

## 生命解析への革新的挑戦

相澤 康則 研究室～分子生命科学専攻



相澤 康則 講師

生物学は現在、最も発展の著しい学問の一つである。そのため、次々と新たな問題が発生し、そのたびに従来の考え方が通用しなくなってくる。

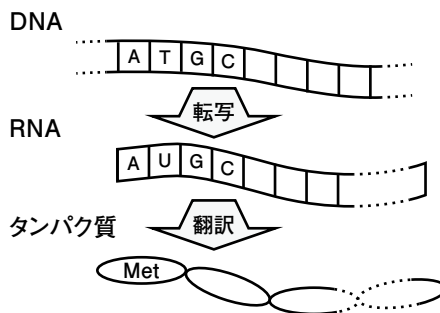
相澤研究室では、従来の常識では解決できないような難問を、ゲノム、RNA、タンパク質、細胞などのさまざまな視点で見ることにより解き明かそうと研究を行っている。先生の研究内容は広い分野にわたっているが、本稿ではその中から、ncRNA からの遺伝子の発見についてと、その遺伝子から見つかった非常に珍しいタンパク質の研究について述べていこう。



## 新たなRNAの発見

大空を翔る鳥、大海を泳ぐ魚、広大な陸地に生い茂る草木、森の中でうごめく昆虫、あらゆる場所に存在する微生物、そして文明を育むヒト、これら地球上のすべての生命の遺伝情報はDNA上にA、T、G、Cのたった4種類の塩基の配列で記録されている。一つの生物がもつ全遺伝情報をゲノムという。技術の進歩によりゲノムの塩基配列を読むことは可能になったが、その配列の意義については現在も未解明の部分の方が多いと言っても過言ではない。相澤先生はヒトゲノムの未知の部分を読み解くため、日々研究を重ねている。

DNAの情報は最終的にタンパク質になり、タンパク質が生体内でさまざまな役割を演じることで生命活動が行われている。タンパク質はアミノ酸が一行に連なったものであり、DNAの塩基配列はRNAを介してタンパク質のアミノ酸配列に変換される。DNAのA、T、G、Cの塩基配列をRNAのA、U、G、Cの塩基配列にうつすことを転写といい、RNAの塩基配列をタンパク質のアミノ酸配列に変換することを翻訳という。このDNA、RNA、タンパク質の関係は生物学のセントラルドグマと呼ばれている(図1)。



DNAからRNA、RNAからタンパク質へと情報が伝達される。

図1 セントラルドグマ

ゲノム上の遺伝子と呼ばれる部分は、ほとんどがタンパク質の情報に対応している。ただし、ゲノム上のすべての塩基配列が遺伝子であるというわけではない。ヒトの場合、遺伝子はゲノムのわずか1～2%程度である。

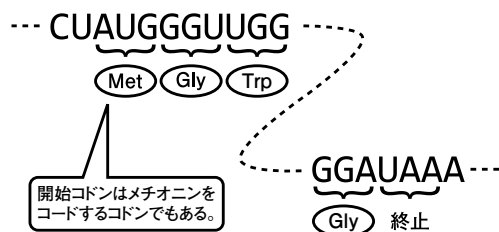
従来は、セントラルドグマののっとり、RNAはDNAの情報とタンパク質の情報を仲介しているものがほとんどだと考えるのが常識だった。しかし、2005年、国内外のグループがマウスやヒトのあらゆる細胞に存在するRNAの塩基配列を調べたところ、従来は遺伝子だと考えられてい

なかった多くのゲノム領域がRNAに転写されていることがわかった。転写された割合は、ヒトゲノムではその約60%にもなった。それまで、RNAに転写されるヒトゲノムの領域は遺伝子の周辺の約25%程度だと考えられていたので、未知のRNAが多数見つかったことになる。

