



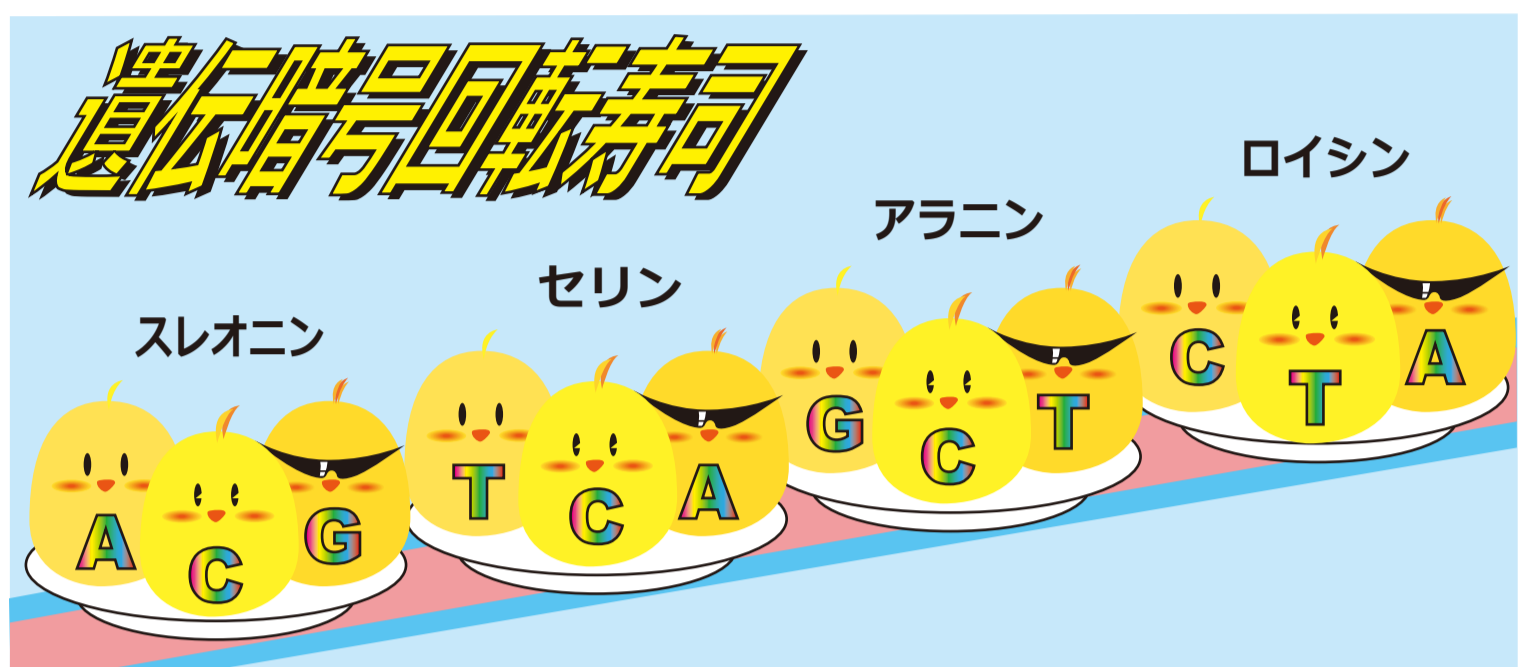
第5回 遺伝子は行間を読まなきゃだめ？ ヒトゲノム計画その後

「ヒトの遺伝子：ゲノムとは？」

私たちの遺伝情報は、DNA（デオキシリボ核酸）という物質で、「二重らせん構造」をしているというの有名な話です。これは1953年にワトソンとクリックによって確かめられました。では、DNAにどうやって遺伝情報が書かれているのか？それはアルファベットにたとえると、A（アデニン）、G（グアニン）、C（シトシン）、T（チミン）の四文字が使われていて、3文字が1組となって、1種類のアミノ酸を指定しているというわけです。

この情報を「翻訳」すると、アミノ酸の長い鎖（蛋白質）を意味しているわけで、細胞内で遺伝情報から蛋白質が合成されることで、遺伝情報が発現したことになります。

1種類の生き物が持っている、遺伝情報そっくりワンセットのことを、「ゲノム」と呼びます。



とはいえ、これは基本中の基本。この説明の仕方だと、通常「遺伝子」と呼んでいるものは、DNAが3文字ひと組でたくさん並んでいるもの、ということになります。ヒトのゲノムは、「30億文字」で書かれているわけです。これは確かに膨大な情報量です。さすが人間様と思いたいところですね。このヒトのゲノムを世界中の研究者が協力して、配列を全て読み取ろうというのが「ヒトゲノム計画」で、読み取り作業自体は2003年に終了しています。

「でも、隙間だらけ。もしくはすっからかん？」

ところが残念ながら(?)、そのうちの10%程のDNA配列しか、アミノ酸を指定していないことがわかっていて残りの配列は、「がらくたDNA」なんて言い方さえあります。

その「無意味なDNA配列(?)」は、すべての遺伝子情報(蛋白質のアミノ酸配列)の中にまで入り込んでいます。1つの蛋白質の遺伝子の中で、アミノ酸を指定している部分を「エクソン」、指定していない部分を「イントロン」と呼びます。実際に遺伝情報が発現するときには、DNAからRNA(リボ核酸)へ情報を転写した後、イントロン部分を切り取ってエクソン同士をつなぎ合わせる「スプライシング」という仕掛けまで用意されています。

