

論文内容の要旨

論文題目 Study on Evolution and Diversity of the Genes with Tandemly-Duplicated Structures

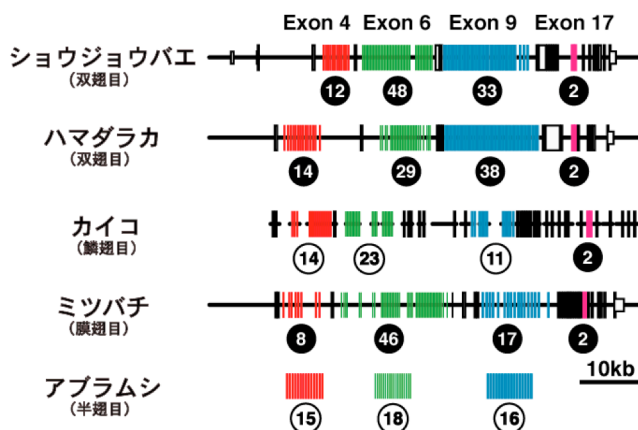
(反復構造をもつ遺伝子の多様性と進化に関する研究)

氏名 大西 啓介

1. DSCAM遺伝子の進化

ショウジョウバエ (*Drosophila melanogaster*) の DSCAM 遺伝子は神経系で重要な役割を果たす遺伝子である。本遺伝子は Exon 4, 6, 9, 17 内の各メンバーが反復した構造をもち (Alternative Exon)、選択的スプライシングによってそれぞれから1つずつのエキソンを選ぶことにより、3万種以上の mRNA を生じうる (図1上)。まず DSCAM 遺伝子のゲノム構造をゲノムのドラフト配列が公開されている *Drosophila* 属の3種 (*melanogaster*・*yakuba*・*pseudoobscura*) で比較解析した。その結果、多くのエキソンは3種間

で高度に保存されているものの、Exon 6 においては、*yakuba* で特異的に失われたメンバーがいくつも存在することが明らかになった。次に、ドラフト配列が公開された昆虫4種 (ショウジョウバエ (*Drosophila melanogaster*)・ハマダラカ (*Anopheles gambiae*)・カイコ (*Bombyx mori*)・ミツバチ (*Apis mellifera*)) に加え、自分自身が degenerate PCR により増幅し、Alternative Exon 4, 6, 9 の配列を決定したアブラムシ (*Acyrtosiphon pisum*) の配列を用い、比較解析した (図1)。その結果、選択的スプライシングを受けないエキシソンのアミノ酸



<図1> 昆虫DSCAM遺伝子のゲノム構造比較

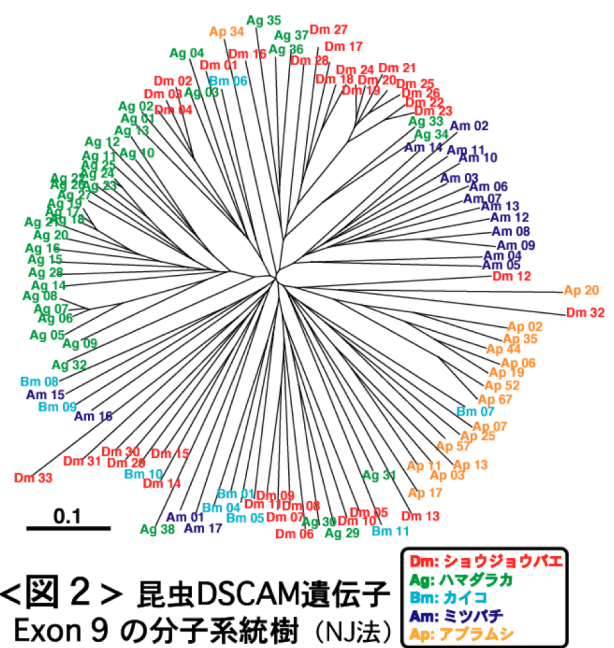
白丸は各エキソンクラスター内のメンバー数が明らかなもの、黒丸はゲノム配列の不完全さにより、実際にはさらにメンバーが多い可能性があるものを示す。