このとき新しく見つかったRNAは、ほとんどすべてncRNAに分類され、データベースに登録された。ある核酸や塩基配列がタンパク質に翻訳される時、それらはタンパク質をコードしているという。ncRNAとはnon-coding RNAのことで、タンパク質をコードしていない領域から転写されたRNAのことである。当時、RNAがタンパク質をコードしているかしていないかは、そのRNAの塩基配列情報をもとに判断された。

DNAとRNAの情報は4種類の塩基の配列で表されるが、タンパク質を構成するアミノ酸は20種類あるので、塩基とアミノ酸の配列の間には変換が必要である。生物は3つの塩基を1つのアミノ酸に変換している(図2)。この3塩基の並びのことをコドンと呼ぶ。コドンは4種類ある塩基が3つ並んだものなので、 $4^3=64$ 種類ある。一方アミノ酸は20種類だから、1つまたは複数の種類のコドンが1種類のアミノ酸に対応することになる。例えば、グリシン(Gly)に対応するコドンはGGU、GGA、GGC、GGGの4種類であり、トリプトファン(Trp)に対応するコドンはUGGの1種類である。

また、コドンには翻訳の開始地点を示す開始コドンと、翻訳の終止地点を示す終止コドンがある。翻訳は開始コドンから終止コドンの直前のコドンまでの間で行われる。具体的には、開始コドンはAUGの1種類、終止コドンはUAA、UGA、UAGの3種類で表される。



三つの塩基に一つのアミノ酸が対応している。

図2 RNAの塩基配列とコドン

次に、タンパク質をコードしているかしていないかを判断するために用いる、ORFマップの説明をしよう(図3)。ORF(Open Reading Frame)とは開始コドンと終止コドンの間を指すもので、つまりは翻訳の読み枠という意味である。ORFマップとは塩基配列中のORFをすべて表示したものである。コドンは3塩基の並びなので、1つの塩基配列から3種類のコドンの並びを作ることができる。

一本のRNA上には複数のORFが存在しうが、その全てのORFがタンパク質に翻訳されるわけではない。一般に、最も長いORFが翻訳される。そのため、あるRNAのORFマップを作り、すべてのORFの長さを調べることで、そのRNAからタンパク質が翻訳されているかどうか、ある程度の判断ができるのである。統計的観点から、最長のORFのコドン数が80以下のRNAはタンパク質をコードしていないと判断されている。つまり、コドン数が80以上のORFをもたないRNAがncRNAに分類されているのである。

以上のように、新しく見つかったRNAの多くは機能の実体が全く分からなかったが、その塩基配列上のORFの長さだけでncRNAに分類された。新たなRNAの発見が、ncRNA研究に世界的な一種のブームを引き起こしたのだった。



## TUFから発見された遺伝子

相澤先生は機能未知のRNAをncRNAと呼ぶことに疑念を抱いた。あるRNAがタンパク質をコードしていないということを証明するのは非常に困難である。なぜなら、そのRNAに含まれるすべてのORFがタンパク質をコードしていないことを証明する必要があるからである。

当時ncRNAと分類されたものは約5500種類

もあった。先生は、最も長いORFが80以下のRNAであっても、それらをすべてncRNAとみなしてしまうのは乱暴だろうと考えた。

相澤先生は、機能未知のRNAを研究するにあたり、それらをTUF(Transcript of Unknown Function)と呼ぶことにした。機能のわからない転写物という正直な呼称である。