この遺伝情報の状態を、身近なものに例えるなら、「見どころの直前になると決まって長いCMが入り、これを何度も繰り返す、ついイラッとするテレビ番組」と言う感じでしょうか？更に異なる遺伝情報同士の間には、延々と「がらくたDNA」が続くのです。つまり「次の番組まで延々といつまでもCMが続く」わけです。

ヒトのDNAって、何故こんな無駄なことをしているのでしょうか？

「がらくた箱をのぞいて見る？」

世間には「物好きなヒト」がいるもので、このがらくたDNA(延々と続くCM)の配列を調べたヒトがいるのです。するとその結果、奇妙なことがわかりました。正常に働いている様々な遺伝子と、ほとんどそっくりな(がらくた?)配列が沢山見つかったのです。中には、大昔に私たちのDNAに紛れ込み、そのまま居付いてしまったウイルスのDNAまで複数見つかってしまいました。

使われない90%のDNA配列の中には、このような配列が膨大な数で散乱しています。新薬開発の目的で遺伝子配列を調べ、「新しい遺伝子を見つけた！」と思っても、大抵はこの「がらくたDNA」の一部だったりします。そういうものは論文発表する価値が少なく、もちろん商売にもならないので、世界中の研究者を、楽屋裏で散々がっかりさせ、泣かせてきたのです。

「細胞1個のDNA量の進化史」

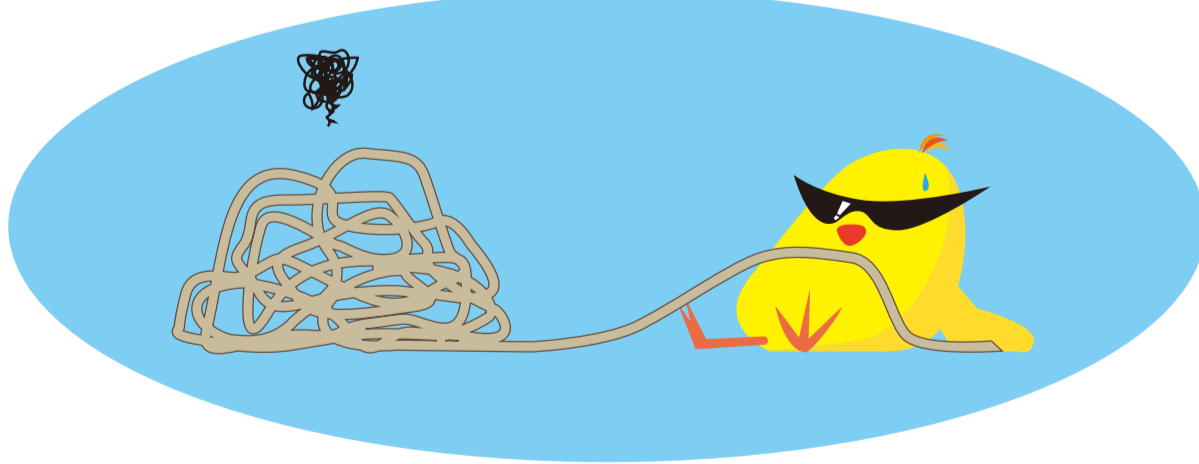
ここで、ちょっと話を変えてみます。ヒトのDNAゲノムは、細胞1個の中に30億文字入っているわけですが、「文字数」は、生き物の種類によって異なります。ところがこの文字数は、生き物が進化していく歴史の中で「順調に増え続けた」わけではなく、必ずしも高等な生物ほど多いという訳でもありません。

ヒトよりも単純な身体なのに、ヒトよりずっと多い文字数の生き物もいるようです。動物の系統進化の様子を描いた「系統樹」は、上へ上へ枝が分かれて伸びていく樹のように描かれますが、これに細胞1個あたりのDNA文字数を当てはめると、まるで「しだれ柳」のようになってしまうのです。

どうやらゲノムの文字数は進化の歴史の中で、ある時期一気に増加し(ゲノムがそっくり重複する”倍数化”)という現象も知られていて、生き物の世界では珍しいことではありません)、そこから分かれた子孫の系統で減少するといったことが頻繁に起きた可能性が有ります。

その「ある時期」とは何なのか？例えば魚が両生類に、といったような生き物が大きく姿を変える「大進化」には、そのような「DNA大盛増加」が必要なのでは？と推測する意見もあります。沢山の遺伝子が重複し、それぞれが時間とともに異なる機能を持った遺伝子に進化していくことで、「新しい生き方」が生まれると言うのです。

私たちの「がらくたDNA」は、そうした進化の経過を物語る「新しい遺伝子開発」の名残りというわけでしょうか？



「一見がらくた。でも実体は？」

しかし、また違った見方もあります。結局「使われているDNA」は10%だけで、そこに書かれている遺伝子(蛋白質)の数は、3万種類程度だろうと言われています。これを「少なすぎる！DNA上の情報は本当にこれだけか??」と疑ってみることで。

実際、遺伝子配列の上流には、その遺伝子発現を制御する他の蛋白質が結合する「領域」が沢山見つかっていて、配列自体が蛋白質の酵素のように働くこともあることが分かってきました。

直接アミノ酸を指定していないから、「無意味ながらくた」と思ったら大間違い。全く別のルールで書かれた「暗号」が、30億文字のなかに相当潜んでいるようなのです。まさに遺伝情報は、アミノ酸指定以外の「行間を読む」必要があったのです。

いずれにしろ、ヒトの遺伝情報全体には一体何が描かれていて、どう働いているのか？「ヒトの身体はどうやって生きているのか？」全体像が掴めるのは、まだ先のことのようにです。

個人的には、30億文字全てを要約すると、「この情報を書き換えてでも生き抜け。」と書いてある、なんて夢想するのですが。