配列は昆虫の種間で高度に保存されていた。しかし、Alternative Exon 4, 6, 9, 17 のそれぞれで分子系統樹を作成したところ、Exon 17 を除き、目・科のレベルではそれぞれの系統特異的に大部分のエキソンが進化していることが明らかになった（例：図2）。さらに、本当にこれら Alternative Exon の大部分が組み合わせによるトランスクリプトームの多様性に寄与しているかどうかを確認するため、ミツバチとアブラムシでそれぞれ 44, 66 クローンの cDNA の塩基配列を決定した。その結果、予想通りいずれの種でも様々なエキソンが組み合わせられていることが確認された。また、ハチにおいてはスプライシングのエラーが多い（Exon 6 のメンバーのうち2つ以上が1つの成熟型 mRNA に含まれていたり、イントロンの切り離しが正確でなかったりする）が、アブラムシではエラーが少ないことも明らかとなった。

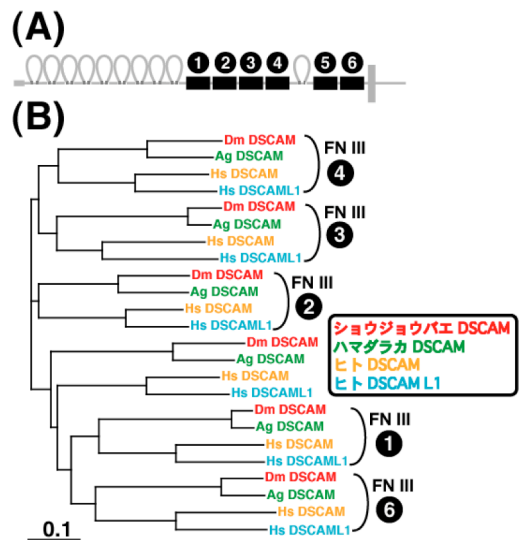
次に、エキソン反復構造をもたない昆虫の DSCAM 遺伝子と同じドメイン構造（図3(A)：10個の Igドメインと6個の FN IIIドメイン）をもつ遺伝子が脊椎動物でも知られており（DSCAM と DSCAML1）、これらと昆虫の DSCAM 遺伝子を用い、解析をおこなった。Igドメインと FN IIIドメインそれぞれのみを用いて作成した分子系統樹（図3(B)：FN IIIドメインの系統樹）において、大部分のドメインが種間でクラスターを形成したことから、脊椎動物の DSCAM と DSCAML1 遺伝子は昆虫 DSCAM 遺伝子と相同な関係にあることを明らかにした。さらに、同じドメイン構造をもつにもかかわらず、エキソン反復構造をもたない遺伝子が昆虫でもいくつか知られており、これらも用いて 遺伝子全長で分子系統解析をおこなうことにより、DSCAM 遺伝子自体の重複は脊椎動物、無脊椎動物の両方でそれぞれ独立に何度も起こり、また Alternative Exon は昆虫の1つの遺伝子においてのみ生じたことが明らかになった。

2. 直列反復構造をもつ遺伝子

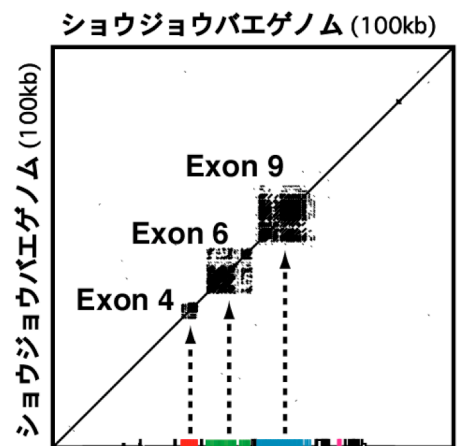
次に私はヒトゲノム中にも、反復構造をもち、かつ脳で重要な機能を果たす未知の遺伝子構造が存在し、それがトランスクリプトームやプロテオーム、インタラクトーム多様性の一端を担っているのではないかと考えた。そこで、ショウジョウバエ DSCAM 遺伝子領域の Dot Plot ではエキソン反復領域特異的に『正方形状』のプロットが生じる（図4）ことに注目した。この Dot Plot を用い、ヒトおよびショウジョウバエの全ゲノム配列中から反復構造の検出をおこなった。Dot Plot を用いることで DSCAM 遺伝子と相同性がない反復構造にまで検出を一



