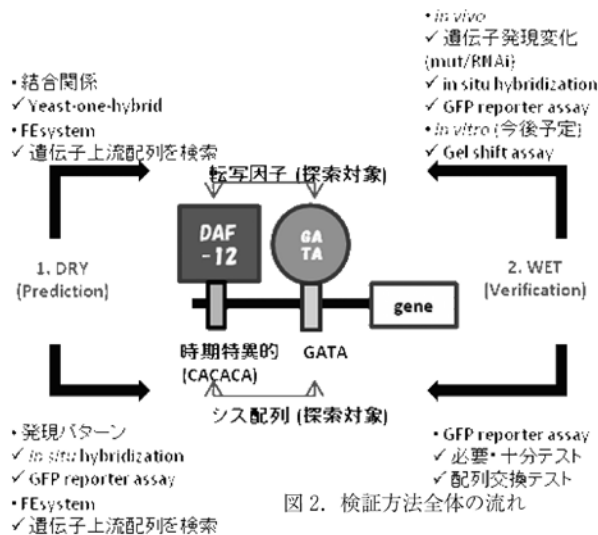
NEXTDB 公開 URL <http://nematode.lab.nig.ac.jp>

B) 発現制御領域の体系的同定、発現制御ネットワークの解析

i) 腸系譜の時期特異的なモチーフの予測と同定 (森明弘)

発生の遺伝子ネットワークシステムの解明には、転写制御を知る必要がある。時期・部位特異的な転写制御は、遺伝子の特定の領域 (モチーフ) にタンパク質 (転写因子) が結合することで引き起こされるため、転写制御メカニズムの解明には、その制御に関わるモチーフと転写因子の同定が必須である。

本研究の目的は、*C. elegans* (線虫) の初期胚で、共通の発現パターンを示す複数の遺伝子群の上流配列比較から時期・細胞特異的制御を行なうモチーフと転写因子の同定である。比較的単純な多細胞生物である線虫を用いても、予測プログラムで複数のモチーフを予測・同定した例もない。そこで、私達は、複数のモチーフを同時に検出し、また複数の遺伝子群に対しても適用できる「FEsystem」を開発してきた。



先行研究から線虫の腸の発生には、複数の GATA 因子が、時期特異的に複数の遺伝子の活性に働いていることが知られているが、その GATA 因子がどのように適当なターゲット遺伝子を見分けているのかは謎である。今回は FEsystem を用いて、GATA サイト以外に、線虫の初期胚での腸細胞で複数遺伝子を時期特異的に制御するモチーフを同定することにした。本研究室の *in situ* hybridization database から 3 つの連続した原腸陥入期 (E2-,E4-E8- 時期) で発現が開始する遺伝子群に、FEsystem でモチーフを予測し、GFP レポーターアッセイ等の検証実験を行なった。結果として、GATA サイトではないモチーフを各時期から同定できた。興味深いことに、同定したモチーフを交換することで、腸系譜での遺伝子発現開始時期を調節できることより、転写には時期特異的に制御するメカニズムが存在することを示唆している。

さらに転写因子候補同定するため、FEsystem で線虫の yeast-one-hybrid のデータセット (Deplancke *et al*, 2006) の中から転写因子結合サイトを予測した結果、それぞれの時期で転写因子候補が見つかった。特に興味深い発現パターンを示した E8 時期のモチーフ (DAF-12 結合サイトに類似) について詳細な解析を行った。GFP レポーターアッセイを *daf-12* 変異体で行ったところ、

E8 時期のモチーフを改編させた construct の GFP 発現と酷似したパターンを示した。さらに、この変異体株で E8 時期特異的な発現パターンの異常は、内在性遺伝子の発現パターンでも異常になることを *in situ* hybridization アッセイで確認した。しかし、この変異体株で E4 遺伝子を発現させても、これらの異常発現パターンは見られなかった。これらの結果は、DAF-12 が E8 時期特異的に転写制御に寄与していることを強く示唆している。また、GATA 因子との関係を調べるため、RNAi を行ったところ、ELT-2 が発現に寄与していることが分かった。以上のことから、時期・部位特異的な転写制御には、「部位」を制御している GATA 因子と「時期」を制御している DAF-12 の組み合わせで発現が制御されていることが示唆された。最後に本研究の重要性について 2 つ述べる。第一に、発現パターンを用いて単一種からの CRM を予測・同定することが可能になった。これは FEsystem の予測精度が高いため、他の高等動物への応用も可能であることを示唆している。第二に、プロモーター配列のデザインの可能性を示唆する。これは、遺伝子のプロモーター領域を解析し、遺伝子発現の時期・場所を予測するだけでなく、プロモーター自体を設計することで、研究者が望む場所・時期で遺伝子の発現が可能になることを示唆している。

ii) 温度感受神経 AFD での発現を制御する調節配列の系統的解析 (鹿兒島 浩)