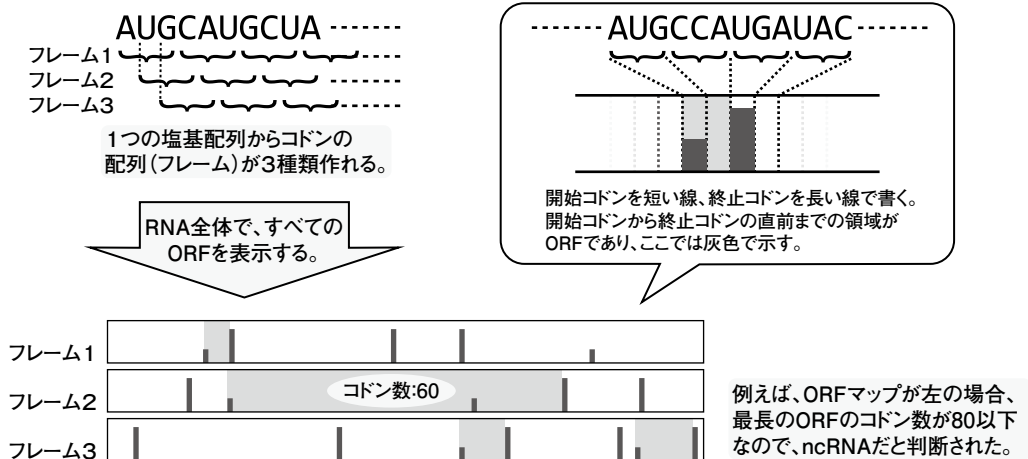


図3 ORF マップ

TUFとして考えられるRNAには3種類ある。一つ目は、RNAそのまま機能するもので、先述した遺伝子を制御するものや、触媒としてはたらくものである。二つ目は、意味もなく転写されているものである。遺伝子の変異が原因で、意味もなく転写されているものは必ず存在する。つまり、すべてのTUFが機能をもつわけではない。三つ目は、アミノ酸数が80以下のタンパク質をコードしているものである。この三つ目の可能性を考慮するために、先生はncRNAと言い切らず、TUFと呼ぶことにしたのである。

相澤先生は、配列情報に頼らずにヒトのTUFを調べる実験を行った。細胞の種類によって発現量が異なるTUFがあれば、そのTUFは転写が調節されているということであり、機能もっている可能性が高い。先生は、TUFの発現量の変化を調べるために、ヒトの脊髄から採取した幹細胞を利用することにした。幹細胞とは、いろいろな細胞のもとになる細胞である。この幹細胞を脂肪細胞と骨細胞に分化させ、発現量の異なるTUFがあるかどうかを調べることにした。あるデータベースにncRNAとして登録されていた約5500種類のうち、ほぼすべてについて発現量の変化を調べたところ、脂肪細胞と骨細胞で発現量が5倍以上異なるTUFが6種類見つかった。

そして相澤先生は、これら6種類のTUFの中から、アミノ酸数が80以下のタンパク質をコードしているRNAを発見したのである。これは哺乳類のTUFがタンパク質をコードしているとい

う初めての例となった。この発見は、新たに見つかったRNAがncRNAであるという先入観に惑わされずに、中立的な視点でTUFを研究した相澤先生の姿勢に起因すると言える。この遺伝子はAGD3と名付けられた。以降、この遺伝子の発見の経緯について述べていこう。以下、遺伝子を示すときにはAGD3遺伝子、AGD3タンパク質と書く。

現在データベースに登録されているタンパク質をコードするORFの長さは平均400～500コドンである。しかし、AGD3遺伝子のORFの長さはわずか63コドンである。言い換えれば、AGD3タンパク質はたった63個のアミノ酸が連なった短いタンパク質だということである。

なぜ、相澤先生はこの小さなタンパク質を発見できたのだろうか。それは、ヒトのAGD3遺伝子のORFの相同領域をマウスのゲノム中に見つけることができたからだ。相同領域とは、異なる生物のゲノム間で似ている領域のことである。ヒトのAGD3遺伝子の相同領域をマウスのゲノム上で調べたところ、ヒトのAGD3遺伝子と長さが一致し、アミノ酸配列も非常に似ているORFが発見された。これは、AGD3タンパク質が哺乳類動物で共通の重要な役割を果たしている可能性が高いことを示している。このような考えのもと、先生はAGD3タンパク質の検出を試みた。そして、実際にそのタンパク質が翻訳されていることを明らかにし、遺伝子の存在を証明したのである。