<図2> 昆虫DSCAM遺伝子 Exon 9 の分子系統樹 (NJ法)



<図3> DSCAM遺伝子の機能ドメインにもとづく分子系統樹 (NJ法)
(A) タンパク質の構造 (FN IIIドメイン)
(B) FN IIIドメインの分子系統樹



<図4> ショウジョウバエ DSCAM遺伝子領域の Dot Plot

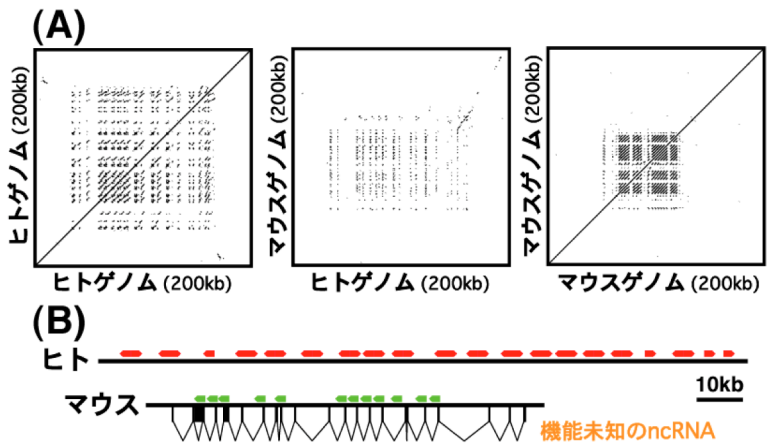
般化した。本解析の結果、ヒトゲノムにおいてショウジョウバエゲノムよりもはるかに多くの反復構造を検出できた(表)。ドメインの反復構造をもつような遺伝子はヒト、ショウジョウバエとも比較的多くを検出でき、これらの中には構造タンパク質が非常に多く見られた。遺伝子の反復構造としては、タンパクコード、非コードともヒトの方がショウジョウバエよりもはるかに多くの構造を検出できた。エクソン(あるいは gene fragment)の反復をもつ遺伝子構造としては、ヒトの方が多くのもを検出できた。しかし、ショウジョウバエの DSCAM 遺伝子のように、複数の

<表> Dot Plotによる全ゲノム解析により
検出された遺伝子構造 (一部)

	ヒト	ショウジョウバエ
ドメイン反復	Titin	Titin
	Collagen	Dumpy
	DAZ	
遺伝子反復 (タンパクコード)	Histone	Cytochrome P450
	HOX (A, B, C and D)	
	Olfactory receptor	
	Zinc finger	
	C/D box small nucleolar	5S rRNA
(非タンパクコード)	TTTY	
	Protocadherin (α and γ)	DSCAM
	Immunoglobulin (κ light, γ , heavy and λ)	
	T-cell receptor (α and β)	
未知の構造	多い	少ない

Alternative Exon のクラスターをもつ未知の遺伝子構造は検出することができなかった。また、ショウジョウバエではほとんど検出できなかったが、ヒトでは、機能未知であるが何らかの機能をもつ可能性が高い、反復構造をいくつか検出することに成功した。

その一例が図5に示した150kb近くにおよぶ反復構造である。本領域はマウスにも対応する領域が存在する。しかしヒトとマウスではプロットの形が大きく異なる(図5(A))。本領域に含まれる反復因子を詳しく解析したところ、ヒトとマウスの因子間では局所的にのみ相同性があることが分かった。また、ヒトでは反復因子が2つ反対向きに並ぶ、回文状の構造をいくつももつが、マウスではすべての因子が同方向であることが明らかとなった



