

博士論文

ミトコンドリア遺伝子の構造の研究

浅川修一

①

博士論文
ミトコンドリア遺伝子の構造の研究

指導教官 三浦謹一郎教授

東京大学大学院工学系研究科工業化学専攻
87169 三浦研究室 浅川修一

目次

I. 序論	1
II. 実験操作	
1. ミトコンドリアの精製	10
2. ミトコンドリアDNA, RNAの調整	11
3. ミトコンドリアDNAのクローニング	12
4. Polymerase chain reaction (PCR) 法による未クローニング領域の増幅	13
5. シークエンスのためのdeletion mutantの作成	14
6. サブクローニングによるテノプレートDNAの調製	15
7. 合成シークエンスプライマーによるシークエンス	15
8. dideoxy sequence法による塩基配列の決定	15
9. 16SrRNAのS1マッピング	16
10. 16SrRNAのプライマーエクステンション法 によるマッピング	16
III. 結果および考察	
1. 制限酵素の切断部位	19
2. 塩基組成と鎖 (Strand)	23
3. 遺伝子の同定	27
4. 遺伝子配置	38
5. 遺伝暗号	49
6. 蛋白質遺伝子	55
7. tRNA遺伝子	70
8. プロモーター配列	88
9. 鎖 (Strand) に特異的な塩基の偏り	93
10. 12S、16SrRNA遺伝子	106
IV. まとめ	111
V. 参考文献	114
VI. 謝辞	120

1. 序論

人から細菌に至るまで多様に分岐しそれぞれが極めて複雑なシステムで運営されている（生きている）生物を物質科学的に一般化された形で説明しようという試みがなされてきた。特に1950年代より発展してきた分子レベルでのアプローチは大きな成功を収めつつあり、DNAの情報をもとにRNAを介してタンパク質が合成される様子があらゆる生物にとって最も基本的な営みである事を示した。これらの発現機構はその微細な部分においては生物により多様化しているが、遺伝暗号^{1,2}、もRNAの基本的な構造^{3,4}、コドンとアンチコドンの対合様式などの基本的な事柄は人から細菌に至るまで普遍的であると思われた。しかし真核生物の細胞中に存在するミトコンドリア（図1）において、これらの最も基本的な事柄について例外がある事が解った。

ミトコンドリアは真核生物の細胞中に存在し酸化的リン酸化により生体のエネルギー源であるATPを合成する細胞内小器官である。ミトコンドリアの本来的な機能は、糖などの代謝産物をクエン酸回路に關与する酵素や、その他の酸化酵素群と電子伝達系酵素により終末酸化を行い、水と二酸化炭素に分解するとともに、それと共役してADPからATPを合成し、生体に供給するというものであるが、これらの代謝的な研究は1900年代前半より進められてきた。

一方、ミトコンドリアの中には核DNAとは別に独自のDNAが存在することや、独自の蛋白質合成システムが働いていることなどが明かになるに従い、そのゲノム構造を調べることや遺伝情報の発現システムを探ることも興味を持たれていった。これらのことは1960年代から次第に明らかになっていき、その遺伝情報の発現機構、DNAの複製機構などについての研究がされていた⁵⁻⁸。そして1970年代後半にDNAの塩基配列の決定法⁹が飛躍的に発展したのを受け、

1981年にイギリスのサンガーらによりヒトのミトコンドリアの全塩基配列が決定されたことにより¹²、ミトコンドリアのDNA、あるいはミトコンドリアの遺伝情報の発現系がさらに注目されるようになった。

その結果は、ミトコンドリアの中ではいくつかの点で従来研究されてきた真核生物の細胞質や原核生物の遺伝情報の発現システムの常識をくつがえすようなシステムが働いている事を示していた。更にその後、ウシ、マウス、アフリカツメガエル、ショウジョウバエや酵母、原生動物などでミトコンドリアの遺伝子構造等が研究されたが、その結果、上に述べたようにミトコンドリアの遺伝情報の発現系が単に特殊であるだけでなく、生物の進化段階によってその特殊性が多様化している事が解った¹⁰⁻²⁹。

それらのうち主に良く研究の進んでいる脊椎動物と節足動物におけるミトコンドリアに特徴的な点をまとめると次の通りである。

1. 遺伝暗号

ヒトやウシの塩基配列が決定される過程で、ORFと予想されるフレームにおいて普遍暗号では終始コドンであるTGAが頻出したため、ミトコンドリアではトリプトファンの暗号ではないかと予想されていた³⁰。さらにAGA、AGGコドンはORFの予想される3'末端(てRNA遺伝子などに隣接している部位)以外には存在しないため終始コドンであると予想された。その後ウシにおいて遺伝子とそれに対応する蛋白質のアミノ酸配列の解析の結果と対応づけられた所、TGAはトリプトファン、さらにATAがメチオニンのコドンである事が確かめられた。一方、酵母、原生動物などのミトコンドリア遺伝子の解析によるとTGAがトリプトファンであることは一致していたが、AGA、AGGは普遍暗号と同じくアルギニンを指定していた。また節足動物であるショウジョウバエのゲノム解析の結果、AGAがORF上に見つかりセリンを指定するものと予想された。AGGは存在していなかった。

2. tRNAの数

ミトコンドリアのゲノムには高等生物において22個のtRNAがコードされている。一方酵母においては24個のtRNAがコードされていると考えられている。また原生動物や植物においても二十数個であると予想されている。植物においては細胞質より一部のtRNAが移入するとの報告もあるが³¹、後生生物においてはこれらの22個のtRNAがミトコンドリアの蛋白質合成に必要なtRNAの全てであると考えられている³²。

tRNAのアンチコドンとmRNAのコドンとの塩基対の形成の可否による識別は遺伝情報を正確にアミノ酸配列に転換する際の必須の過程である。従来、ミトコンドリア以外の系ではこの塩基対の形成はコドンの1、2文字目とアンチコドンの3、2文字目の対合においては厳密にワトソン-クリック型のA-U、G-Cによる塩基対が必要であるが、コドンの3文字目とアンチコドンの1文字目の対合においてはG-U塩基対、I-A、C、U塩基対などのwobble（ゆらぎ）塩基対まで対合が可能であるとされていた³³⁻³⁵。しかしこのことを考慮しても60-62通りの遺伝暗号に対応するためには最低31~32個のtRNAが必要となり、ミトコンドリアのtRNAの数では対応しきれないことになる。そこでミトコンドリアの系では各4コドンファミリーのコドンに対応するtRNAが各1種類しか存在しないことから、その1種類で4コドンファミリーのコドン全てに対応しているものと考えられた。それらのtRNAにおいては、そのアンチコドンの1文字目が未修飾のUであることがtRNAの解析により示されており³⁶、コドンとアンチコドンの対合は、U-N wobble塩基対の形成によると言う説と³⁷、アンチコドンの1文字目のUはあまり関係なく、コドンの1、2文字目とアンチコドンの3、2文字目だけが対合すれば良いとする、3中2説などが提唱されているが³⁸、正確な対合を行うための必要十分条件はまだ詳しくは解っていない。

3. tRNAの構造

脊椎動物や節足動物においてtRNAの遺伝子から推定される二次構造はセリンのAGY (Y=U, C), AGX (X=U, C, A)のtRNAを除けば、一般的なクローバーリーフ型の二次構造であると推定された。しかしその二次構造において、ステム中には多くのG-U塩基対、ミスマッチ塩基対が見い出された。またそのループについてはGG配列、TTCRA (R=A, G)配列などの高次構造形成に必要な共通塩基に欠くものがほとんどであった。さらにtRNAの塩基分析により、それらは修飾塩基が少ない構造であることが解った。またセリンのAGY, AGXに対応するtRNAの構造はD-loopを欠くという異常なものであった。(図2)

4. mRNAの構造

mRNAには原核生物や真核生物にみられるような先導配列がほとんど存在しなかった。遺伝子中には開始コドン以外に多くのメチオニンのコドンが存在しており、翻訳開始のメチオニンのコドンとして唯一存在する情報は5'端近傍に存在するということであると考えられている。さらにその開始コドンとしてはAUN (N=A, C, G, U)、さらにGUGが存在すると推定されている。また遺伝子中に終始コドンの存在しないものもあるが、それらの終始コドンはmRNAのポリアデニル化によって生成されるものと説明された³⁹。

5. RNAのプロセッシング

tRNA、mRNA、rRNAはポリシストロニックに転写される。そしてmRNA、やrRNAは独立にプロセッシングされるのではなく、tRNAのプロセッシングに共役して起こるという省略化され、重複化したシステムで生じることが解った。

6. ゲノムの高集積化

これらの高等動物ミトコンドリアゲノムには90%以上の領域に遺伝子がコー

ドされており、介在配列やイントロンの多く含む核のDNAとは異なり、集積化が進んでいる。プロセシング機構の重複化やポリアデニル化による終始コドンの生成もその一端と考えられる。

以上に示した特徴を端的に述べると、ミトコンドリアの蛋白質合成系は極めて単純化、省略化されており、rRNAのプロセシングやポリアデニル化による終始コドンの生成などの事柄については機能が重複化されていると言うことができる。

現在の分子生物学の成果のほとんどは原核生物とそれを宿主とするファージ、真核生物とそれに宿主とするウイルスの研究によって得られたものである。それらの研究は実に良く進んでおり、それまで物質的な意味でのメカニズムが良く解らなかつた生物のかなりの部分を分子レベルで解き明かしてきた。なかでも遺伝情報から蛋白質を合成するという過程のメカニズムの解明については良い成功を取めている。しかしこのように良く研究の進んだ対象においても、分子と分子を単に関連づけることには良い成功を取めているが、まだそれらの関連づけを微小な構造のレベル、化学的なレベルあるいは物理的なレベルで説明づけることには多くの問題が残っている。例えばmRNAがリボソーム上で翻訳の指令書であることは解っているが、効率良いイニシエーションを行うためにはどのような先導構造が必要か、さらにその先導構造とリボソームの間はどのような分子認識が行われているか？などの命題や、あるいはあるtRNAをアミノアシル化するのはそれに対応するアミノアシル合成酵素であり、それが多数存在する他のtRNAの中で、目的のtRNAを間違えずに認識できる⁴⁰、ということは実験的に確かめられているが、ではそのような高い分子認識の信頼性を得るためには具体的にtRNAのどこをどのように見分けているのか？などの命題についてはまだ良く解らない部分が多い。これらの例以外にも多くの解明されるべき基本的な事象が存在している。

ところがこれまで研究の主流であったこれらの対象では、表現型的に進化していると思われる真核生物はおろか、表現型的には下等であると思われる原核生物についても、分子レベルではかなり複雑に進化している。その意味は分子間の認識において、最も基本的な認識機構以外に様々なオプションが付加しているということである。したがってこれらは蛋白質合成にとって最も初歩的、基本的な機構を調べるという意味では必ずしも適当な題材でないかもしれない。

それに対してミトコンドリアの系は上で述べたように、大変単純化された蛋白質合成系である。それらは様々なオプションを廃した系であると思われる。従って、それを調べることから蛋白質合成に必要な最低限のファクターにいつそう接近できる可能性を持っている。例えば共通塩基は従来のミトコンドリア以外の対象において、*tRNA* に必須のファクターであるように考えられていたが、ミトコンドリアの*tRNA* を見る限り最低限必要なものではないことが解る。また *tRNA* SerAGY などは *tRNA* 構造を大胆に簡略化しても、*tRNA* としての機能を保てることを示している。さらに4コドンファミリーにおける U-Nwobb1e 塩基対はアンチコドンとコドンの対合の最低限の ΔG を求めることに良い指標を与えてくれるかもしれない。あるいはミトコンドリアに存在する変則遺伝暗号について、その差異の原因を詳しく調べれば、遺伝暗号生成の良い情報を得られる可能性がある。

