

植物病原菌のcDNA Lottery (2)

- 平成13年度化学教室への講座配属3年生の研究報告 -

(植物病原菌 / cDNA / 塩基配列)

吉田 浩*・飯塚真理*・平林崇樹**

cDNA Lottery of a Phytopathogenic Fungus (2)

- A Report of Studies of the Third Year Grader Allotted to the Department of Chemistry in the Year 2001 -

(Phytopathogenic fungus / cDNA / Base sequence)

Hiroshi YOSHIDA*, Mari IIZUKA*, Takaki HIRABAYASI**

The third year grader in the year 2001 allotted to the Department of Chemistry carried out the following studies. The student was given arbitrarily a clone from the cDNA library of a phytopathogenic fungus *Fusarium moniliforme* and sequenced the insert of the clone. The clone contained a partial length cDNA encoding a hypothetical protein of unknown function. The protein has a highly charged region which presumably assumes an α -helix. Similar structural motifs have been reported in various proteins including mammalian caldesmon.

はじめに

われわれは昨年度の本紀要で、講座配属3年生向けに計画した実験と、その結果を報告した(1)。すなわち学生は、植物病原菌 *Fusarium moniliforme* のcDNAライブラリからめいめい任意の1クローンを選び、そのインサートの塩基配列を決定する。これをアミノ酸配列に翻訳し、得られた配列のホモロジー検索を行って、かかってきた既知タンパクの構造や機能について文献調査をする。そのねらいは、今や生化学・分子生物学の中心となったクローニング技術と、最近ますます重要性を増しつつあるバイオインフォマティクスを学習することにある。どんなタンパクに当たるかは予測できないので、これをcDNA lottery (cDNAくじ) と名付けた。平成13年度は配属学生が少なく、1名のみであったが、なかなか興味深いタンパクが当たったので、以下に報告する。

実験

実験方法、指導側の準備については昨年すでに述べたが、要約すれば次の通りである。植物病原菌 *F. moniliforme*

のcDNAライブラリから、指導側で予め2000 bp程度のインサートをもつクローンを選別しておき、学生に与える。学生はそのクローン・プラスミドをミニプレップし、制限酵素によりインサートを切り出して電気泳動でおよその長さを測定したのち、塩基配列を決定する。得られた塩基配列をアミノ酸配列に翻訳し、そのタンパクについてホモロジー検索を行う。かかってきた相同タンパクについて構造や機能の文献調査をする。それぞれの実験で学習する項目も昨年詳述した。

結果と考察

Lottery 5 (番号は昨年度からの通し)

電気泳動から推定したインサートの長さは1910 bp、塩基配列の決定では1801 bpで、両者は比較的良好一致していた。配列決定には3回の反応を要した。配列は日本DNAデータバンク (DNA Data Bank of Japan, DDBJ) に登録番号 (Accession number) AB092343として登録した。

このインサートには43 bの長いPoly(A) tailがあり、不思議なことにその中ほどにTが1つ挿入されている。この意味は不明で、あるいはcDNAの調製に用いたプライマーに由来するartifactかもしれない。このcDNAの特徴は5'末端側に続く長いプリン塩基A、Gに富む配

*化学教室

Department of Chemistry

**平成13年度3年生 Third Year Graders in the Year 2001

列である。こんなに偏った塩基組成のメッセージがありうるかと、半信半疑で翻訳したところ、開始コドンATGからではないが、終止コドンTGAで終わる451アミノ酸残基のタンパク読み枠 (1 - 1356) が取れた。したがって、何らかのタンパクをコードしている不完全長cDNAと予想された。5'末端側の塩基組成の偏りを反映し、タンパクのN末端側アミノ酸組成は極めて特異で、グルタミン酸 (E, コドンGAA, GAG) リシン (K, コドンAAA, AAG) およびアルギニン (R, コドンAGA, AGG) に富む高度に荷電した配列が現れた。ところが翻訳タンパクをホモロジー検索にかけたところ、この特異な配列がいろいろなタンパクにみられることがわかった。比較的相同性が高く、機能がわかっているウサギ平滑筋カルデスモンのSpace regionと比較した結果を図1に示す。150残基ほどの間に68残基、45.3%の相同性があり、明らかに関連タンパクである。

```

1' EAERKRKEEEERQRREEEKRKKQQR--EAEEKA-----RREEAEKQRIE
   **** * ***** . * * * * * ***** * * * * *
55" EAERGRLEEEERLKAQEKKRAEQARIEAEKKAQERERREAEEREKEE
44' EERKKKEEEQKQREEEERLHREQLER-EAAEEARLRREEERKERERER
   ** . * * * . * * * * * * * * * * * * * * * * * * * * *
105" RRREAEERERREKEERERREAEKERREAEERERREKEERERREKEERE
92' RAHREDMER-----KRAAREAE--QRRFREEQERIRLDKLPPLLRWLDTC
   * . * . * * * . * * * * * * * * * * * * * * * *
155" RREKEERERIKEEERRAAKEAATGQKGRKEKARLEEQKLLKQLEEKKT