生物は一次元情報に過ぎないゲノム配列を使って、どのようにして空間的・時間的な遺伝子発現制御を行っているのだろうか? この謎に迫るため我々はモデル生物、線虫 *C. elegans* の温度感受神経 AFD を研究材料として選び、この細胞での特異的な遺伝子の発現制御機構の解明を試みている。

このために我々は、AFD 神経で特異的に発現している 11 種類の遺伝子プロモーター :GFP コンストラクトを構築、または収集した。続いてこれらのコンストラクトに系統的な欠損を導入し、AFD 特異的な発現に必要な最小限な調節領域 (50 bp~600 bp) を決定した。これらの過程で、全く同一の発現パターンを示すプロモーターであっても、転写調節メカニズムが異なる場合があること、神経細胞特異的な発現の制御には、転写の活性化だけでなく抑制調節もまた重要な役割を果たすこと、離れた位置にある複数の調節配列が必要になる場合があることを明らかにした。

この一連の解析の中で、我々は特にグアニル酸環状化酵素 *gcy-8*, *gcy-18* 遺伝子のプロモーターに着目した。これらは完全に同一の発現パターン (前コンマ期から成体に到るまで AFD 神経だけで排他的に発現) を示し、またそれぞれ 50 bp, 68 bp という非常に短い調節領域 (コア配列) だけで AFD 特異的な発現を再現できた。さらにこれらは遺伝子重複によって生じた遺伝子であり、その発現の制御機構もまた保存されている可能性が高く、比較解析の材料として最適であると考えられた。

温度感受神経 AFD では 4 種類の転写因子 (*ceh-14*, *ceh-23*, *nhr-38*, *ttx-1*) が発現しているが、このうち 2 つの転写因子、LIM 型、及び OTX 型ホモボックスである *CEH-14*, *TTX-1* をコードする遺伝子の変異だけが、この神経の機能である温度走性に異常を示す。そこでこれらの変異体での *gcy-8*, *gcy-18* プロモーター :GFP の AFD 神経での発現を観察したところ、*ceh-14* または *ttx-1* どちらか一方の変異体では発現レベルがやや低下に過ぎなかったが、これらを二重変異にすると発現が完全に消失した。この結果を受けて次に、これらの転写因子を別の臭覚受容神経 AWB に強制発現させたところ、本来 AFD 神経だけでしか発現しない *gcy-8* と *gcy-18*:GFP を、AWB 神経で発現させることに成功した。このことから、*gcy-8*, *gcy-18* の発現を制御する上流転写因子は *ceh-14* お

および *tbx-1* であり、この両者が共同して転写制御を行っている可能性が示唆された。

近年、精力的に進められている *Caenorhabditis* 属の線虫のゲノムプロジェクトによって *C. elegans* に加え 4 種類の近縁線虫 *C. briggsae*, *C. brenneri*, *C. remanei*, *C. japonica* の配列データが得られた。これらを元に現在、これら 4 種類の線虫の *gcy-8* と *gcy-18* の相同遺伝子の上流領域についての発現解析を行ったところ、*C. elegans* の場合と全く同様に AFD 特異的 (排他的) な発現を示すことが判った。また、これらの近縁線虫のゲノム配列データを使った比較解析では、*gcy-8*, *gcy-18* 相同遺伝子群の上流配列中に種間で高度に保存された配列があり、実験的に求めた *gcy-8* の 50 bp、および *gcy-18* の 68 bp のコア配列の一部と一致した (図 3)。そこでこの保存配列を欠損、または置換したプロモーター-GFP コンストラクトを作成し、AFD 神経での発現を解析したところ、保存配列への変異導入は発現に大きな影響を与えることが判った。我々はこの配列が CEH-14、TTX-1 タンパク質の結合配列である可能性が高いと考えており、現在、精製 TTX-1、CEH-14 タンパク質を使ったフットプリント、ゲルシフトなどの *in vitro* 実験系で検討を行っている。