しかし現状ではミトコンドリア蛋白質合成系をその諸因子の再構築などによって分子レベルで調べることは、最も研究の進んでいる酵母や脊椎動物についても、あまり進んでいない。

脊椎動物などの研究の成果につづいて解析された、数種類の下等生物、カビ、植物などの結果は、脊椎動物や節足動物で解明された変則遺伝暗号などの蛋白質合成に基本的ないくつかのファクターの特殊性があらゆる生物のミトコンドリアにおいて必ずしも共通するものではなく、ミトコンドリア間においてもその生物

の進化段階によって、これらのファクターが変化していることを示した。従って様々な進化段階に位置している生物を調べると、それらの変化の過程を追跡できる可能性がある。それにもかかわらず下等生物と脊椎動物の間の進化段階の生物種についての研究は遺伝子レベルにおいてもほとんど行われていなかった。

本研究においては、キョク皮動物であるヒトデの遺伝子構造を明かにした。ヒトデを選んだ理由としては、上で述べたように下等生物と脊椎動物のような進化的に大きなギャップのある生物種間の比較ではなく、ある程度の高等動物と脊椎動物などの間においては遺伝暗号や異常tRNAの構造にどの程度の差異があるか？ その差異が下等生物から脊椎動物に至る進化段階によって連続的に変化したものであるのか？あるいは進化段階によって連続的に変化するのではなく、独自の方向に進化したファクターなどがあるのか？などを調べることであった。

脊椎動物ではいくつかの生物種においてゲノムの全構造が調べられており、その情報を元に蛋白質合成系の研究が進められている⁴¹。また最近かなり広い範囲の生物種についてミトコンドリア蛋白質合成系の基礎情報が報告されてきている¹⁰⁻²⁹。ミトコンドリアの蛋白質合成系を調べるにあたっては、ミトコンドリアゲノムがさほど大きくないことから、ゲノム構造の知見を得ることから始めるのが効果的な道筋である。キョク皮動物についてはこれまでゲノムの構造やtRNAの構造などが知られておらず、まずゲノム構造を調べることによりキョク皮動物ミトコンドリア蛋白質合成系の諸要素のデータを得ることを目的とした。また得られた情報を他の生物と比較し、その共通構造、差異構造を調べることにより、分子認識の機構、進化などについての基礎的なデータを得ることを目的とした。

動物細胞の断面模式図

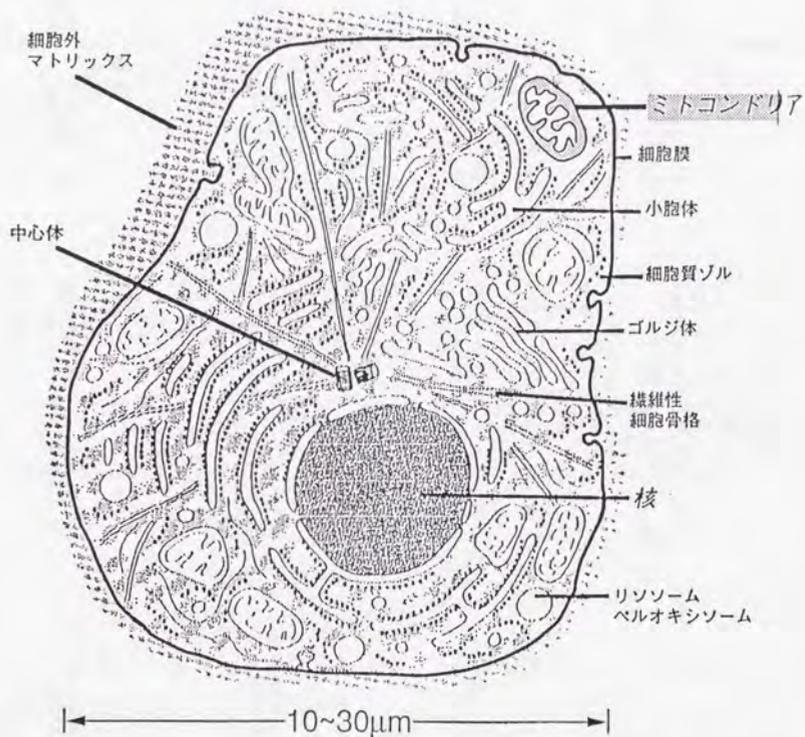


図1 動物細胞の断面模式図

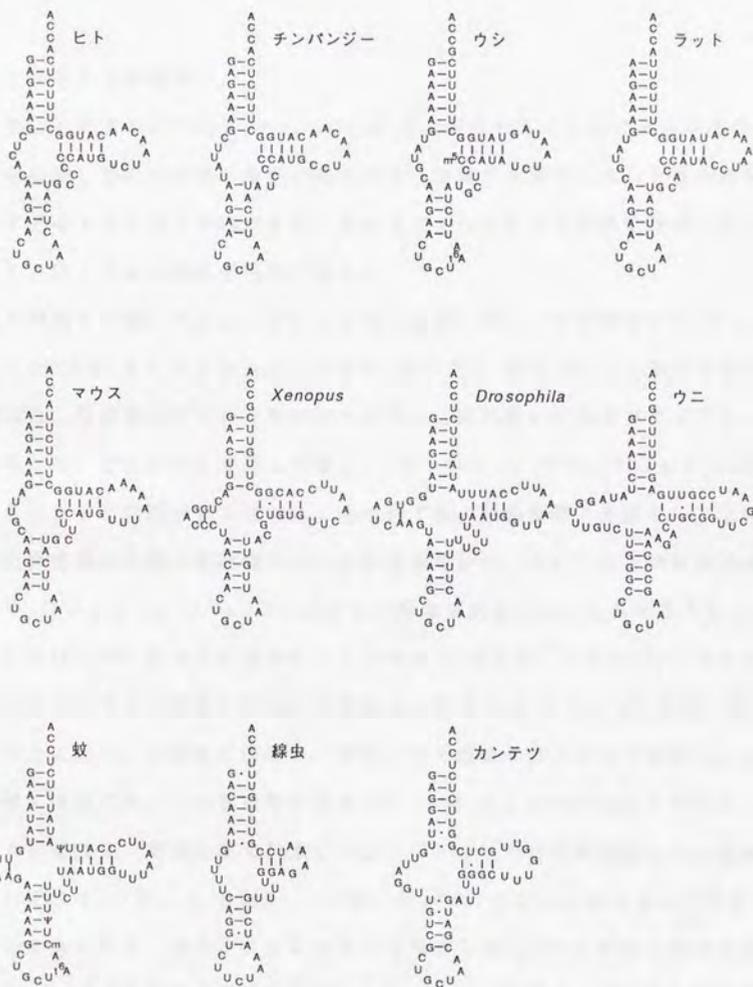


図2 種々の動物のミトコンドリアのAGNに対応する tRNA Ser のクローバーリーフ構造

II. 実験操作

1. ミトコンドリアの精製²⁰

イトマキヒトデ (*Asterina pectinifera*) を産卵前の五月上旬に三浦半島の馬堀海岸から採取した。その場で解剖し卵巣あるいは精巣を摘出した。それをその場でドライアイス上に置き凍結させた。その後は -80°C の冷蔵庫で保存した。1回(年)で約2Kgの卵巣が回収できた。

保存した卵巣を1回に10g~200gづつ分取し約7~8倍量のSE buffer (0.25M Sucrose, 1mM EDTA pH 7.5) 中で氷解させ、懸濁した。懸濁液をテフロンホモジナイザーに数回通してホモジナイズし、細胞を破壊した。これを低速で遠心分離 ($1,000 \times g$, 0°C , 10分) した。上精(ミトコンドリア画分)を回収し、沈澱物(核、細胞壁等)を捨てた。上精に対しこの低速遠心分離を数回繰り返して沈澱物を除いた。次にこの上精に高速遠心分離 ($8,000 \times g$, 0°C , 20分) し沈澱物を回収した。これをSTE buffer (0.3M Sucrose, 10mM EDTA, 10mM Tris-HCl pH 7.5) に懸濁し数回、高速遠心分離 ($8,000 \times g$, 0°C , 20分) することにより、沈澱物を洗浄し、回収した(粗ミトコンドリア画分)。上精(脂質等)は捨てた。この沈澱物を数倍の0.3M Sucrose+TSE bufferに懸濁し、蔗糖密度勾配遠心によりミトコンドリアを精製した。密度勾配はそれぞれ2.0M、1.8M、1.0M、0.6M sucrose+TSE bufferを4ml、4ml、12ml、10mlづつ、ベックマンSW28用の遠心チューブ(38ml程度入る)に下から静かに重層し、最上層に懸濁液数mlをのせた。これに超遠心分離 ($25,000 \text{rpm}$, 4°C , 30分) した。ミトコンドリアは2.0Mと1.8Mの層の境界に層を形成しており、その部分を

回収した。回収液を5倍量のTSE bufferで薄め、高速遠心分離(8,000×g、0°C、20分)により沈澱物を回収した。(精製ミトコンドリア画分)これらの粗、精製ミトコンドリアを以後の実験に用いた。

2. ミトコンドリアDNA²⁰、RNAの調製

得られた粗ミトコンドリアを1% N-sodium lauroyl sarcosinate、200mM EDTA、溶液に懸濁した。溶液1ml当たり0.5mgのprotenase Kを加え37°Cで保温した。オレンジ色の懸濁液が黒ずんできた時点で保温をやめ、フェノール、クロロホルム抽出を行った。0.1倍量の酢酸ナトリウム(pH6.0)を加え、さらに2倍量のイソプロパノールを加えた。しばらく室温におくと、黒ずんだ下層と白い上層の二層に分かれるので、黒ずんだ下層を回収し12,000rpmで遠心分離し沈澱を得た。この沈澱物を70%エタノール-水でリンスし、真空乾燥した。これをTE bufferに溶かし、RNase Aで消化しRNAを分解した。0.6倍量のPEG溶液(20% PEG 6000、2.5M NaCl)を加え、1時間氷冷した。12,000rpmで10分間遠心分離し、上清を捨てた。沈澱物を70%エタノールでリンスし、真空乾燥した。この段階ではミトコンドリアDNAは核DNAとの夾雑物であるがクローニングにはこれを用いた(粗ミトコンドリアDNA)。精製ミトコンドリアDNAを得るためには以下の操作によった。

この粗ミトコンドリアDNAを適量のTE bufferに溶かし1.02倍量(w/v)のCsClと1/14倍量(v/v)の10mg/mlエチジウムブロマイド溶液を加え日立の卓上超遠心機で遠心(100,000rpm、18°C、20時間)した。チューブに針をさし、下から0.3~0.4mlずつ分画した。各画分をアガロースゲル電気泳動レクロズドサーキュラー(CC)mtDNA画分を確認した。この画分を再び上述の条件で遠心し、CC画分を回収した。これ

からブタノール抽出によりエチジウムブロマイドを除いた。これに2倍量のTE bufferを加えた後に2倍量のエタノールを加え遠心分離した。沈澱を70%エタノールでリンスし、真空乾燥した。これをそのまま、または滅菌水、あるいはTE bufferに溶かし-20°Cで保存した(精製ミトコンドリアDNA)。

ミトコンドリアDNAを単離するには、このようにCsCl-エチジウムブロマイド中において超遠心分離により精製するのが常法であり、何度か試みたが絶対量が少なくなるため、最終的にクローニングには採用しなかった。

ミトコンドリアのトータルRNAの回収には精製ミトコンドリアを0°Cで1% SDS、10mM Tris-HCl pH7.5、25mM EDTA溶液中に懸濁し、10分間放置した後、500μg/mlになるよう、protease Kを加え、37°Cで30分間保温した。それからフェノール、クロロホルム処理し、エタノール沈澱し回収した。