```

図1. *F. moniliforme*のタンパクとウサギ (*Oryctolagus cuniculus*)のカルデスモンSpace regionのアミノ酸配列比較。後者の配列は、GenPept-updated (GPU) : translated protein sequence from GenBank-updatedより得た (Locus, AF421381; Author, Zhan E.Y. and Chacko S.; Direct submission(2001))。配列比較はFASTAの結果に基づき、ホモロジーのある部分のみ示した。上段、配列番号に'を付してあるのが *F. moniliforme*、下段、配列番号に"を付してあるのが *O. cuniculus*。後者は268残基からなる。*は同一残基、. は類似残基を示す。

しかし、カルデスモンは高等動物の平滑筋および非筋細胞においてCa²⁺依存性アクチン ミオシン相互作用を制御しているタンパクであり、カビにそのようなタンパクがあるとは考えにくい。また、相同性があるのは上述の高度に荷電した配列領域のみで、その他の部分ではまったく相同性がない。二次構造予測では、この高度荷電配列はヘリックスと考えられ、親水的ヘリックスという構造モチーフとしていろいろなタンパクに出現するものと思われる。比較は示さないが、最も相同性が高かったのは *Babesia bigemina* のp200という200 kDaの抗原であった。*B. bigemina*はウシなど家畜の赤血球に寄生する病原虫で、p200は報告された部分配列1108残基の中央部800残基ほどがE, K, Rに富む変わったタンパクである (2)。家畜への感染を診断するうえで有用なタンパクであるが、繊維タンパク

と予想されているだけで、機能はわかっていない。他にもこの構造モチーフをもつタンパクはシロイヌナズナ、ショウジョウバエ、ヒトなどゲノム研究が進んだ多くの生物で報告されているが、残念ながらほとんど機能未知で、遺伝子から推定されたHypothetical proteinの域を出ていない。おもしろいことに、微生物としてはコウボについて報告があるのみで、その他のカビ類や細菌では知られておらず、本報告が初めてである。

上述の特異な配列領域を除いてホモロジー検索にかけると、相同性は低いものの、たった1つだけかかってきたタンパクがあった。*Boophilus microplus*という、ウシにつくダニのエステラーゼ (3) で、配列比較を図2に示す。この場合も相同性は限定的で、543残基のエステラーゼの63-180位に24%ほどの相同性がみられるに過ぎない。関連タンパクでもこの程度の相同性しかないことはあるが、分子全域にわたる相同性ではないので、われわれのタンパクがエステラーゼとは断定できない。結局ここで見いだした *F. moniliforme* のタンパクの機能は推定すら難しいが、おそらく高度荷電ヘリックスを構造モチーフとしてもつ何らかの酵素ではなからうか。なお、DDBJへの登録にあたっては、内容を示すタイトルが求められる。そこで、“Cloning and Sequencing of a cDNA Encoding a Caldesmon-like Protein of *Gibberella fujikuroi*”とした。相同タンパク中、機能が明確だったのはカルデスモンだけだからであるが、やや不正確だったかもしれない。より正確には、“a Protein with Caldesmon-like Structural Motif” とすべきであったろう。なお、*Gibberella fujikuroi* は *F. moniliforme* の有性世代の名称である。われわれは無性世代の菌を用いているので、その名称 *F. moniliforme* をずっと使ってきた。本論文でもそれを用いたが、DDBJへの登録では、*G. fujikuroi* の使用を求められる。

```

336' EDKPLAQTSSSTPFLDKKVPKRRGLVQVLPD-----DPDYTRICLEQG
   ** . * * * . * * * * * * * * * * * * * * * *
61" FRKPSPKVSWVPRWTRTFGKRCMQEIVLDATEAHEYAEGPGMSEDCLYLN
378' LEHLINGHLSPVNVNGIHSSPMSQKSMSTSVGPPMNGMIKALTPGSTSESL
   . * * * * * * * * * * * * * * * * * * * * *
111" VWSPVNDNSTKPVVMVWIHGGNF--KAGSSDSPAFDGSVLAATQNVVVVVSF
428' TISNAGHETLVNGINGNTNGNPAH
   * . * * * * * * * * *
159" NY-RLGFFGFLNAVHGNASGNAGL