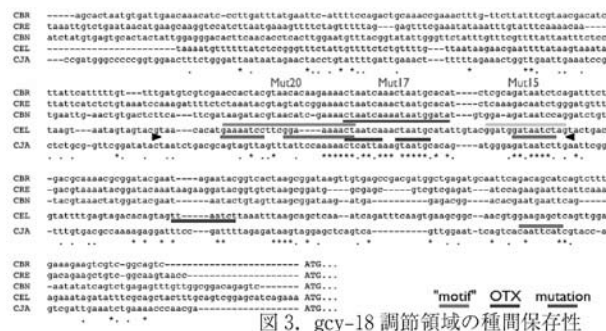


図 3. *gcy-18* 調節領域の種間保存性

iii) 内胚葉誘導における Src シグナル経路の解析 (住吉栄輔)

Src シグナル経路は分化、増殖、細胞接着等の制御に寄与し、ヒトの癌化に於いても重要な役割を果たしている。我々は Src シグナルの解析を線虫 *C. elegans* の内胚葉誘導の系を用いて行った。*C. elegans* の内胚葉誘導は、4 細胞期に於いて P2 割球より EMS 割球の後ろ側に対して Src および、やはり分化と増殖に寄与する Wnt シグナルが伝わり、EMS の娘細胞のうち後ろ側に位置する E 割球において内胚葉遺伝子群の発現を誘導する事により起こる。内胚葉誘導に於ける Wnt シグナルについては多くの知見が得られているが、Src シグナルについてはチロシンキナーゼレセプターである MES-1 によって P2 割球から EMS にシグナルが伝えられ、これに応じて SRC-1 が E 内でシグナルを伝達する事しか判っていない。我々は NEXTDB より抽出した初期胚で発現している遺伝子を RNAi によって破壊し遺伝子破壊胚の初期発生を録画したライブラリーより、Wnt および Src シグナルに関連する可能性のある遺伝子を検索した。その結果 β -catenin をコードする *hmp-2* 遺伝子が内胚葉誘導関連遺伝子として同定された。

多くの生物で β -catenin は Wnt signal の刺激によって核移行し、転写因子 Tcf を活性化する事で下流の遺伝子発現を制御する。しかし、*C. elegans* の内胚葉誘導では他の β -catenin である WRM-1 が Wnt シグナルに応じて Tcf である POP-1 の E に於ける核外移行を促進する事でシグナル伝達を担っており、HMP-2 タンパクは細胞接着に特化し、シグナル伝達の機能を持たないとされていた。

遺伝学的解析により、HMP-2 は *C. elegans* の Src である SRC-1 の下流に於いて内胚葉誘導シグナルの伝達に貢献している事がわかった。*hmp-2* を Wnt シグナル構成員と同時に破壊する事

で、Wnt シグナル単独の破壊に比べて高い割合で E の核での POP-1 の増大をもたらした。この時、E の子孫細胞に於いて中胚葉の遺伝子である *tbx-35* の発現が起こり、内胚葉の遺伝子である *elt-2* の発現が阻害され、内胚葉誘導が阻害されていた。HMP-2 は E の核に多く、E の姉妹であり、中胚葉を形成する MS 割球の核には弱く局在する事を明らかにした。*src-1(RNAi)* 胚ではこの HMP-2 の核局在が阻害されていることから、HMP-2 は Src シグナルに応じて E の核に移行し、POP-1 の核外移行を促進していると考えられる。

多くの生物で、 β -catenin は Src によるチロシンリン酸化によって、細胞膜上に局在して細胞接着を担うカドヘリン複合体から遊離する事が知られている。リン酸化チロシンは HMP-2 と同様に SRC-1 に依存して E の核に多く局在し、*hmp-2* の RNAi によってこの局在が阻害される事を示し、さらに、HMP-2 が SRC-1 に依存して *in vivo* でチロシンリン酸化される事を確認した。また、HMP-2 に結合するカドヘリンである HMR-1 を破壊すると SRC-1 非存在下でも HMP-2 が核に局在し、Src および Wnt の両シグナルの破壊による内胚葉誘導阻害が抑圧される事が判った。これらの結果から、HMP-2 は E に於ける Src シグナルに応じたチロシンリン酸化によってカドヘリン複合体から遊離して核移行し、POP-1 の核外移行を通じて中胚葉遺伝子の発現抑制を行い、内胚葉誘導に寄与している事が明らかになった。

C) 翻訳制御・局在化のメカニズムの解明

i) *pos-1* mRNA 局在化機構の解析 (野口浩毅)