# 全長天然変性タンパク質の特性

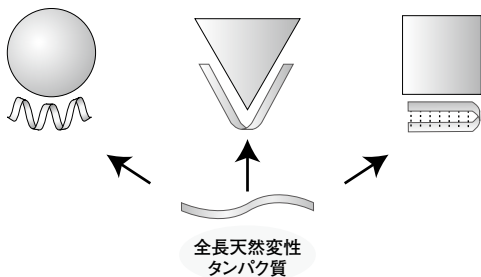
相澤先生はTUFが小さなタンパク質をコードする可能性があるという新たな考え方を示した。その後、AGD3タンパク質を深く解析した結果、AGD3タンパク質は全長天然変性タンパク質であることが明らかになったのである。

天然変性とは、生理的条件下で一定の構造をとらない性質のことであり、全長天然変性タンパク質とは、天然変性が全長にわたっているタンパク質のことを指す。

従来は、タンパク質はある特定の安定な立体構造をとることが多いと考えられてきた。ほとんどのタンパク質では、アミノ酸残基間の水素結合などによって、直鎖であったペプチドが折りたたまれてらせん構造をとったり、シート状の構造をとったりしている。

それに対し、全長天然変性タンパク質は、それ単体では生理的条件下でそのような一定の構造をとらず、他の物質に結合してはじめて安定な構造をとる。また、結合する対象の物質によってその構造は異なる(図4)。AGD3タンパク質は、わずか63のアミノ酸からなるというだけではなく、全長にわたって天然変性であるという非常に珍しいタンパク質である。

部分的に天然変性なタンパク質は今までに数多く見つかっていた。一定の構造をとる部分は特定の物質にのみ結合するが、天然変性な部分はいくつかの異なる物質に結合する。部分的に天然変性なタンパク質は、一定の構造をとる部分が特定の物質に結合し、天然変性の部分が複数の種類の物質をそこに誘引するという仕組みをもっているこ



安定な構造は結合する対象によって異なる。

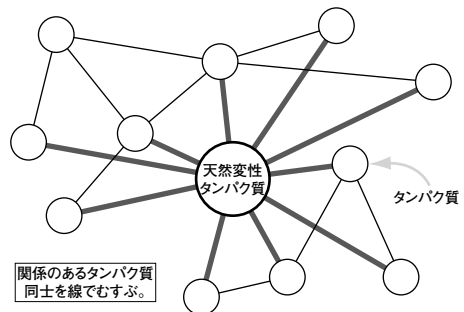
図4 全長天然変性タンパク質の多様な結合

とが多い。このことから、全長天然変性タンパク質は、非常に多くの種類の物質と結合し、相互作用することが予想される。

先生は、AGD3タンパク質は細胞内シグナル伝達に関わる機能があるのではないかと推察している。研究の結果、細胞内シグナル伝達に関わるタンパク質に付加される化学物質と同じものが、AGD3タンパク質に付加されていることがわかったからだ。シグナル伝達とは細胞内での情報のやり取りのことで、シグナル伝達に関わるタンパク質は小胞体という細胞内小器官から細胞内のいろいろな場所に輸送される。実際にAGD3タンパク質に緑色蛍光タンパク質をつけて観察したところ、小胞体の表面とその周辺に蛍光が確認できた。このことから、AGD3タンパク質が単に拡散しているわけではなく、細胞内輸送されていることが明らかになったようだ。

今後の研究では、AGD3タンパク質の天然変性が、シグナル伝達という機能にどのように関与するかを明らかにしたい、と相澤先生は考えている。

天然変性タンパク質は数多くの物質と結合できるので、いろいろな化学反応に関わることができる。その結果、生体内で広く影響を及ぼしやすく、細胞内ネットワークのハブという重要な役割をもっていることが多い(図5)。相澤先生の研究が進めば、全長天然変性タンパク質の細胞内ネットワーク上の位置づけについても明らかになっていくだろう。