3. ミトコンドリアDNAのクローニング

イトマキヒトデのミトコンドリアDNAに対する数種類の制限酵素による切断パターンは姫野らによりすでに報告されており²⁰、BamHI A、Cの二断片はpBR 327にクローンされた形で姫野氏から恵与された。クローニングの出発材料としては粗ミトコンドリアDNAを用いた。その理由は精製ミトコンドリアDNAは極めて量が少なくアガロースゲル電気泳動ではほとんど確認できなかったこと(もっとも粗ミトコンドリアDNAも少なかった)、制限酵素パターンがある程度解っていること、すでにクローニングされた断片からサザンハイブリダイズのプローブが得られることなどによる。粗ミトコンドリアDNAの存在はサザンハイブリダイズにより確認した。ミトコンドリアDNAをBamHI-XbaIで消化しpUC18⁴²にライゲーションし強制クローニングを試みた。このベクターは巻き戻りをふせぐため、アルカリホスタファアーゼ処理したものをを用いた。

その他、以下に用いるベクターはダブルダイジェションのものを含めて（一方の酵素で切れてないものも多いので）アルカリホスタファアーゼ処理し、巻き戻りのバックを減らした。その結果2個のEcoRIのサイトを含む約2Kbpの断片を得た。制限酵素パターンからその断片はBamHI B断片中の一部であると推定され、シーケンスの結果、ミトコンドリア由来のものと断定した。残りの2Kbpに当たるBamHI-XbaI断片は200個のプラスミドをスクリーニングした中には見あたらなかった。次にPstI-BamHI、EcoRI-BamHIで強制クローニングを試みた。これらについては、それぞれEcoRI-EcoRI、EcoRI-XbaIの断片からランダムプライマーエクステンション法⁴³によりプローブを作成し、コロニーハイブリダイズによるスクリーニングを行った。約2,000個のコロニーからはクローンは得られなかった。さらにEcoRI-XbaIのプローブを用いてEcoRI-SalI断片のクローニングを試み、コロニーハイブリダイズによるスクリーニングを行った所、クローンを得た。最後に残ったSalI-BamHIのクローニングをするため再びXbaI-BamHIのコロニーのライブラリーに対しXbaI-SalIプローブとしてスクリーニングをおこなったがクローンは得られなかった。

4. Polymerase Chain Reaction (PCR) 法^{44,45}による未クローニング領域の増幅

クローニングできなかつたSalI-BamHI部分の両側にそれぞれのサイトから約70bp離して、合成DNAプライマー（24mer、25mer）を作成しプライマーとした。テンプレートのmtDNA（量不明）および両プライマーを等量（mol）混合し、PCRを行った。PCR条件は変性を94℃、1.5分、アニーリングを48℃、2分、伸長反応を72℃、3分とし、25サイクル行った。増幅された断片をアガロースゲルより回収し、直接シーケンスのテ

ンプレートとして用いた。(図3)

5. シークエンスのための deletion mutantの作成^{46,47}

BamHI A断片はEcoRI、HindIIIで消化し、0.7Kbp~2.2Kbpの断片に切り縮め、pUC18、19にサブクローニングした。これらやBamHI-XbaI、EcoRI-SalIはそのまま、あるいは平滑末端化してSmaIあるいはHincIIサイトに挿入しなおして、ミュータント作成に用いた。ミュータントの作成は以下のように行った。

まずインサートの入ったプラスミドをアルカリ法により大量調製した。つぎにシークエンスプライマーのbinding サイト、3'突出型制限酵素サイト、5'突出型制限酵素サイト、挿入断片の位置関係にある2種類の制限酵素を選び(挿入断片中にないもの)、約10~20 μ gのプラスミドを3'突出型の制限酵素で最初に消化した。アガロースゲルで完全消化したことを確認後、充分量の5'突出型制限酵素による消化を行った。フェノールクロロホルム処理、エタノール沈澱後、100 μ lのExoIIIヌクレアーゼ bufferに溶かし、ExoIIIヌクレアーゼで消化した。最適な酵素量は、予備実験により見積もった。1~20分間にわたり1分ごとに5 μ lづつ分注し、氷中で反応を停止させた。1分~5分、6分~10分、11分~15分、16分~20分のものをそれぞれまとめて(25 μ l)、75 μ lの1.3 \times mung beanヌクレアーゼ bufferに混合させ、65 $^{\circ}$ C、5分間インキュベートし、ExoIIIヌクレアーゼを失活させた。次にmung beanヌクレアーゼにより37 $^{\circ}$ Cで30分間消化した。この段階でアガロースゲル電気泳動を行い、経時的にdeletionが進んでいるかを確認した。うまくいっていない時は酵素量、反応時間等を調整して再度、実験を行った。フェノールクロロホルム処理、エタノール沈澱、乾燥後、T4 DNAポリメラーゼ bufferに溶解し、T4 DNAポリメラー

ぜにより末端を平滑末端にそろえた。これを *ligase* により環状化し、大腸菌のコンピテントセルを形質転換した。コロニーを100個程度サンプリングして、プラスミドを調製し、階段状にプラスミドを選択した。mung beanヌクレアーゼの代わりにS1ヌクレアーゼを用いる場合は、分注後そのまま熱処理によって、ExoIIIヌクレアーゼを失活させ、エタノール沈澱を行った後にS1ヌクレアーゼによる消化を行った。その後の操作は同様に行った。

6. サブクロニングによるテンプレートDNAの調整

適当な *deletion mutant* の取れなかった領域に対するテンプレートの作成は、様々な制限酵素断片をサブクロニングすることにより一部補完した。

7. 合成シークエンスプライマーによるシークエンス

5、6の操作によってカバーできなかった領域に対しては、その近傍に合成プライマーを作成し、シークエンスを行った。またPCR断片のシークエンスはすべて合成プライマーを用いた。

8. *dideoxy sequence* 法による塩基配列の決定^{9,48}

シークエンスは市販のキットを用いて行った(宝酒造の7-DEAZA Sequencing Kit、BRL社のSequenase ver. 2)。二本鎖テンプレートはプラスミドの場合、アルカリ変性を行い、PCR断片については94℃で熱変性させた。用いたテンプレート量は約1 pmolであった。プライマー量はプラスミドの場合0.5 pmol、PCR断片の場合5 pmol用いた。テンプレートとプライマーのアニールはプラスミドの場合、変性させたテンプレートをプライマーと混合後、65℃、5分間インキュベートした後、室温におい

て1時間放置した。またPCR断片の場合、テンプレートの熱変性の際にプライマーを同時に混合させておき、変性後、すぐに氷冷し1時間放置した。シーケンス反応はキットのマニュアルに従った。アクリルアミドゲルは0.4 mm×30 cm×60 cmのゲルを用い、読みたい領域に応じて5%~8%に調整した。

9. 16S rRNAのS1マッピング⁴⁹

16S rRNAのマッピングにはトータルミトコンドリアRNAを用いた。S1プローブはサイズの予想される3'端からそれぞれ200bpほど離して、合成DNAプライマー2本を作成し、PCR法によってこの間の約400bpの断片を増幅した。次にこの断片に対し16S rRNAに相補的な方のプライマーだけを用い、同時に³²P dCTPのラベルを導入して非対称PCRを行った。これを7M Ureaアクリルアミドゲルから回収して、S1プローブとした。次にこのプローブと数μgのトータルRNAを80%ホルムアミド、0.4M NaCl、50mM PIPES pH6.4、1mM EDTA溶液中で混合し、90°Cで3分加温した後、46°C、48°Cで3時間保温し、その後充分量のS1 buffer+20~30μgサケ精子DNAを加え、1、10unitのS1ヌクレアーゼにより30分消化した。これをアクリルアミドゲル電気泳動によって塩基長を検定した。

10. 16S rRNAのプライマーエクステンション法によるマッピング^{50,51}

予想される16S rRNAの5'端から約150bp下流に相補的な合成DNAプライマーを作り、末端をラベルする。これをトータルミトコンドリアRNAと10mM Tris-HCl pH8.3、1mM EDTA、0.25M KCl溶液中で混合し、60°Cで1時間保温後、室温で1.5時間放置した。その後これを希釈、Mgイオン、dNTPの添加など溶液を逆転写酵素(AMV)の反応

条件に合わせて、逆転写酵素を30分間作用させた。こうして伸張したDNA断片をアクリルアミドゲル電気泳動によって塩基長を調べた。



図3 PCR法による長クローニングの増幅

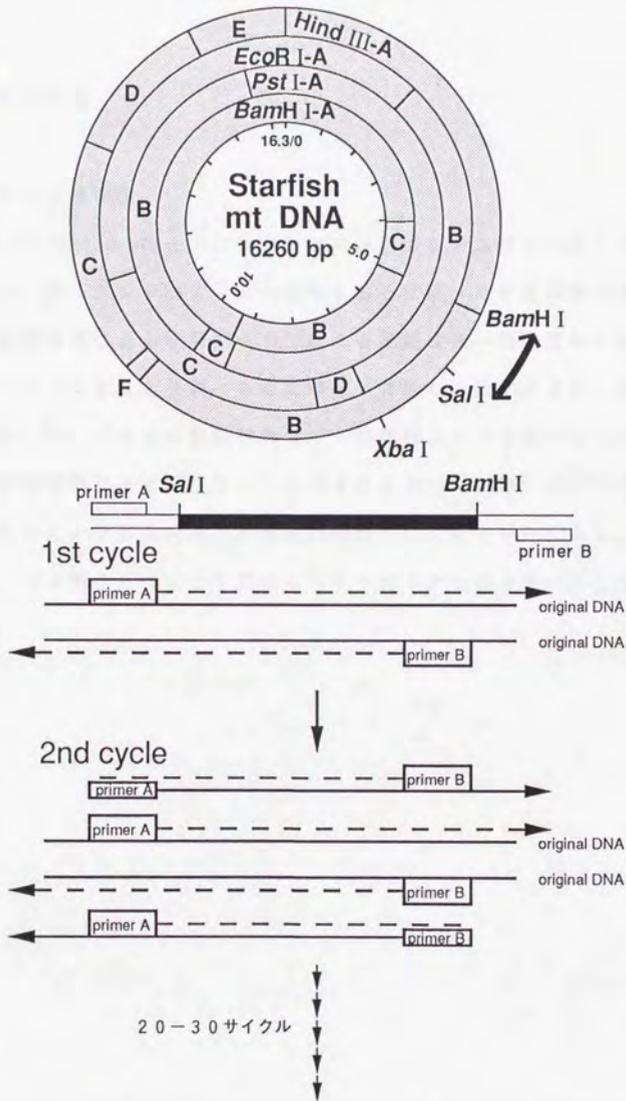


図3 PCR法による未クローニング領域の増幅

III. 結果および考察

1. 制限酵素の切断部位

*Asterina pectinifera*のミトコンドリアDNAのゲノムサイズは16、260 bpであった。表1にシークエンスの結果をもとに調べた制限酵素による切断部位を示す。制限酵素による切断パターンによる解析は同一生物種などにおいて遺伝子の多型やその分散などを調べるのに簡潔で有効な手段であるが、本研究では特に調べなかった。これらの制限酵素サイトの情報はサブクローニング等に利用した。この制限酵素サイトはクローニングされた単一（正確にはいくつかの断片に分けてクローニングしたので、それぞれの断片について単一である）の個体のものである。また荒木らによって調べられたキヒトデとのパターンとの共通性は乏しかった。