<図5> ヒトの機能未知の反復構造の一例

(A) ヒトとマウスのDot Plot (B) 反復因子のゲノム上の構造
 ◀と ▶ はそれぞれヒトとマウスの反復因子をあらわす。
 互いの因子間には局所的に相同性がある。

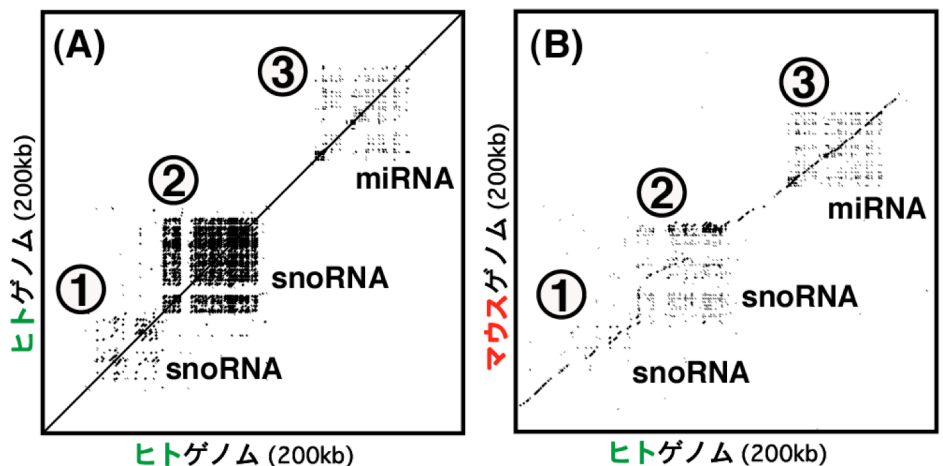
(図5(B))。さらにマウスにおいて

は、反復因子の大部分をイントロン内に、一部をエクソン内に含むような、機能未知の non-coding RNA (ncRNA) が存在することも明らかとなった。

以上のヒトゲノム解析により発見された機能未知の反復構造のうち、もう1カ所に関し、次に詳細に解析をおこなった。

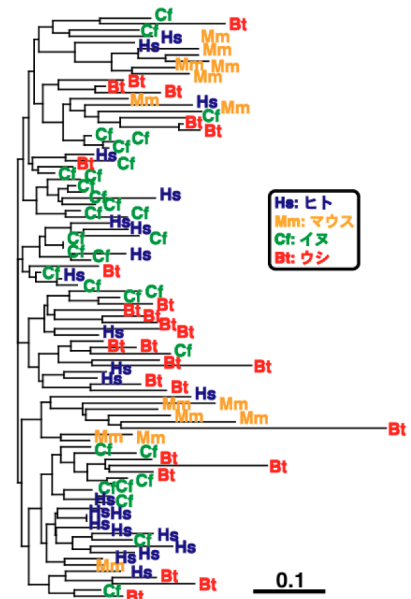
3. ヒトのncRNAクラスター

Dot Plot によるヒトゲノム解析により見つかった領域のうち、最も興味深いのが図の領域である



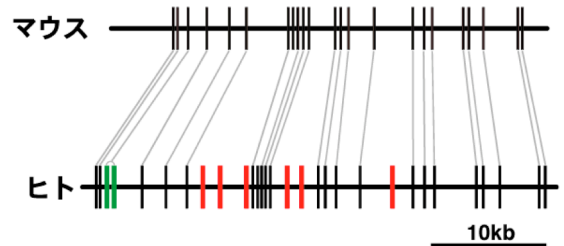
<図6> ncRNA複合領域のDot Plot (A) ヒトーヒト (B) ヒトーマウス