多くの多細胞生物において、初期発生過程は卵母細胞に蓄えられた母性 mRNA 等の母性遺伝子産物により制御されている。このため、初期発生過程の理解には母性 mRNA に対する転写後制御機構の解明が重要である。本研究では線虫初期胚において生殖細胞系譜へ局在化する母性 *pos-1* mRNA の局在化機構に着目した。POS-1 タンパク質は他の母性遺伝子の翻訳制御を通じて生殖細胞系譜の確立に働くが、*pos-1* mRNA の局在化はこの POS-1 タンパク質の局在制御に関与していると考えられる。本研究で我々は、1. 局在に必要な mRNA 上の配列 (シス因子) 同定のための *in vivo* アッセイ系の構築、2. *in vivo* アッセイ系を用いたシス因子の同定、3. シス因子に結合して局在を制御するタンパク質 (トランス因子) の探索を行った。

1. *in vivo* アッセイ系としては、遺伝子導入により検出用のタグ配列と *pos-1* 配列の融合 mRNA を発現する形質転換株を樹立し、*in situ* hybridization により融合 mRNA の局在を調べる方法を採用した。線虫では、形質転換体からの母性発現はサイレンシングにより抑制されるため、この抑制を受けない形質転換体を作成する方法として知られる Bombardment 法を茂木文夫、杉本亜砂子両博士の協力を受け導入した。また、*in situ* hybridization での検出に必要な発現量を確保するため、強力な *pos-1* プロモーターを単離・採用し、さらに、タグ配列である GFP を GC 含量が高いものへ変更する等の工夫をした。

2. シス因子の解析は欠損解析と置換解析により進めた。この結果、*pos-1* mRNA の局在化には 3' 非翻訳領域 (UTR) 上の近接した二つの保存配列 (CCCACA が二回繰り返すタンデムリピート配列、及び 30 残基の配列) が必要と判明した。他生物における研究から、mRNA の局在制御は多くの場合 3' UTR 上のシス因子により制御されている事が示されているが、この結果は線虫においても同様の機構が mRNA の局在制御に働いている事を示唆するものである。

3. トランス因子については、Yeast three-hybrid 法を利用し、2. で同定したシス因子と結合するタンパク質を線虫 cDNA ライブ

ラリーから探索した。得られた18個の候補遺伝子をさらに絞り込むため、RNAi処理個体または変異株を用いて、候補遺伝子の活性低下が *pos-1* mRNA の局在へ与える影響を調べた。この結果、候補遺伝子の一つ、*mex-1* が *pos-1* mRNA の局在に必要な事が分かった。2細胞期胚では、野生株では *pos-1* mRNA は生殖細胞系譜である後側割球 (P₁) へ明瞭に局在するが、*mex-1* 変異株では胚全体に一樣に分布し、局在化しなかった。これらの結果は MEX-1 タンパク質が *pos-1* mRNA 局在化のトランス因子である事を強く示唆している。MEX-1 タンパク質は *pos-1* mRNA と同様に生殖細胞系譜へ局在化する RNA 結合タンパク質であり、初期発生過程における細胞運命の決定に必要な事が分かっているが、分子的な作用機序はまだ詳しく調べられていない。今後の解析により、MEX-1 による *pos-1* mRNA 局在化の分子機構が明らかになる事が期待される。現在 MEX-1 と *pos-1* mRNA の結合を確認し、また結合配列等を詳細に解析するため、ゲルシフトアッセイを実施中である。

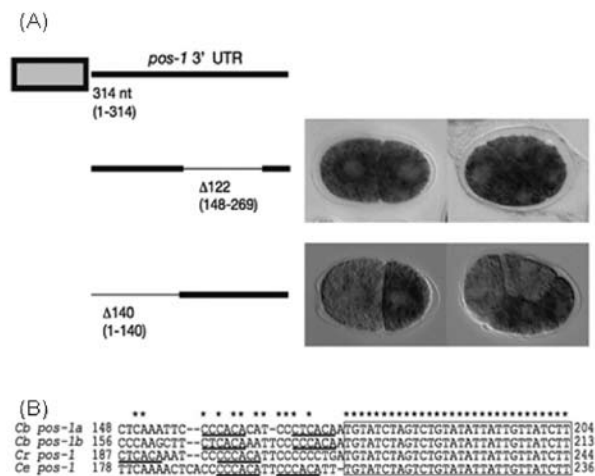


図4. *pos-1* 局在化シス因子の同定

欠損解析 (A) と置換解析 (省略) により *pos-1* mRNA 局在化シス因子の同定を行った。同定されたシス因子 (B) は近縁種間で高度に保存されていた。