Mati	TGCCGA	1: 6813	5195					SfII	GGCCNNNGGCC	0: 7088	15120		
MatiF	CCTNAGG	2: 1545	2222	2430	3547	4408	5311	5820	CCCGGG	3: 4576	6587		
MvAI	CCLLGG	18: 1075	7225	7706	10103	11331	11555	11940	GTATAC	4: 5185	1287	1652	10801
		18887	7825	12573	13708				TACTAC	3: 8159	5688	14857	
		0: 12583	12506						ACTAGT	1: 12584			
NaeI	GGCGCC	1: 10553							CCATCC				
NarI	CCGCCG	11: 7088	7089	7250	9370	12898	13454	14023	AATATG	17: 1144	2544	2982	3656
NciII	CCMGG	14801	15120	15121	15887					17: 8619	8986	10450	11028
NcoI	CCATGG									14294	14903	15580	11028
NdeI	CCATGC	4: 180	1447	4898	13024					3: 1500	8392	13362	
NfoI	RCCGCT	1: 10553								15: 507	1750	3225	3762
NheI	CTACAC	7: 257	13889							7342	9798	10412	12999
NhaIII	CATC	25: 7211	9411	11318	11378	12101	13081	14788		4: 4045	8613	12688	12787
		14282	12565	12319	13080	13401	14089	14438		15: 1317	1298	5531	8632
		14560	14857	15421	15922					0: 12871	13332	14821	
NotI	GGCGCCGC												
NruI	TCCGCA	0: 131	3047	6552	7896	8188	10303	13070					
NspIII	CJCGCG	8: 131	3047	6552	7896	8188	10303	13070					
NspI	RCATGY	1: 1442	7512	12564									
NspI	7524IV	3: 946	3455	3959	4805	7084	9213	14730					
PmaCI	CACGTG	1: 3107	3455	3959	4805	7084	9213	14730					
PvuII	RGGLGCT	4: 3107	6257	8220	15315								
ParII	CGGLCCG	1: 10742											
PstI	CTCCAG	3: 6172	11883	12497									
PvuI	CGATCG	1: 5638											
PvuII	CGATCG	2: 131	3047										
RsaI	GTAC	25: 43	53	924	1513	2043	3874	4549					
		758	850	854	883	1193	1539	1759					
		13983	15158	15406	15792	10459	11305	12794					
SacI	GAGCTC	3: 10540	12416	15149	15792								
SacII	CCCGCG	5: 6252	7896	8188	10303	14542							
Sall	GTCGAC	1: 15369											
SauIAI	GATC	4: 22	852	1113	2095	2969	3250	4303					
		4822	4852	5264	5495	5513	5639	5643					
		6655	7651	7718	7822	8053	9080	10290					
		12413	12926	13273	13570	13971	14287	14348					
		15195	15275	15710	14441	14494							
SauII	CCTNAGG	2: 15445	5195	12575	12710	14441	14494						
ScaI	AGACTT	5: 3873	6176	8979	12753	15197							
SclI	GCAC	7: 6514	6884	8665	10554	12277	15382	13993					
SerFI	CCNGG	29: 1075	2222	2430	3547	4408	5311	6520					
		6887	7088	7089	7250	7628	7706	8370					
		10403	11331	11555	11940	12283	12506	12673					
		12898	13454	14023	14601	15120	15121	15708					
		0: 15867											
SfuI	GGCCCC	3: 1327	13172										
SfuII	GGCLCC	3: 1584	13016	16056									
SfuIII	GLCCCC	8: 1056	1455	2996	4070	10640	12416	15749					
SfuI	GLCLCL	18: 118											
SecI	CCNNGG	38: 230	607	698	1259	1521	1687	1750					
		2430	3225	3413	3547	3644	3762	4047					
		7139	7532	7511	9683	7087	7088	7089					
		10412	12372	12309	1349	6186	6196	10303					
		15120	15391	15882	1449	14249	14542	14600					
SfaNI	GCATC	15: 254	634	4102	5053	5752	5899	6116					
		7700	9274	10833	11375	11577	11646	14109					
		14785											
SfaNI	GATCC	10: 1267	3215	10092	10943	11462	12363	12774					
		13458	14106	14592									

L=A/T ,R=A/G ,J=A/C ,E=T/G ,Y=T/C , M=G/C ,N=A/T/G/C

2. 塩基組成と鎖 (Strand)

イトマキヒトデミトコンドリアゲノムの塩基組成を表2に示す。表2には数種類の脊椎動物、2種類のウニおよびショウジョウバエの塩基組成も示した。表で示されるように脊椎動物等では相補鎖間でかなりの塩基の偏りがみられる。

AloniらはヒトHeLa CellのミトコンドリアDNAをアルカリ条件下で相補鎖を分離し、CsCl密度勾配遠心により沈降度を調べた⁵²。その結果、これについては、相補鎖が沈降度により明確に分離され、重い方をHeavy Strand (H鎖)、軽い方をLight Strand (L鎖)と定義した。AloniらはさらにRNA転写物の解析を行い主要なRNA転写物がH鎖にハイブリダイズすることを示した。Brownは脊椎動物、非脊椎動物など数種類の生物のミトコンドリアDNAについて同様の実験をおこない、沈降係数と塩基組成の関連を調べた⁵³。それらの結果、脊椎動物ではH鎖、L鎖に明確に分離され、それぞれの鎖と塩基組成に、L鎖にGが少ないなどの、関連があることが解かった。さらにこの頃から数種類の脊椎動物の全塩基配列が報告されはじめ、その遺伝子構造が明かになるに従い、ゲノムにコードされている遺伝子のほとんどはL鎖にコードされていることが示された。従って脊椎動物ではL鎖をSense Strandと呼ぶことも多い。一方ショウジョウバエでは相補鎖の密度的な差はほとんどなく、またウニにおいてもその差は小さいものであった。

一般に直鎖状の2本鎖DNAのCsCl中の浮遊密度は塩基組成から求める実験式がいくつか提唱されているが、生体サンプルについては良く実験事実と合う^{54,55}。1本鎖DNAの密度と塩基組成とを詳しく調べた例はあまりないが、数種類の1本鎖ファージの浮遊密度が解っている。しかしそれらはG+C量と浮遊密度の相関として求めてあるため、あまり参考にならない。G+C含量という観点からみると、2本鎖DNAはワトソン-クリック型のベースペアを形成しているわけであり相補鎖のG+C含量は等しく、ミトコンドリアDNAもこの例外では

ない。従って2本鎖DNAについてはG+C含量が解れば一意的に各塩基の含量が決まり、それはA=T、C=Gとなっている。それに対し1本鎖DNAの場合、G+C含量という観点から論じるのは、不適當である。1本鎖DNAの浮遊密度を塩基組成と関連づけるのには、A、G、C、Tの正確な含量を求める必要があるであろう。さらに2本鎖DNAにおいても合成オリゴヌクレオチドによる浮遊密度の研究の結果、塩基組成だけではなく塩基の順序が浮遊密度に大きく影響していることが報告されている⁵⁶。それに対し生体サンプルの場合CsCl中ではほとんど例外なく、G+C含量と浮遊密度は直線関係に乗っている。数bpのオリゴヌクレオチドの場合、特定の配列に依存するスタッキングの強弱などの関与が大きいであろう。それに対し生体DNAサンプルについては、部分的にみると配列に依存する効果は存在するであろうが、マクロな観点から見て、G+C含量という統計的な観測に沿っているものと考えられる。一本鎖のDNAの場合のこれらの状況は基本的に同じであろう。

2本鎖DNAにおける、それぞれの鎖別の浮遊密度を考える場合、互いのA+T、G+C含量は等しいので、それぞれAに対するT、Cに対するGという観点で考察するのが都合が良い。脊椎動物の塩基組成を見るとL鎖においてはCの存在率はGのものより10%から20%多くなっている。逆にいうとH鎖においてはGがCより10%~20%多くなっている。さらにAとTについてもL鎖ではAがTより5%~6%多いことが脊椎動物に共通している。相補鎖間の浮遊密度の差に対してこれらの2つの要素が加算的な効果を及ぼしているのか、減算的にきいているのかは、実験的に確かめるべきことであるが、少なくともこのうちのどちらかは浮遊密度の差の原因となっているであろう。ヒトデにおける塩基組成を脊椎動物と比較すると一方の鎖においてCがGより約10%多くなっている。またAはTより約4%多くなっており、脊椎動物と同じ傾向となっている。これらの値は2つの要素が加算的に働こうが、減算的に働こうがに互いの相補鎖間で

浮遊密度に脊椎動物と同程度の差があることを予想させる。またシヨウジョウバエやウニにおいて相補鎖間の浮遊密度にあまり差がないのは、これらのG-C, A-T間の塩基組成の差が脊椎動物に比べ相当小さいことで説明がつくであろう。

本研究では実際に浮遊密度を求めていないが、脊椎動物のL鎖の塩基組成と同様の塩基組成となっている方をヒトデミトコンドリアDNAのL鎖と定義し、その相補鎖をH鎖とした。鎖(Strand)間の塩基組成の差やその生成原因についての詳しい考察は9章で述べる。

表2

高等動物ミトコンドリアゲノムの塩基組成

	A	C	G	T	%
ヒト	30.9	31.2	13.1	24.7	L鎖
ウシ	33.4	25.9	13.5	27.2	
マウス	34.5	24.4	12.3	28.7	
アフリカツメカエル	33.1	23.5	13.5	30.0	
ヒトデ	32.4	24.6	14.1	28.9	
ウニ(P.I)	30.8	22.5	17.2	29.5	
ウニ(S.P)	28.8	22.7	18.4	30.2	
シュウジョウバエ	39.5	12.2	9.3	39.1	

3. 遺伝子の同定

図4に *A. pectinifera* の全塩基配列と同定された遺伝子を示す。また蛋白質遺伝子から推定されるアミノ酸配列を付記した。同定された遺伝子は22個の tRNA 遺伝子、13個の蛋白質遺伝子〔NADH dehydrogenase subunit (ND) 1~6、4L; Cytochrome oxidase subunit (CO) I、II、III; ATPase 6、8; Cytochrome b (Cyt b)〕、および2個の rRNA 遺伝子 (12S、16SrRNA) であった。これらの遺伝子によりゲノム全体の約95%が占められており(表3)、その他の部分には遺伝子はコードされていないものと考えられた。この遺伝子構成は脊椎動物、節足動物、ウニで報告されているものと一致していた。

遺伝子間には tRNA Thr 遺伝子と 16SrRNA 遺伝子の間に約450bp、tRNA His 遺伝子と ND4 遺伝子の間に140bp、ND5 遺伝子と ND6 遺伝子の間に51bpの3カ所に比較的大きなスペーサーが見られた。さらに tRNA Ser (UCN) 遺伝子と COIII 遺伝子、tRNA Leu (UUR) 遺伝子と tRNA Gly 遺伝子、tRNA Pro 遺伝子と COI 遺伝子の間にそれぞれ30bp程度のスペーサーが存在した。しかし、それ以外のスペーサーは13bp~1bpの小さなものであった。また tRNA Ser (AGN) 遺伝子と ND5 遺伝子の間、ND1 遺伝子と tRNA Leu (UUR) 遺伝子の間、tRNA Cys 遺伝子と tRNA Trp 遺伝子の間など9ヶ所においては遺伝子が完全に隣接しており、スペーサーのない構造になっていた。さらに COIII 遺伝子と tRNA Ser (UCN) 遺伝子、tRNA Val 遺伝子と tRNA Cys 遺伝子、tRNA Ala 遺伝子と tRNA Leu (CUN) 遺伝子など5カ所では遺伝子が互いにオーバーラップしていた。また植物のミトコンドリアに見られるようなイントロンは存在しないと考えられた。