```

図2. *F. moniliforme*のタンパクとダニ (*Boophilus microplus*)のエステラーゼのアミノ酸配列比較。後者の配列は、PRF protein sequence databaseより得た (Locus, 2624222A; 文献 (3))。配列比較はFASTAの結果に基づき、ホモロジーのある部分のみ示した。上段、配列番号に'を付してあるのが *F. moniliforme*、下段、配列番号に"を付してあるのが *B. microplus*。後者は543残基からなる。*は同一残基、. は類似残基を示す。

*F. moniliforme*のcDNAの特徴として、Poly(A)付加シグナルが正規のAATAAAでないことが多いと前報で報告した。今回もまたそうで、3'非翻訳領域にAATAAAという配列はない。可能なPoly(A)付加シグナルとして、1622 - 1627のAATGAA, 1734 - 1739と1741 - 1746のAATCAA, および1745 - 1750のAACAAAの4つが考えられる。このうち最初のもはPoly(A)付加点から137 bも離れているので、25 b以内にある残り3つのうちのいずれかがシグナルとしてはたらいっていると考えられる。

類似の実験プログラムについて

本実験プログラムの意図については昨年度の報告で述べた(1)。その眼目は、小なりとはいえ、自らの手で新しいことを見だし、データベースにAuthorとして1ページを加える喜びを学生達に味わわせるというにあった。類似の実験プログラムが考えられていたとしても不思議ではないが、内容的にほとんど同じ報告(4)がなされていたので、それを紹介がてらこのプログラムの意味を再考してみたい。国際生化学分子生物学連合の教育誌にLaboratory ExercisesとしてBrigham Young大学のR. L. Kasparによって報告された、“Integrating Internet Assignments into a Biochemistry/Molecular Biology Laboratory Course”がその論文である。生化学もしくは関連分野の学部4年生を対象に計画したというが、材料として市販のヒトcDNAライブラリを使うことを別にすれば、内容的にはわれわれのプログラムに酷似している。すなわち、1) プラスミドDNAの精製、2) 得られたDNAの制限酵素消化、3) DNAの塩基配列決定、そして4) 決定した配列のヒトゲノムデータベース中の既知配列との比較、というものである。ただしタイトルにもある通り、主眼はインターネットの使用法の習得にある。著者によれば、インターネットに蓄えられた膨大な生化学データベースを使いこなすことは、これからの生化学研究に必須であるにも拘わらず、多くの学生はそのことを知らないし、知っていたとしてもアクセスの方法がわからないので、それを教えるのが非常に重要な課題になっているという。著者はまずInternet Assignmentと称する課題を与え、National Center for Biological InformationのEntrezというサイトを中心に塩基配列データからタンパク立体構造データなどへのアクセス方法を学習させ

た後に、彼がWet-lab moduleと呼ぶ、塩基配列データを実際に生み出す実験へと進んでいる。ただしこのプログラムでは、ヒトcDNAライブラリを使うので、ほとんどは既知配列に行き当たる。したがってモデル実験、いってみれば研究の疑似体験である。もっとも著者は、「もしヒトゲノムデータベースにない配列に行き当たったとすれば、それこそ本当の教育機会である。」と述べてはいる。われわれの場合は、仮に既知タンパクにきわめて似ていたとしても、新しい生物種についての報告であるから、Neuesである。前報でも論じたように、既知の配列に行き当たるのと新しい配列を見いだすのでは感激がまったく違うので、その点ではわれわれの方が優れていると自負している。また今回のように、機能の予測もつかない新しいタンパクに出会うと、本格的に追求してみたいくなる。こうなると指導側としても、次はどことなく当たるかと楽しみになってくる。ただ、われわれはWet-labに重点をおいているので、インターネット使用法を体系的に教えることはしていなかった。この点はKaspar論文に学ぶところが多いし、むしろ正規のカリキュラムとしてそうしたプログラムを考える必要がある。

謝 辞

講座配属学生のために、図書館で文献検索法の講習会をわざわざ設けていただいたことに深く感謝する。

参考文献

- 1) 吉田 浩, 飯塚真理, 石ヶ坪 潤, 小田耕平, 小川 祥子, 福島理恵 (2001) 島根医科大学紀要 24, 33-38
- 2) Tebele N., Skilton R.A., Katende J., Wells C.W., Nene V., McElwain T., Morzaria S.P. and Musoke A.J. (2000) J. Clin. Microbiol. 38, 2240-2247
- 3) Hernandez R., He H.Q., Chen A.C., Waghela S.D., Ivie G.W., George J.E. and Wagner G.G. (2000) Insect Biochem. Mol. Biol. 30, 969-977
- 4) Kaspar R.L. (2002) Biochem. Mol. Biol. Education 30, 36-39

(受付 2002年12月17日)