(A)欠損解析による局在化シス因子の同定。*pos-1* 3'UTR後半を欠失した融合mRNA (Δ122)は2細胞期胚、4細胞期胚で局在しないが、*pos-1* 3'UTR前半を欠失した融合mRNA (Δ140)は2細胞期胚、4細胞期胚で内性*pos-1* mRNA同様に局在した。この結果に基づき、詳細な欠損解析と置換解析を行い、局在に必要なシス因子を同定した。(B) 同定されたシス因子の近縁種間での比較。シス因子はCCCACAのタンデムリピート(図中、下線部)。近縁種ではCYCACA)と、種間で高度に保存された30残基の配列からなる(図中、四角で囲まれた部分)。Cb *pos-1a*、Cb *pos-1b*は近縁種の*C. briggsae*の、Cr *pos-1*は*C. remanei*の、Ce *pos-1*は線虫*C. elegans*の局在化シス因子近傍の配列を示す。

ii) *mex-3* mRNA局在化のメカニズム (金野宏之)

mRNA の局在化は多くの生物において観察されている現象であり、線虫 *C. elegans* の初期胚においても特定の母性 mRNA が特定の割球へと局在化することが観察されている。しかしながら、酵母やショウジョウバエにおいて mRNA の局在化機構が精力的に研究されているのに対し、線虫においては mRNA の局在化機構の研究はほとんど手付かずの状態に残されており、線虫を用いた研究により、mRNA の局在化機構に関して新たな知見が得られることが期待される。

線虫初期胚において、前側割球へと局在化する母性 mRNA と

して有名なものに母性 *mex-3* mRNA があり、本研究では母性 *mex-3* mRNA の局在化機構について解析を行なった。

まず母性 *mex-3* mRNA の局在化に必要なシス因子の解析を行なった。GFP の垂種である VENUS の CDS の下流に *mex-3* 遺伝子の様々な部分配列を融合させたコンストラクトを作製し、これを biolistic transformation 法により、線虫の生殖腺で発現させ、その後 VENUS の antisense 鎖をプローブとした *in situ* hybridization を行なうことにより、導入コンストラクトから転写された VENUS 融合 mRNA を検出し、初期胚における局在パターンを調べた。

その結果、*mex-3* 遺伝子の 3' 側非翻訳領域上 (3' UTR) の 179nt が母性 mRNA の局在化に十分であり、179nt を欠失した *mex-3* 遺伝子から転写された VENUS 融合 mRNA は局在化しないことが分かった (図5)。この 179nt はステムループを構成し、特にループ付近を構成する 35nt が mRNA の局在化に重要であることが明らかとなった。この 35nt は線虫種間でも高度に保存されていた。

また、*mex-3* CDS::*mex-3* 3' UTR というコンストラクトから翻訳された MEX-3 タンパク質が初期胚において前割球へと局在化するのに対し、そのコンストラクトの 3' UTR から 179nt を欠失したものでは MEX-3 タンパク質の前割球への局在化が見られなくなるを見出した。このことは、母性 *mex-3* mRNA の局在化が MEX-3 タンパク質の局在化に重要であることを示唆しているものと思われる。

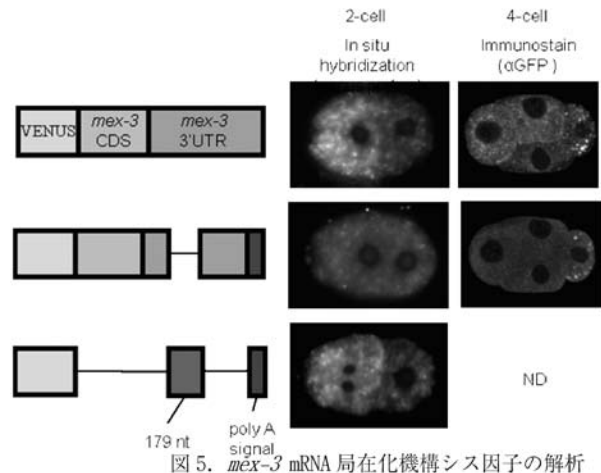


図5. *mex-3* mRNA 局在化機構シス因子の解析

一方、母性 *mex-3* mRNA の局在化のトランス因子候補の探索を行なうために、機能喪失により母性胚性致死を示す主な RNA 結合タンパク質の各変異体の初期胚に対し、*mex-3* の antisense 鎖をプローブとした *in situ* hybridization を行なうことにより *mex-3* mRNA の局在パターンを調べた。これまでのところ、*mex-3* mRNA の局在化が完全に消失するものは見つかっていないが、*mex-1* 変異体と *mex-5* 変異体において、*mex-3* mRNA の局在化が非常に弱まることを見出した。今後、トランス因子候補の探索をさらに進める予定である。