このようにヒトデのミトコンドリアゲノムは各遺伝子がたいへん密にコードさ

れており、経済的な配置といえた。表4に様々な生物のゲノムサイズを示す。この表に示したように高等植物、カビ、酵母、原生動物などはゲノムサイズがヒトデよりはるかに大きく、また構成遺伝子もヒトデより多い。しかしゲノムサイズが大きい主な原因は遺伝子間に存在するスパーサーや遺伝子内のイントロンなどがゲノムサイズの大きな部分を占めていることにある。それに対して脊椎動物、節足動物、ウニなど高等動物はすべてゲノムサイズが16 Kbp前後となっている。またこれらの遺伝子構成も一致しており、イントロンがなく、スパーサーがほとんどない構造となっている。さらにいくつかの遺伝子のオーバーラップはこれらの高等動物すべてに見られる。つまりゲノムの縮小化、経済化はこれらの高等動物に共通しており、そのゲノムは非常に経済的な遺伝子配置となっている。さらに線虫においてはゲノムサイズが約14 Kbpであり、ATPase 8 遺伝子が欠けている点で他の高等動物と異なっているが、経済的な遺伝子配置になっているという点では全く同様である。ヒトデにおける遺伝子配置の経済化はこれらの高等動物の事情と基本的に同一と言えた。

しかし細かく見るとウニはゲノムサイズが15.7 Kbpであるのに対しヒトでは16.6 Kbp、アフリカツメガエルでは17.5 Kbpなどのゲノムサイズの分散が見られる。これらの高等生物のゲノムサイズの分散の最も大きな原因は遺伝子サイズの分散というより、スパーサー領域のサイズの差であるが、その中でも特にD-loopの大きさの差によるものと思われる。ウニではD-loopのサイズが120bpであるのに対し、ヒトでは1 Kbp、アフリカツメガエルでは約2.1 Kbpとなっており、このD-loopのサイズの差がほぼゲノムサイズの差に一致していることが解る。全ゲノムからスパーサー領域を除くとこれらの高等生物ではすべて15.3 Kbp~15.5 Kbpに分布しておりコーディング領域のサイズはほぼ等しくなっている。ヒトデの場合、非翻訳領域の合計は約0.8 Kbpであり、これを除いたコーディング領域は15.5 Kbp (表3)と

他の高等動物ミトコンドリアゲノムとほぼ等しくなっていた。非翻訳領域のうち16S rRNA遺伝子とtRNA Thr遺伝子の間に存在する約450bpが大ききから見て、ヒトデミトコンドリアゲノムのD-loopに相当するものと思われる。

ヒトデにおいて特徴的なのはこのD-loop以外にも、のべ約350bpもの非翻訳領域がある点である。脊椎動物ではD-loop以外の非翻訳領域としてはtRNA Asn遺伝子とtRNA Cys遺伝子の間にある20~30bpのL鎖の複製開始の認識部位が共通して存在する⁵⁷以外は数bp以下の小さなスパーサーしか存在しない。さらにこのL鎖複製開始点のスパーサーはトリには保存されておらず¹⁸、トリにおいてはD-loopが唯一の大きなスパーサーとなっている。ショウジョウバエのミトコンドリアゲノムではD-loop以外に10bp以上のスパーサーが8ヶ所存在するが最大のもので31bpである。ヒトデでは10bp以上のスパーサーは7ヶ所存在するが、特にtRNA His遺伝子とND4遺伝子間の140bp、ND5遺伝子とND6遺伝子の間の51bpなどD-loop以外の大きなスパーサーはヒトデにしか存在せず特徴的である。同じキョク皮動物であるウニのミトコンドリアゲノムではヒトデに見られるこの2ヶ所のスパーサーは存在せずヒトデとの差異となっている。しかし、これらの中で、ND4遺伝子とtRNA His遺伝子間のスパーサーはイトマキヒトデ以外の2種類のヒトデでも保存されていることが報告され、何らかの機能的な役割を果たしていると考えられる。またヒトデとウニでは数ヶ所において10bp以上のスパーサーが保存されており8章で詳しく考察する。

以上のように細部においては高等動物間で差異があるものの、基本的に高等動物はゲノムの縮小化の方向に進化したのであろう。

3010 3020 3030 3040 3050 3060 3070 3080 3090 3100 3110 3120
 CCCCACACCTTTAAAACTCTCCCTAGTGGTCTTAAATAGCCCGACTGGGAAGTCTGCCCAATTTGGTCTACACCGTACTCCCGCCCAATGGAGGTCACACCTCTCTCTC
 FHTFLNMLLLLVGLLIIAAAGKBSAQFLHLPWLPFAAHEDGPTFLV

3130 3140 3150 3160 3170 3180 3190 3200 3210 3220 3230 3240
 AGCATTCTCAGCAGTGCATATGGGTAGGCTGAATTTCTCTCATTGCACTAAGCCCTTATGCGCTTCTAGAAATTTAAAACTCTAGCCCTAACTCTGGATCAATAACA
 SIRHSSSTHVVVAGIIFLLLIRLSPFLYACSSNFFHTMCLILIGSLT

3250 3260 3270 3280 3290 3300 3310 3320 3330 3340 3350 3360
 GCTATTTGGCCGACACCGCTATCTCCAGCAGCTAAAGAGATAGTGGCTATTCTACTAAGACACTAGACTGATGGTGGTCTATGGACTAACCCACCAADATA
 AIFAAATTAISQHDIEKIVAYSTTSQLGLMHWVAIIGLNQPSI

3370 3380 3390 3400 3410 3420 3430 3440 3450 3460 3470 3480
 GCATTTCTCATITGCACTCAGCTTCTCAGGCTGACTATTTTATCTCCGGGAGAAATATACATAGCTTAAAGAGGAGCAGACATTCGAAATATGGCCGACTCTCATCTT
 ALFHICTHAFKIAFFKALFLFLBSSGSIIRHSLNLDLHSLNGLHF

3490 3500 3510 3520 3530 3540 3550 3560 3570 3580 3590 3600
 ATCTCTCCACACAGCGCTGTATTTATCTGGGAGAGCTGGCTTCTCAGDAATCTTCTCTGGGTTCTACTGGAGGATTAATCCGAAATAGGACTCCAAATCTCTCT
 ILFNHTAACIILIGSLALSIGIPLFLPGFYSKDLILLEIGLTFNS

3610 3620 3630 3640 3650 3660 3670 3680 3690 3700 3710 3720
 AAATTTGGCGATGCTTACTACTAGCACTATTAACCTGGCTTATCTCATTGGAATAATCTTTGCTCATTAAAGACCTTTCTTTCCCTCTGCCCGAATAAAA
 NFMGIVLVLGLLATLTLTSTVYSFRSIFTFCFIKPFIKXWYSPFLAPLX

3730 3740 3750 3760 3770 3780 3790 3800 3810 3820 3830 3840
 GAAGAAAAAATAGCTAACTAACTCTAAACCGTATGCGCTGGAGATTTATCGGATGGATCTCCTAACTACAGCTCTGGCCCAATTTATCAACTAGCTCTGATTA
 EKNHMLTNRLLNRLALGLTTLISGWLTLNLTLLVPLVPIITISBSVLE

3850 3860 3870 3880 3890 3900 3910 3920 3930 3940 3950 3960
 AAGACTGGCAGCTCTTAACATACAGAGGATCTTTTCAATTCATCTTACAGACTTACCTAAAATCTGACCCCTACTGCATATAAGCAATCATTTACCATAAA
 KTAALLLTLITGVLFSLISILQRETLNLSPTTAYNTNSFTTN

3970 3980 3990 4000 4010 4020 4030 4040 4050 4060 4070 4080
 CAATGATCTACGAAAAATACACACTCTTATTTCTATTCTTATACATATCTTCTTAGCAGCTGAAAAATAGCCGAGGATGAAGAGAAAACTTTGGTGCACAGGA
 QWPHYENISHLIFLFLYYSFTTISLSLSTRNIDRGGWSENI GAQDQ

4090 4100 4110 4120 4130 4140 4150 4160 4170 4180 4190 4200
 ATAGCAATCCCTCTAAGCAGCAAGATGCACTTCCCAACAGCGCTACATAAGCAATCTCCTATTTCTTTTAACTACTATTTATAGCATATCATCTCA
 IATITSSNNAASQSYLTSQTYIKQYLLFLPSLILIIIALGLT

4210 4220 4230 4240 4250 4260 4270 4280 4290 4300 4310 4320
 ACCATAAGCAGCTTTCCCAAGATTTATCTACACACCTTATCTCATACAACTTTATAAAATGGACCTTACTTTGCTTATAAAGCTTTAAGCTATATATCCGCC
 TISQLSLPSFDI *LAKLISYDSG

4330 4340 4350 4360 4370 4380 4390 4400 4410 4420 4430 4440
 ATATGTTAGCAAAAGCAGCACTAAGCCACCAATATATAGCTCCAAATTAATAATCTCCCACTCTCTGATAAATTCGGCCCAACAACTCTTTGTATCAA
 YTLVLA VVLLAVLLIYGGIILLYGGMHSSYLRBSAGVLDSTNTV

4450 4460 4470 4480 4490 4500 4510 4520 4530 4540 4550 4560
 ACCCATCTTTCCCTCACTTAATAAGGGTAACTTAAACCCCAAGNAATAACAGAGACTAAGCACTAAGATCTGGTAAATTTACTAATCTTGGGTACCCCTTCC
 GMSNLEVNILPDLDFNLVLVGMBSIVLSLVLSESLVLEFNFSVSPYREA

4570 4580 4590 4600 4610 4620 4630 4640 4650 4660 4670 4680
 CGATATGGCTTAGATATACAAAAGCAGCACTCTCCCACTATATAAATTAACACCAAGACTATAAAGAAACCCCAACAGCTGCAATACAAAGCCAGACATGCCAC
 SYIATSYVYVFLMGMGYIILVLAIZFGLLISCLCGSLAV

4690 4700 4710 4720 4730 4740 4750 4760 4770 4780 4790 4800
 TAAACTAGCACTAAGCAGCACTAATAAGGGGACAACTATAAAAAACCAAGCTCTCCAAAGGAACTATAAACCAAAACCGTATAAAAAATCTATATATAA
 LVLEGLACGYPPSLSYFVLTLSGLFMDLVLVTYFIH TMTG

← ND 6 Cyt b →

4810 4820 4830 4840 4850 4860 4870 4880 4890 4900 4910 4920
 CCCCCTGAAAGAACACTTTATTTAGATATAAAGACTCTCTTATGATCTCCACCAAGTAACTCTCCATGGGAAATTTGGCTACTTCTAGGACTAGCTTATA
 PLRKNHPLFLFSIINNSLIDLPSFNLNLSIWNHFGSGLGLCLII

4930 4940 4950 4960 4970 4980 4990 5000 5010 5020 5030 5040
 ATACAATAGCTTACCGCTTCTTAGCAATGCACTCTCGAGCTCTCCCTAGCATCTCTCCGTAAGTACATATGCCGCGATGAAATAGCGGCTGCTTGGCGAATC
 IQILITGLFLAMHNTYSDVSLAFSSVSHICRDVNWGLLRNI

5050 5060 5070 5080 5090 5100 5110 5120 5130 5140 5150 5160
 CAGCCCAACAGCATCTTTCTTCTGCTTATCTTCACTTGCAGCAGATTTACTATGCTCTACGTAAAAAGAACTGAAACAGAGAGGATTTCCACTCTCTTA
 HANTASFPFLCLCYLPHIGRGIYYGSSYVNNETWNIGVILLLL

5170 5180 5190 5200 5210 5220 5230 5240 5250 5260 5270 5280
 ACCATGCTAACCGGTTGGTAGTACTGCTACTCCCTGGCTCAAACTGCTCTGAGGAGCACTGAATACTAATATCTGCGGCCCATATAGGACTCTATAGTACAA
 TMLTAFVGVYVLFNQMHSFWDATVITNLSAVVYIGSSIVQ