天然変性タンパク質は、数多くのタンパク質と関係をもつため、細胞内ネットワークのハブとなっている。

図5 ハブとしての天然変性タンパク質



## 相澤研究室のこれから

相澤先生はTUFの中からAGD3タンパク質という全長天然変性タンパク質を発見した。現在、相澤研究室ではAGD3タンパク質の他にもTUFなどから見つかった数十の候補の中から全長天然変性タンパク質を探している。全長天然変性タンパク質は、今まで見つかったタンパク質とは性質が大きく異なるタンパク質である。よって、従来見つかったタンパク質とは全く違う次元で細胞内制御をするタンパク質群がTUFなどから新たに見つかる可能性があると考えられる。全長天然変性タンパク質の研究が進めば、さまざまな応用につながる可能性もあるだろう。

例えば、薬への応用を考えてみよう。あるタンパク質をもとにして創薬をする場合、タンパク質をペプチドに切断してから化学合成することが多い。そのため、タンパク質をペプチドに断片化し、そのペプチド群の中からもとのタンパク質と同程度の活性をもつペプチドを探し出すという手法が採られる。しかし、多くのタンパク質は全体で立体構造を作っているため、ペプチドを途中で切ってしまうと立体構造を作れなくなり薬として使え

なくなることが多い。

一方、特定の構造をもたない天然変性タンパク質ならば、病気の原因となる物質と結合する部分だけを切り出しても結合力を失わない可能性が高いので、その部分が薬となりうる(図6)。つまり、通常のタンパク質に比べて容易に創薬への応用ができるのだ。

一言で説明できない生命現象を説明しようと思ったときにどの知識を使えばいいのかを瞬時に判断できる能力、それが生物学的センスであり、それは数学や物理の高度な問題を解くときに使う数学的センスや物理学的センスと何ら変わりはないと相澤先生は言う。生物学的センスを意識していれば、生物学を単なる暗記科目ではなく、生物のシステムを解読する学問だとよりの確に捉えることができる。生物のシステムは工学的なアイデアにもつながっていく、と先生は付け加えた。常識にとらわれず、生物をゲノム、RNA、タンパク質、細胞といったさまざまな観点から見ることにより、相澤先生が今後どのような活躍を見せるのか、注目したい。

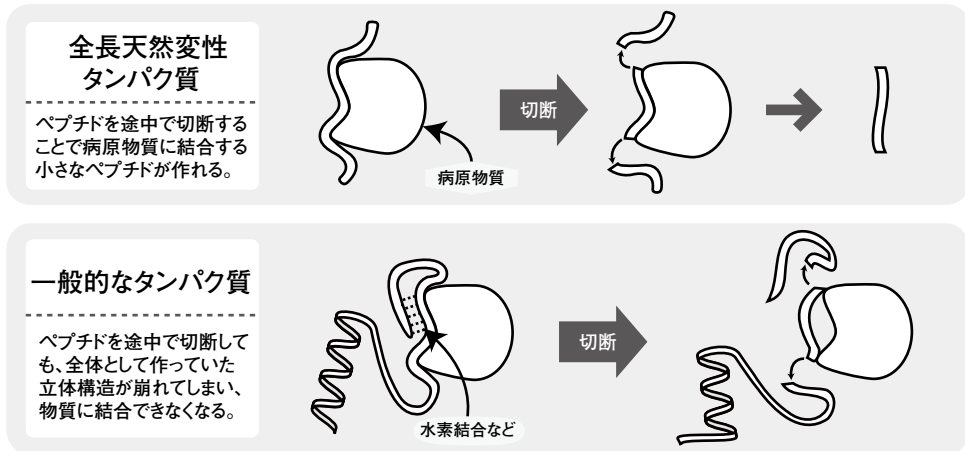


図6 全長天然変性タンパク質からの創薬

今回取材させて頂いた際、記事の内容以外にも知的好奇心をくすぐるような興味深い研究の話がたくさんして頂き、勉強になりました。

お忙しい中、取材のためにお時間を割いてくださった相澤先生に心より感謝いたします。

(渡来 直生)