(D)モデル化に向けた胚の細胞配置データ取得 (平木秀明)

線虫 *C. elegans* は、透明で細胞数が少ないため、発生の過程を詳細に観察することができる優れた系である。発生においては細胞間相互作用が重要な役割を果たすので細胞の形状・配置・隣接関係の把握が非常に重要である。本研究では、細胞膜を蛍光染色して共焦点顕微鏡で観察した時系列3次元画像をもとに、計算機を用いて半自動的に細胞の形状モデルを構築するシステムの開発を進めてきた。

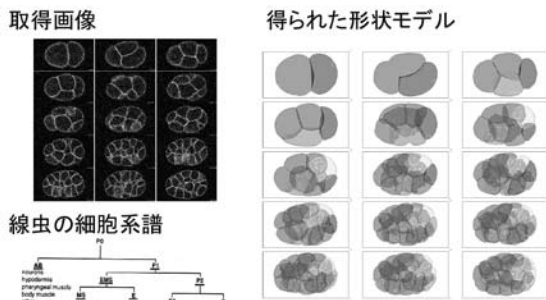


図6. 2-28細胞期の細胞形状モデル

このシステムは、ある時点における細胞形状を3次元画像上での領域拡張法により抽出するプログラムとそれに必要な種点を対話的操作で修正するためのユーザーインターフェースプログラムで構成されている。ここで得られる形状はボクセルで表現されていて画像のノイズによる細かな凹凸があるためそのままでは表面積が求められなかったが、形状を三角メッシュ表現に変換し動的輪郭法を用いて滑らかな細胞形状を抽出する後処理システムを開発し、隣接する細胞間の接触表面積の値などが得られるようにした。このシステムを適用して複数個体の形状モデルを構築するため、細胞膜に局在するGFP融合遺伝子を発現する線虫株の撮影条件を検討した。その結果、多光子共焦点顕微鏡を用いて撮影後も発生が進行し孵化するような条件で、1~28細胞期の胚の画像データが取得できるようになった。

正常発生する胚10個体の画像データから形状モデルを構築して解析したところ、初期胚の細胞間の隣接関係には、その接触表面積が増えるものと減るもの、増減が個体ごとに異なるものがあることが分かった。また、MS細胞と接触しないために新しい運命に誘導されないとされていたABarp細胞が実際にはMS細胞と接触していること、その接触表面積は誘導が起こるMS細胞とABara細胞の間の接触やMS細胞とABalp細胞の間の接触よりも小さいことが観察された。さらに、*par-2*遺伝子をRNAiによりknockdownした胚5個体についても同条件で撮影して解析し、姉妹細胞の体積の非対称性が*par-2*遺伝子のRNAiにより失われることを確認できた。

本研究により、発生しつつある胚全体の細胞の形状をコンピューター内にモデル化できるようになり、コンピューターグラフィックスを用いてさまざまに見方を変えて観察することや、各細胞の体積、隣接する細胞間の接触表面積、それらの変化などを数値化することができるようになった。得られるデータは、個体差や遺伝子変異による影響を含む時系列データである。このデータを細胞分化や遺伝子機能などの生物学的な知見と関連づけていくためには、大量のデータをうまく整理して取り扱い、解析に有用な規則性を見出す手法を考えていく必要があるだろう。一方データ取得に関しては、より定量的な解析をするために画像処理に用いるパラメータの有効範囲について検討しさらに自動化を進めること、より後期まで撮影するための条件を検討すること、細胞内の状態を示す核や細胞骨格の画像データと同時撮影すること、などが今後の課題として挙げられる。今後さらにシステムの改善を進め、他のRNAi等による変異体や近縁種間の比較に応用する予定である。

2) 近縁線虫のゲノム/cDNA解析

発生過程の多様性および可塑性に関してその遺伝的基盤を比較解析するために、原腸陥入までの卵割およびその配置が異なる近縁線虫 *Diploscapter coronatus* (E. Schierenberg, Univ. Koeln より供与) をサンプルとしてゲノムおよびcDNA解析を行った。