5290 5300 5310 5320 5330 5340 5350 5360 5370 5380 5390 5400
 TGAATTTGAGGAGGATCTCCGTTCAATGCTACCTCAACCGATTTTGGCTTCCATCTTATTTCCCTTATCAGTGGCTGCTTGTCAATTTCACTATTTCTATCA
 M V W G G F S V L D N T L R F F A H F L P F F I I A A L S I L H L F T L H Q

5410 5420 5430 5440 5450 5460 5470 5480 5490 5500 5510 5520
 ACCGATGCAAAACCGCTGGATAGATCAAGTTTGGCAAGACCCCTTTACAGTACTTTCACTAAGGCAACAGCCATTTATTTCTATTTATGCTATTTAGATGAAATA
 TGSNNPTGIESKPKDPTFPHTYFSTKDJTGFIIILFHLSSII

5530 5540 5550 5560 5570 5580 5590 5600 5610 5620 5630 5640
 GTCTTACTACCCCACTACTAAAGACCCAGCAATTTAAACCCCAAGCACTTGTAAACCCCACTATACAAACCGGATATTTCTTATTTGCTATGCACTTCACTGA
 V L L E S P N L L N D P E R H Y N F A H P L V T P I H I Q P E N Y P L F A Y A I L R

5650 5660 5670 5680 5690 5700 5710 5720 5730 5740 5750 5760
 TCGATCCAAAAAAGCTGGAGGCTAGTAGGCTCTGACTCTCAATCAATGCTTATTTCTTCCCTTTCTACACACTTAAAAATCAAGCCACACATTTCCACAGCATCAAA
 SIPNKLGLGVVLAISILVLLVPLVPLHSTNLRFPASQ

5770 5780 5790 5800 5810 5820 5830 5840 5850 5860 5870 5880
 AGCTGTTTGGCTCTTACTAGATATTACTATCTACTCTGAGGAGACACCCAGTAGAAGCCATTACTACTAGGACAAATTTGCAATAGCATCTGGCTTATCT
 SLFWLLTSTVFIILTWLGLSQPVEDPFIILGQIISSIATYALF

5890 5900 5910 5920 5930 5940 5950 5960 5970 5980 5990 6000
 TTTTACATCTCCCACTAGCATCTGAATGAAAAAGACTAAATTTGCCAAGATAGCTTAATGCAAAAGCAAGCACTGAAACCTGAAATAAAGGTTTAAATCTTTCTTAGC
 FLAPFLASLELNRLLIF

tRNA Phe →

← 16S rRNA

9010 9020 9030 9040 9050 9060 9070 9080 9090 9100 9110 9120
CTAGCTCTTAAATATTTTAAATATATTAATTATTTAATATGATATAAACCAGGTAACAACAATACCAATATCTTATTCGCAATTTACTCTCACACACACCTTCAGA
* M L L S I F V F L C V D L G L S V L F L V G L S L L V L G L S
9130 9140 9150 9160 9170 9180 9190 9200 9210 9220 9230 9240
GGTAAAGCCCCAGTGTGATGGGATTTACTGCTCAGCTCACTAAAAGCTACCAAAATGTTGTGGAAACAAGTCAAACTCTTTAAAGGCTATTGGAGATAAAAAAACCACTTAACA
T F G G T Q N R R S R W S F V V L S H Q P F L T L S S N F A I R L Y F F L S L L
9250 9260 9270 9280 9290 9300 9310 9320 9330 9340 9350 9360
TCTAGCCATAACCAAACTGAGTAAATGAGGATCTTTCTGACTAGGACCTTAAGATATAACTTTATCAAAAACCAAAAGGGCCCAACCCCTTAAGATAAGATCTT
S G M V L V G A L I F C G N E V L C K L S I F K I L F G F L P F L P L G L S L I S
9370 9380 9390 9400 9410 9420 9430 9440 9450 9460 9470 9480
CAAAACACCCCGCTCCCCACAAATATAAAGGATAAACCAGCCACATGAGATAAGACTTTAAATCAAAGCTTTAGCCATAAAAAAACCCAGAGTATTATATATGAAAG
L V L G A G G V L L S L R G V H S L S K L D F S N A M F V S S N I I T Y A G
9490 9500 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600
AACCCAGATAACCGAGCAACTCTACAGAAATACCACTGCTACAAATCAACCAATATGAGCTATAGAAGAAAGCAAAATCTCTCCGATTAGGCTGATTTAAACCTCTCA
V L M V C A V S V S Y G V T S C I W G I H A I S S F A F I K R H Q T Q N L G G W
9610 9620 9630 9640 9650 9660 9670 9680 9690 9700 9710 9720
CCCTCTACAGAGCGGACACTCTTAACAGTGAATAAAGTTATATTTAATCTTTCTACCACTATAGACAACAACAGGCTCTATTTTGGCAAGTATAGATAACCAAGCC
G C V L V S L V G L S A L L S I L N L S E V V Y V L V A F P A I K Q M T S L V L G
9730 9740 9750 9760 9770 9780 9790 9800 9810 9820 9830 9840
CTGTAAACCCACCTCTTGAATCAACTCAGGAATAAATGATGTCAGCAAGAAACAAGCCAGCTTAAAGCCATAAGCCAAAGGATGAGCTGTTTAAAGCTTCAT
Q I F G V G Q I V Q I V D F F W Y H C P F L G L K L G L A L I L M L F S T A Q H L F E M
9850 9860 9870 9880 9890 9900 9910 9920 9930 9940 9950 9960
CAAGCAACAGAGATAATCAACCAAGCTGATTAACCACTCAATACATCCCGCTCAACAGATTCACTAAAATAACTTACCGACACTCACTTACGAGGAGTAAAGCT
L S M S S Y L M A Q I I V V N L V I G A S V S Q V L F Y K V S S E V N R P T F G
9970 9980 9990 10000 10010 10020 10030 10040 10050 10060 10070 10080
ATAGAAGAAGATCCGCAATATAGACAAGTTCATGATGTCACCCAGCAAGGTAACAAGTGGTCTCTCAAGAGCAATAGAGTCCCAATACAGATACAGATTAATACCAT
F S L I F L I S L S T M E L D G V M L T F P M H S S L V L T G L E V V N A I L V M

← ND 2

10090 10100 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200
CAAAAATTTAGGATGATAGAGCTGGAGGAATTAACCTTCATCTAATCTAATTAAGCTAGATGCTTACCTTAAGGCAACAGC
* L E N R N H D * L F C S S P F L

← IRNA Ile

10210 10220 10230 10240 10250 10260 10270 10280 10290 10300 10310 10320
ACTATCCAAACACCTAAAGCCACAAACAAAACACAGAAAGACACACTAAGAGGCGATACTCTTCCAGTAAAGTACATTAATCTATATGCAATATGCAAAACCCGGATAGAGCTCC
S D L V G L L V V F F V L F S L S L F L Y K K W T L Y N L Q D Y R P R P Y S A R
10330 10340 10350 10360 10370 10380 10390 10400 10410 10420 10430 10440
AACCTCAAAAATAAAGAACCAAACTCCAGCTTAAAGCTACTACTAAAGACTTAACGGAAAGATCCCCCAATGGCGAAGACCCCAAGAAAGAGAGACATAGATACCAT
V M L F L L V L V G A K L S E V V L V N V P F F E C H M F S S G G L F L L V S L L N
10450 10460 10470 10480 10490 10500 10510 10520 10530 10540 10550 10560
GATAAATAATATAGGCTGCTGCTGATAAAGAACTAGCAAAAGGGCCACAGCATACTACTTGTAACTGACACAACTTACAGCTCCCTCAGTAAAGCTCAAAAGGGCCCG
I F I I N A Y E A I F F H A F P G G A Y E V N Y G S V I E S E E T G L D F P P R

10570 10580 10590 10600 10610 10620 10630 10640 10650 10660 10670 10680
ATTCTCTCTCCCAAGTGAAGAACAACTATATTCGACAAAGGAAACCACTAAACATTAACCAAGCAACTTGAAGACTCTCAATACCTTACATTAACCCCCAGCAACAG
M T E A L T S V F W I A F L P F C S F M L W S R S Q S S E I V E V N F G G A F L
10690 10700 10710 10720 10730 10740 10750 10760 10770 10780 10790 10800
CACCAACCTTAAAGGATCAAAACCGAGCTTACTGCTAGAAAGAGCTGAGCTACCGCTCGGACCCCTTAAGAAAGATAATAGGAATTAGAGAACCAACCAACCCCAACAG
V V G L L L I L G L S I E Y S V T O A V A R V A G L F S Y S N S N S W G S G L L S
10810 10820 10830 10840 10850 10860 10870 10880 10890 10900 10910 10920
GTATACAGATACCTAGATAAACCTATATTAAGCTCAGAGATATTTACTCTACCAAGATACTTCACTGGAATTACCGACCAAGCAGGACAGAGATTTCCCAAGAAAGAAAGAG
Y V S L S L S I L M L S L N V S V V S Y K V P I V S L V L A I G L F L V P

10930 10940 10950 10960 10970 10980 10990 11000 11010 11020 11030 11040
AGAAGAAGAAATAAAGGAGATGCTTTAGAGGCTTTAAAGTCTCTTAAATTAATAGCTTAAACCATCCCAATAGGCTGATATATCCAAAGGCCCAAGATATTGGACCTCT
S F E F L Y P S A R S P K L E K I L L K F G D A I P Q L L G F P G V I M P G K
11050 11060 11070 11080 11090 11100 11110 11120 11130 11140 11150 11160
ACGAAAGCTGATATAACCCAGAACCTTCTGCTCACTAGGGTCACTAAGCAACCCCAACAAACAGGACTATAAAGCAATAGATTTAACAAAACAACAACTCACTCACTCA
R L Q H Y G L V K R E V L T L L A V A L L V P V I F L I S N V F F V F W D V C

← IRNA Leu(UUR)
11170 11180 11190 11200 11210 11220 11230 11240 11250 11260 11270 11280
AAGAGAGTTGAACTCTCGATAAAGAACTTAAAGCTTCTGCAATACCACTTGGTATTTCAAGCAATCATATATATAGTATCTTATACGATTTGATTTGGTGTITTAGCA

← IRNA Gly ← IRNA Tyr
11290 11300 11310 11320 11330 11340 11350 11360 11370 11380 11390 11400
AAACCTTCTGGTGGAGGCGAATGCTGCTATACTCACTGAATTTGACCGAAGAGACTTGAACCTTCTCACTAGATTTAGAAATCTATCCGATCCCTATGCCCAACCAATAC
X

IRNA Asp
11410 11420 11430 11440 11450 11460 11470 11480 11490 11500 11510 11520
ACGAAGCTAGTAAAAAATACCCGAGATTTGACAGCTGGAAATATCTGTAAACCCCTTGATGCTTTTCACAAATAAAGGGAGGATTTATCTTCCATTTTGGAGTATGAC

← IRNA Met ← IRNA Val
11530 11540 11550 11560 11570 11580 11590 11600 11610 11620 11630 11640
CCAAAAGCTTTTATAGCTTACTTTACCTTTCCTGGGGAGCTTGAATGTAAGCAATCTCTTTACACAGAGCTGATATTTGCAAAATCAAAATTTGCTTTTCTGTTTTCAGCGAGAT

← IRNA Cys ← IRNA Trp
11650 11660 11670 11680 11690 11700 11710 11720 11730 11740 11750 11760
TACCCGCATCTATAGTTTGCATCTCAAAATGTAATACACACAAACCCAGAGGTAAGAAAATTTCCGATCTTATTTAAAGCTTTGAAAGGTTTATGTTAGATTAACCTAAACCC