(図6(A): ncRNA 複合領域)。1辺が20kb以上のプロットが生じる頻度はヒトゲノム中であれにもかかわらず、本領域はこのようなプロットが200kb以内に3つも並ぶという、昆虫DSCAM 遺伝子領域(図4)に似た、極めて特異な構造をもつ。しかし、DSCAM 遺伝子とは違い、3つのプロットとも明らかに読み枠(ORF)をもたない。本領域はフグ・ニワトリ・オポッサムのゲノム中には相同性のある領域が確認できなかったが、マウス・ラット・イヌ・ウシのゲノム中には明確に1:1に対応する領域が見つかった(図6(B): ヒト・マウス)。解析の結果、クラスター①と②はsmall nucleolar RNA(snoRNA)の反復構造であることが明らかとなった。これらの反復因子の塩基配列を哺乳類4種において取り出し、分子系統解析をおこなった。まずクラスター①では、ヒト・イヌ・ウシでは比較的保存性が高かったが、マウス・ラットではそれぞれの系統特異的な進化を果たしたメンバーが大部分を占めた。一方、クラスター②では、ヒト・イヌ・ウシのそれぞれの系統間でも昆虫のDSCAM 遺伝子(図2)と同じように、大部分のメンバーが特異的な進化を果たしたことが明らかとなった(図7)。



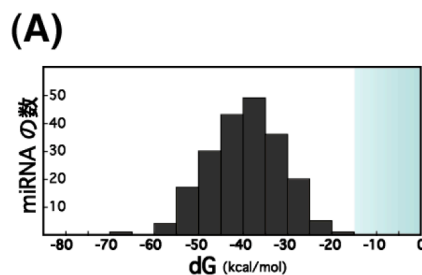
<図7> snoRNAクラスター(図6の②)の分子系統樹(NJ法)

次に、クラスター③のメンバー1つがマウス脳由来の既知microRNA(miRNA)と完全に一致したことから、クラスター③は驚くべき構造をもったmiRNAクラスターであることが明らかとなった。本クラスターのヒトパラログ間の塩基配列の相同性は平均60%と多様性が高い。しかしマウス23個中22個でヒトに相同な配列があり、前駆体領域において大部分の配列が高度に保存されていた(図8:黒)。それ以外のうちヒト7個はマウスゲノム中に対応するものが見つからず、そのうちいくつかはイヌゲノム中にも存在しないことから、ヒトの系統特異的に生じた可能性が高い(図8:赤)。



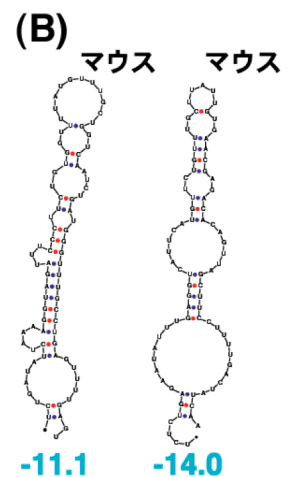
<図8> miRNAクラスター(図6の③) 赤は霊長類にあっげっ歯類にはないもの、 緑は霊長類でコピー数が多いものを示す。

次に、前駆体2次構造の自由エネルギーの値を既知miRNAのその分布(図9(A))と比較した。既知miRNAにおいてはこの値が-20kcal/molより大きいものはほとんど存在せず、-15kcal/molより大きいものは1つも知られていない(図9(A):青い領域)。本miRNAクラスターに含まれるメンバーの大部分は-30kcal/molよりも小さな値をとる。しかしマウスの2つに関しては、この値が-15kcal/molより大きく、非常に不安定である(図9(B))。



<図9> miRNA前駆体2次構造の自由エネルギー解析

(A) ヒト既知miRNAの前駆体2次構造の自由エネルギー分布 (B) miRNAクラスター内で前駆体2次構造が不安定なもの



またヒトにおいてもこの値に近い、かなり不安定な前駆体をもつものが2つ存在する。このことから、いくつかのメンバーにおいて、ヒトとマウスそれぞれの系統で独立にmiRNAの偽遺伝子化が起こった可能性が高いことが明らかとなった。