num clone	gene	CDS	description
3,063			(weak hit to <i>P.pacificus</i> genome)
2,335	<i>his-60</i>	F55G1.11	Histone
1,570			(novel)
1,522		MTCE.26	cytochrome oxidase subunit I
1,140		F57F5.1	cysteine protease
1,106		T01B11.4	ADP/ATP translocase
1,053	<i>act-1</i>	T04C12.6	actin
1,024	<i>gpd-2</i>	K10B3.8	
1,009		MTCE.31	cytochrome oxidase subunit II
855		F46H5.3b	
732	<i>rps-19</i>	T05F1.3	Ribosomal protein S19e
681	<i>rpl-19</i>	C09D4.5	60S ribosomal protein L19
644			(novel)
620	<i>his-12</i>	ZK131.6	histone H2A
610			(novel)
563			(novel)
563	<i>his-41</i>	C50F4.5	histone H2B
543	<i>his-41</i>	C50F4.5	histone H2B
532	<i>vit-4</i>	F59D8.2	
516			(novel)
495		R02F2.4	
491		C42D4.1	

表1. 高頻度に出現する遺伝子とその *C.elegans* ホモログ

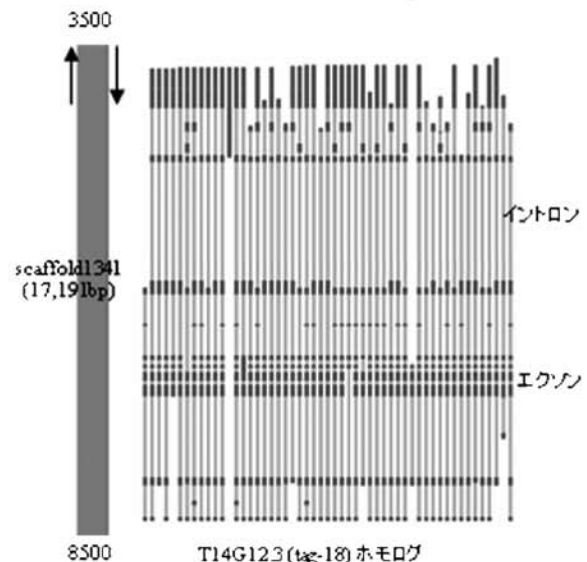


図7. スプライシングパターンが示唆される例

cDNA解析では通常法、オリゴキャップ法、V-capping法によりライブラリーを構築し、得られた計147,256クローンについてEST解析を行った結果、13,245のユニークなcDNA種に分類され、うち6,351については *C.elegans* にホモログが見つかった(表1)。一方、非常に高頻度(3,063/147,256)に出現するが、既知のタンパク質に相同性が見られないcDNA種も見られた。全長cDNAクローンの5' EST先頭を解析することにより、*D. coronatus* においてもスプライスリーダー配列が見られることが判明した。高頻度であったのは *C.elegans* におけるSL1、SL2と同一の配列であったが、そのほかにこれらと僅かに異なる配列パターンも一定の頻度で見られた。

ゲノム解析では、全ゲノムショットガン(WGS)法によるゲノム配列再構成を試みた。ショットガンクローンの両端をサンガー法によりシーケンスして120万リードを得、高クオリティ領域抽出およびベクタ削除を行って総延長657Mbasesを得たが、これは推定ゲノム長より×4相当と考えられる。本リードセットをPCAP (Huang, Iowa State Univ.) を用いてアセンブルし、36,000コンティグ(最長48,519 bases、N50 8,945 bases)、21,873 scaffolds(最長61,666 bases、N50 13,166 bases)、総延長

175Mbases のアセンブリを得た。本アセンブリ上より *C.elegans* ホモログを探索することで、EST 解析では未発見の 3,066 ホモログが見つかった。また、cDNA 解析で得られていた EST 配列をアセンブリ上に正確にアラインメントすることで、一部の遺伝子についてエキソン構造を得ることができ、さらに、複数のスプライシングパターンの存在を示唆する結果が得られた (図 7)。

ゲノムアセンブリの質的向上のため、次世代型シーケンサ SOLEXA リードを用いたハイブリッド法を試行した。具体的には、*D. coronatus* ゲノムを SOLEXA によりシーケンスし、得られた 1.7 億本、総延長 650Mbases のリードを VELVET (Zerbino, EBI) によりアセンブルし、得られたコンティグを WGS リードセットに加えて、全体を PCAP でアセンブルした。新アセンブリを、WGS リードのみを素材としたアセンブリと比較した結果、16,684 コンティグは伸長され、4,167 コンティグ組はその間がブリッジされたことが分かった。