IRNA Ala
11770 11780 11790 11800 11810 11820 11830 11840 11850 11860 11870 11880
TTCTTGTGACTTATTTAAAGATAAACAAATGATTTGCAATCAGTAACCAAGTTTAACTCTCGAGTCCAGTCAAGAAAGGATTTGACCTTGAATCTGATCTGATCTAAATC

← IRNA Leu(CUN) ← IRNA Asn ← IRNA Gln
11890 11900 11910 11920 11930 11940 11950 11960 11970 11980 11990 12000
TACTCGATTAACCACTTTCGACCCAGCTTGAGTTGACAGATTTGATCTGTTATCTCTGTTTAAAGGCAATTTGCTTCCATTTAGGCTACACCAACATGAAAGGATGATTAAGCC

← IRNA Pro

12010 12020 12030 12040 12050 12060 12070 12080 12090 12100 12110 12120
TAAAGCAGAGATTTGGATTCTTAAATATAGTACCTACTTACTTCTTCTAAGCAGAAAAGCAGACTTATAGTCTACACCAAGAGCTCCCAAGGATCATCTTCCATCAATCAATTTT

CO I →

12130 12140 12150 12160 12170 12180 12190 12200 12210 12220 12230 12240
CTGTATACATATATATACAAATATAAACCTACCACTAACAGATGATCTTTCCACCAACCAAGACATGGACCTTATCTTATCTTTGGGCGCTGACCGGATATGCC
M Q L S R W F F S T N H K D I G T L Y L I F G A W A G H A

12250 12260 12270 12280 12290 12300 12310 12320 12330 12340 12350 12360
GGACAGCAAGAGCGTAAATACAGACAGACTAGCCCACTGGATCCTATGCAAGCAGCAAAATACAAGATATATGCTACACGCCCTAGTCTGATCTCTTATG
G T A M S V I I R T E L A Q P G S L L Q D D D Q I Y N V I T A H A L H I G F F M

12370 12380 12390 12400 12410 12420 12430 12440 12450 12460 12470 12480
GTGATGCCAATCATATGGGAGTCTGGCACTGACTAATACCTETAATGATCGAGCTCCAGACATGSCCTCCCCCAATGAAATATGATGATCTGACTAATCCCTCTCTT
V H P I H I G G F T G K W M L I P L M I G A P D N A F P R N H M S F W L I P Y S F

12490 12500 12510 12520 12530 12540 12550 12560 12570 12580 12590 12600
CTCTCTTATAGCTCTCGCAGGTGTGAAAGAGGTGTGGCAAGGATGACACTTACCTCTCTTGTCTAGGGGATAGCCGATGCGGGCATGAGTAGCTGGCGATATGCC
L L L L L A S A G V E S G A G T G W T I Y P P L S S G L A H A G G S V D L A I F S

12610 12620 12630 12640 12650 12660 12670 12680 12690 12700 12710 12720
CTTCCAGTGGGCGAGCTCCCTCAATCTTCTGTCTAAATAATATACCACTGTTATAAATGGCAACCCCTGGATTCATTTGACCCAGTCTCCCTATTGTCTGATCAATTT
L H L A G A S S I L A S I N F I T T V I N H R T P G I S F D R L P L F V M S V F

12730 12740 12750 12760 12770 12780 12790 12800 12810 12820 12830 12840
GTAAAGCTTCCCTCTTATCCCTACAGTCTGCAAGTGCATACGATGCTCTTACAGACGAAAGTACACCTTTTTCAGCCCGCAGAGGGGGAGACCCATT
V T A F L L L L L V L A G A E T H L L T D A K V N T T F F D P A G G G D C P I

12850 12860 12870 12880 12890 12900 12910 12920 12930 12940 12950 12960
CTATTCAACACTTATGCTCTCGCTACCAGAACTCATTTATCTTCCGGATTGCGAATGATTCCTCATGATATGCGACTACCGGAAAGAGAACCTTGGT
L F Q H L F W F F G H P E V Y I L I L P G F G M I S H V I A H Y A G K S E P F G

12970 12980 12990 13000 13010 13020 13030 13040 13050 13060 13070 13080
TATCTCGAATGTACCGCTAGTCTGTATGATCTAGGATTTAGCTTGGCCCACTATGTTACAGTGTGATGATACACCGGATCACTTTACAGCCGGACCC
Y L G H V Y A I V S I G I L G F L V W A H H M F T V G M D V D T R A Y F T A A T

13090 13100 13110 13120 13130 13140 13150 13160 13170 13180 13190 13200
ATGATATAGCACTACCACTGTATAAAGTATTTACTGATGACACCCCTACAGGGGAGAACCTTCATGATGGACACACTTTATATGGCCTTGGATCTTATCTTATACC
H I I A V P T G I K V S M W A T L Q G S N L R W D T F L L W A L G G V F L F T

13210 13220 13230 13240 13250 13260 13270 13280 13290 13300 13310 13320
ATGGAGGGCTAACCGGATTTCTGCAAAATCACTCGAGCTACCTCCACGACACTACTAGTGTGCCACTTTTCAATGATATATCAATGGGACATTTGCCATA
I G G L T G V I L A N S S I D V I L H D T Y Y V V A H F H Y V L S M G A V F A I

13330 13340 13350 13360 13370 13380 13390 13400 13410 13420 13430 13440
TTCCAGGATTCACCCCTGGTCCCACTATCTCAGGGTACCGCTACACCCCTATGAGGATACACTTACTCTCATGTTATGGTGTACCTACCTCTTCCACAGAT
F A G E T H H E F F L S G V G L H P L M G N I H F T L M F I G V N L T F P D Q H

13450 13460 13470 13480 13490 13500 13510 13520 13530 13540 13550 13560
TTCTAGGATACCGGATGCCACAGCACTCCGACTCCCGACCGCTACACACTGTAACACAGTCTCTCCATTTGGAGAAACATATCTTATCGCAAGTGTGATTTCTTA
L L G L A G M P R R Y S D Y P D A Y T L W N T V S S I S G A S T I S L I A T L V Y L

13570 13580 13590 13600 13610 13620 13630 13640 13650 13660 13670 13680
TTATCTTATGAGAACTTACTCCAGGCACTGCTCAACCTGATTTCCACTCTCTCTGAAAGAAATTCAGTGTTCCTCCCTCCACCAAGATTTGATGAACT
F I L W E A F T S O R T A L O P E F S T S S L R W Q Y S S F P P S H H T F D E I

IRNA Arg →

ND4L →

13690 13700 13710 13720 13730 13740 13750 13760 13770 13780 13790 13800
CCTTCACAGTTTACTAATAAGTAAAGGATGATTTAAACAAAACCTTAATTTCCGGCTAAAAAATTTGGTAAACACCCCGAACTCTTATACATCACTTCATCAFTA
P S S V Y L I K A

13810 13820 13830 13840 13850 13860 13870 13880 13890 13900 13910 13920
CCATACAATTTTACTTAGTCTCTGGAACTGTATAAAGCACTACACTCTTCAATCTACTAGCTACCTCAATATATCTGTTTGTGAGGCTAACCGATAGG
I T I P Y I L G L L G I L I N R L H F L S I L L C L A G L L I S L F L S L T V W W

13930 13940 13950 13960 13970 13980 13990 14000 14010 14020 14030 14040
CCATAAAGCCAAACACCTTCTATTGCAAAATAACTTACTCTCTACTCTATACCTGAGACTAGCCAGGCTTTCCCTCATGGTAGCCCTTCCCGACACACTCCS
I N A N T T F L L T N N L I L L T L S A C E A S A G L S L M V A L S R T H N S D

CO II →

14050 14060 14070 14080 14090 14100 14110 14120 14130 14140 14150 14160
ACCTTGTCTCAACATAAATCTCCAGCAATAAATGGCAACTGAAACCAGTACGAC TACAAGATGATCTCTCCCACTAAGGAAGCACTATATCTTACAGATATACTTAA
L V S T I N I L Q Q A H A N W T Q L G L Q D A S S P L H E E L I Y F H D Y T L I

14170 14180 14190 14200 14210 14220 14230 14240 14250 14260 14270 14280
TTATCTCAACCTAAATCAATATAGTATTTAGCGACTCTGCTACTATCTCTCCCAAACTAACCGATTTTCTGGAAAGCAAGGATGAAGACTTCCGCAACATCAT
L I L L T I L I L V F Y G L A S L L F S S N T N R F F L G G Q L E L T V W T I F P

14290 14300 14310 14320 14330 14340 14350 14360 14370 14380 14390 14400
CTGGCTAATCAATATTTATGCCCTCCCTCTTACCACTATACTTATGGACAGGTAATAAATCCACTACTAGCAAGGCGATAGGACCACTGATCACTGAAGTAGC
A V I L I F I A L F S L Q L Y L L H D R V N N P Y L T I K A I G Q W Y I T B

14410 14420 14430 14440 14450 14460 14470 14480 14490 14500 14510 14520
AATATCCAGACTAGCGAGTATGATTCGACTCTCACTTCAACTGATTAACCTCGGAAACCTGAGCTTGAAGATGAAACCACTCACTTCCCTCCCAACCG
Y A D Y R E L E F D S Y H I P T S D L T S G N P R L L E V D N R L T L P A Q T P

14530 14540 14550 14560 14570 14580 14590 14600 14610 14620 14630 14640
CCATCCGATTTTATGCTCTCCCGAGCTCTACACTGATGAAATCTCTCTGGGATTAAGATGATGCCCTCCCGGAGCTTAAACCACTCAATCTTCAATTTCTGAT
I R V L V S S A D V F L S W A I P S L G I K N D A V P R L N Q V N F I S R C

14650 14660 14670 14680 14690 14700 14710 14720 14730 14740 14750 14760
GTGGCTATTACGCCAGTCTCAGAAATCTCGCGAGCCAAACAGTCTTATGCCCATAGTAAATGAATCACTTCCAGCTCGAACTGAGTACTAAATTTATTAACCT
G L F Y D Q C E I C G A N H S F M P I V I S E V F S T F E M W S N F I T B

IRNA Lys →

ATPase 8 →

14770 14780 14790 14800 14810 14820 14830 14840 14850 14860 14870 14880
AACTTTGATAGCTTAAGCAAGCAGCAACTTAAITGAAATAAAGTGAATCCACCCATCAAAAGTCCACCACTAAACTGATGATGATCAATTAATCTTCT
M P Q L N L A W L L L F F L

ATPase 6 →

14890 14900 14910 14920 14930 14940 14950 14960 14970 14980 14990 15000
TAGTGGACCTCTTACTAGTAAATTTATGAACTAGTAACTGAAACACTGAAACAGCAACACCCAGTACAGCAACCCCAACCTAAAGTAAAGCTGAACT
G M T S L L V I F I L L N T S I T N T A N I S T T S T N T P L L M N L L L

15010 15020 15030 15040 15050 15060 15070 15080 15090 15100 15110 15120
 AAAAAGATCTTTGGCCAATCTCCCGAGACCTCGTATTATTATTCTATGACCTTACAGCGCGTCTTCTTAATCTTAGTGGTATCTATTTCCTCAAGCATTAAATGACTCCCGCTC
 N S I F G Q F S P D L V L F I F M T L T A V F L N L S W L S I S N P S R W L P S

15130 15140 15150 15160 15170 15180 15190 15200 15210 15220 15230 15240
 CCGGGCAACCTCTCACTACTATCTTTTACCAAGAAGTACTAAAGATATTGTTCCACAACAACCCCAACAGCCCGCTTGAGTACAGCTTTTACCGGCATATTTACTATTAAT
 R A N L L I L S F Y Q E V L K I L F Q Q T N P N T A P M V S A F T A I F I L I F