上記のゲノム/cDNA 解析結果は、NEXTDB と併せて公開できるよう、WWW 上のデータベースとして整備した。

<国内外での成果の位置づけ>

・線虫 *C.elegans* トランスクリプトームについて世界のセンターとして、引き続き機能した。

<達成できなかったこと、予想外の困難、その理由>

・近縁線虫の研究については、RNAi の効きが非常に悪い。2 本鎖 RNA の導入に加え、siRNA, shRNA など試してみたがこれまでのところ効果が得られていない。一般に *C.elegans* 以外の多くの線虫は似た傾向であるが、このためホモログの機能解析が進んでいない。

・成果は多数得られたが、多くの成果が論文化の最中となった。

<今後の課題>

- ・NEXTDB の拡充
- ・体系的な発現制御ネットワークの解明
- ・母性 mRNA 局在・翻訳制御ネットワークの解明
- ・近縁線虫ホモログの機能解析系のたちあげ

<研究期間の全成果公表リスト>

1) 論文

1.1002062044

Tobias Langenhan, Simone Prömel, Lamia Mestek, Behrooz Esmaeili, Helen Waller-Evans, Christian Hennig, Yuji Kohara, Leon Avery, Ioannis Vakonakis, Ralf Schnabel, and Andreas P. Russ: Latrophilin signalling links anterior-posterior tissue polarity and oriented cell divisions in the *C. elegans* embryo. **Developmental Cell** 17(4), 494 - 504 (2009)

2.1002062046

Xin Wang, Yongjun Zhao, Kim Wong, Peter Ehlers, Yuji Kohara, Steven J Jones, Marco A Mara, Robert A Holt, Donald G Moerman and Dave Hansen: Identification of genes expressed in the hermaphrodite germ line of *C. elegans* using SAGE. **BMC Genomics** 2009, 10:213

3.1002062050

Kawano T, Zheng H, Merz DC, Kohara Y, Tamai KK, Nishiwaki K, Culotti JG.: *C. elegans* mig-6 encodes papilin isoforms that affect distinct aspects of DTC migration, and interacts genetically with mig-17 and collagen IV. **Development** 136, 1433-1442 (2009)

4.1002062052

James Douglas McGhee; Tetsunari Fukushima; Michael W Krause; Stephanie E Minnema; Barbara Gosczyński; Jeb Gaudet; Yuji Kohara; Olaf Bossinger; Yongjun Zhao; Jaswinder Khattrra; Martin Hirst; Steven J Jones; Marco A Marra; Peter Ruzanov; Adam Warner; Richard Zapf; Donald G Moerman; John M Kalb: ELT-2 Is the Predominant Transcription Factor Controlling Differentiation and Function of the *C. elegans* Intestine, from Embryo to Adult. **Developmental Biology** 327, 551-565 (2009)

5.0802072142

Hiroshi Kagoshima, Rachael Nimmo, Nicole Saad, Junko Tanaka, Yoshihiro Miwa, Shohei Mitani, Yuji Kohara and Alison Woollard: The *C. elegans* CBFb homologue BRO-1 interacts with the Runx factor, RNT-1, to promote stem cell proliferation and self-renewal. **Development** 134, 3905-3915 (2007)

6.0705142123

Kagoshima H, Shigesada K, Kohara Y.: RUNX regulates stem cell proliferation and differentiation: Insights from studies of *C. elegans*. **J Cell Biochem.** 2007 Apr 1;100(5):1119-30

7.0702251625

James Douglas McGhee Monica C. Sleumer; Mikhail Bilenky; Kim Wong; Sheldon J. McKay; Barbara Gosczyński; Helen Tian; Natisha D. Krich; Jaswinder Khattrra; Robert A. Holt; David L. Baillie; Yuji Kohara; Marco A. Marra; Steven J.M. Jones; Donald G. Moerman; A. Gordon Robertson: The ELT-2 GATA-Factor and the Global Regulation of Transcription in the *C. elegans* Intestine. **Developmental Biology** 302, 627-645 (2007)

8.0602111806

Elena Lazakovitch, John M Kalb, Reiko Matsumoto, Keiko Hirono, Yuji Kohara, Richard M Gronostajski: nfi-1 affects behavior and life-span in *C. elegans* but is not essential for DNA replication or survival. **BMC Developmental Biology** 2005, 5:24 (20 October 2005)

9.0602111726

Hiroshi KAGOSHIMA, Hitoshi SAWA, Shohei MITANI, Thomas R. BÜRGLIN, Katsuya SHIGESADA and Yuji KOHARA: The *C. elegans* RUNX transcription factor MAB-2/RNT-1 is required for asymmetrical cell division of the T blast cell. **Developmental Biol.** 287, 262-273 (2005)

2) データベース/ソフトウェア

NEXTDB <<http://nematode.lab.nig.ac.jp>>