15250 15260 15270 15280 15290 15300 15310 15320 15330 15340 15350 15360
 TTCATTAAGTCTTGGCTCTCTCCATCCGACTTCCCTCTCAAGTCACTATGATTAACAGACTCAATGGGGTCCCGCTTGAATGCTCCCAAAATGACTGGATTTACTTACGC
 S I N V L G L L P Y A T F T S T S H I S L T V S I G V P L W H S V N I L G F Y L A

15370 15380 15390 15400 15410 15420 15430 15440 15450 15460 15470 15480
 ATTCACAGTCGACTAGGACACTTAGTCCCAAGGAACCCCTCGTACTTATACCATTCATGGTAATATAGAAACAATCAGCCCTTTTGCACAACCATAGCATAGGACTACGACT
 F H S R L G H L V P Q G T P S Y L I P F H V I I E T I S L F A Q P I A L G L R L

15490 15500 15510 15520 15530 15540 15550 15560 15570 15580 15590 15600
 TGCAGCCAACTCAGCCCGGACACTTACTCATATTCCTATTATCTACCGCAATCTGAACCTGTCTAGCAGCCCGCTCCATAGCCAGAATACCTGTAAATATTTTCTCCTCTCTC
 A A N L T A G H L L I F L L S T A I W T L S S S P S I A S I T L L I F F L F L

CO III →

15610 15620 15630 15640 15650 15660 15670 15680 15690 15700 15710 15720
 CCTAGAAATAGCGGTAGCGTCCGATTCAGCCCTACTGATTCCAGCCCTAGTAACTCTTACCTATCCCAAACTTTAAAACTTATGACCCATACACCCCGTACCCTGTTGACAA
 L E I G V A C I Q A Y V F T A L V N F Y L S Q N L M T H Q H P Y H L V D Q

15730 15740 15750 15760 15770 15780 15790 15800 15810 15820 15830 15840
 AGCCCTGACCTTAACTGGGGCAATAGAGCTCTAATGATGACTCCGGTCTTATTCTATGATTTATACCAACAGAACCCCTCTTTAGCCGGACCATCTACTCTCTTAACC
 S P M P L T G A I S A L M M T S G L I L W F R T N S N R L L L A G T I L L L L T

15850 15860 15870 15880 15890 15900 15910 15920 15930 15940 15950 15960
 GTTATAATGATGACGAGATGTAATCCGGGAAGCAACTTCCAGGAAGACACCCCTCCCTGTAACACAGGATACGATACGGCATGATATTTTATAACATCCGAATAGCTTC
 V I N W R D V I R E A T F Q G S H T L P V N T G L R V G M I L F I T S E V C F

15970 15980 15990 16000 16010 16020 16030 16040 16050 16060 16070 16080
 TTTTTCCTTCTGAGGATCTTTATAGAAAGCTTACGCCCCACAGTGAATGGAGTATCCTGACCTCAACTGGGATAGTCCCAAACTTTCTCTAGTGCCTTACTAAC
 F F A L F W A F F H S S L A P T V E L G V S W P P T G I S P I N P F L V P L L N

16090 16100 16110 16120 16130 16140 16150 16160 16170 16180 16190 16200
 ACAGCGCTTACTCTCTCAGGAGTACAGTACTCGAGCACACAGGATACTAACCAAAATCGAAGCAATACAGGACTATCTCTACAGTAATCTGGAAATTAATCTTC
 T A V L L S S G V T V T W A H H S I L T Q N R T E A I Q G L F L T V I L G I Y F

16210 16220 16230 16240 16250 16260
 ACTGGCTTCAGCGCTGAGAATACTATGACTCCCGCTTACCATCCGACAGAGATAT
 T G L Q A W E Y V D S F P T I A D S V Y

ヒトデミトコンドリアゲノムの遺伝子構成

タンパク質 (bp)	tRNA (bp)	rRNA (bp)
ND1 981	tRNA ^{Leu} 73	16S rRNA 約1530
ND2 1065	tRNA ^{Ala} 72	12S rRNA 897
ND3 333	tRNA ^{Phe} 72	
ND4 1383	tRNA ^{Ile} 72	
ND4L 297	tRNA ^{Met} 72	
ND5 1926	tRNA ^{Asn} 72	
ND6 489	tRNA ^{Thr} 72	
CO I 1554	tRNA ^{Leu} _{CUU} 72	
CO II 688	tRNA ^{Leu} _{UUR} 72	
CO III 783	tRNA ^{Ser} _{UCN} 72	
ATPase6 693	tRNA ^{Gln} 71	
ATPase8 165	tRNA ^{Val} 71	
Cyt b 1138	tRNA ^{Asp} 71	
	tRNA ^{Pro} 70	
	tRNA ^{Arg} 70	
	tRNA ^{Trp} 70	
	tRNA ^{Tyr} 70	(合計) 約15460bp
	tRNA ^{Glu} 69	Overlap 19bp
	tRNA ^{Gly} 69	
	tRNA ^{His} 69	
	tRNA ^{Ser} _{AGN} 68	$\frac{15460}{16260} = 95.1\%$
	tRNA ^{Cys} 67	

表3

Mitochondrial Genomic Sizes for Various Organisms

Organisms	Genomic sizes (bp)
Vertebrates	
Human	16,569
Mouse	16,295
Frog (<i>Xenopus laevis</i>)	17,553
Invertebrates	
Sea Urchin (<i>S.purpuratus</i>)	15,650
Starfish (<i>A.pectinifera</i>)	16,260
Fruit Fly (<i>D.yakuba</i>)	16,019
Protozoa	
Paramecium	40,469
Higher Plants	
Maize	~484,000
Wheat	~210,000
Filamentous Fungi	
<i>A.nidulans</i>	~32,000
<i>N.crassa</i>	~62,000
Yeast	
<i>S.cerevisiae</i>	68,~76,000
<i>S.pombe</i>	~19,000

表4

4. 遺伝子配置

図5にヒトデミトコンドリアゲノムの遺伝子配置を示した。3. で述べたようにヒトデの遺伝子構成は他の高等動物のものとは一致している。しかしその遺伝子配置は脊椎動物、節足動物のものとはかなり異なっていた。図6にヒトデ、ウニ、脊椎動物（ヒト、ウシ、マウス、ラット、アフリカツメガエル）、シヨウジョウバエ、線虫の遺伝子配置を示す。脊椎動物の遺伝子配置は比較的保存されているがウシガエル、トリ、トカゲなどで一部の遺伝子の転座が報告されている。

ヒトデの遺伝子配置が脊椎動物に比べて特に異なっている点は以下の通りである。

(1) 13個のtRNA遺伝子がクラスターを形成しており脊椎動物のように遺伝子間に分散していない。

(2) 脊椎動物ではND6遺伝子が唯一、他の12個のタンパク質のコードされている鎖(Light Strand; L鎖; Sense Strand)の相補鎖(Heavy Strand; H鎖)にコードされているが、ヒトデではND6遺伝子以外にND1遺伝子, ND2遺伝子が相補鎖にコードされている。

(3) 12SrRNA遺伝子と16SrRNA遺伝子が互いに相補鎖上に存在している。

などの点である。

一方、ヒトデの遺伝子配置を同じキョク皮動物のウニと比較すると良く保存されている。ヒトデとウニの唯一の相違点は16SrRNA遺伝子から13個のtRNA遺伝子のクラスターまでの約4.6kbpの領域が反転している点である。この反転領域およびそれを除く約11.6kbpの領域における遺伝子配置はヒトデとウニで完全に一致していた。遺伝子以外では特にヒトデのND4遺伝子とtRNA His遺伝子間に存在する140bpのスペーサーがウニには存在しない。

キョク皮動物の遺伝子配置の特徴を脊椎動物と比較すると、tRNA遺伝子の

クラスターの存在が最も特徴的である。8章で詳しく考察するがゲノムの縮小化した高等動物ミトコンドリアにおいてはtRNAは単に蛋白質合成においてその本来的な役割を果たすだけではなく、ポリストロニックに転写されたRNA前駆体において、mRNAやrRNAのプロセシングの目印としての役割を担っている^{39,50}。従ってtRNA遺伝子が蛋白質遺伝子やrRNA遺伝子を区切るように配置していることはミトコンドリアの各RNAのプロセシングにおいて重要な意味を持っている。

脊椎動物のミトコンドリアゲノムにおいては13個の蛋白質遺伝子のうち6個がtRNA遺伝子によって両側を区切られている。また6個がどちらか一方を区切られている。また2つのrRNAについてのその両側にtRNA遺伝子が存在する。

それに対しキョク皮動物では多くのtRNA遺伝子がクラスターとなっているため脊椎動物に比べ、他の遺伝子を良く区切っていない。例えばヒトデでは両側にもtRNA遺伝子が存在する蛋白質遺伝子はND1遺伝子とCOI遺伝子の2個だけであるし、ウニではわずかにND1遺伝子のみである。さらに16SrRNA遺伝子もその両側にtRNA遺伝子が存在しない。

蛋白質遺伝子においては、その3'端は一般に終始コドンの存在で認識が可能であるため、終始コドンがあれば隣接するtRNA遺伝子の存在は必ずしも必要ないが、脊椎動物などの例では蛋白質遺伝子の3'端に終始コドンがなく、tRNA遺伝子の存在が必要とされるものもある。けれども特に問題となるのはその5'端の認識であろう。

脊椎動物においては蛋白質遺伝子の5'隣接部分にもtRNA遺伝子の存在しないものは3個であり、ショウジョウバエでは3個、回虫では4個であるが、ヒトデでは6個、ウニでは7個となっている。以上の議論は隣接するtRNA遺伝子の向き(コードされている鎖)を考慮しておらず、これを考えると脊椎動物は5

個、シヨジョウバエが6個、回虫は4個、ヒトデが8個、ウニが8個となる。

脊椎動物などのRNA転写物の解析の結果からすると、おそらくアンチセンスのtRNAもプロセシングの目印になっていることが示唆されるため、考慮すべき結果は前者である。これらのうち例えばATPase 6遺伝子とCOIII遺伝子についてであるが、ATPase 8遺伝子、COIII遺伝子の位置関係は脊椎動物、キョク皮動物、シヨウジョウバエで保存されている。脊椎動物ではATPase 6とATPase 8のmRNAは連結したまま存在しており、その状態でmRNAとして働いているのであらうと示唆されている。おそらくこの部分については広い種間で保存されたメカニズムが働いているのであらう。それ以外の部分についてはキョク皮動物に特有のRNAのプロセシングや翻訳制御を考慮する必要があり後章で詳しくのべる。さらに16S rRNAについてもキョク皮動物はその両末端にtRNA遺伝子がコードされていないため、そのプロセシングについての考察が必要である。特にrRNAは蛋白質遺伝子と異なり3'端についてもそれを特定する事ができない。特にヒトデの場合その3'端隣接部位が非翻訳領域であるため、DNAのシークエンスからは3'端を特定できない。従って16S rRNAのマッピングを行ったので10章で報告する。

ヒトデとウニの反転領域を除くCOI遺伝子からtRNA Thr遺伝子までの約12Kbpの領域については蛋白質遺伝子に限るとキョク皮動物と脊椎動物で良く保存されている。唯一の相違点はND4L遺伝子が転座している点である。またND1遺伝子とND2遺伝子が同一方向に並んで存在している点もキョク皮動物と脊椎動物で保存されている。キョク皮動物、脊椎動物、シヨウジョウバエにおいて、蛋白質遺伝子の相対的な位置関係はtRNA遺伝子に比べ比較的良く保存されている。それらに対してtRNA遺伝子の移動は顕著である。従って遺伝子配置の変遷のメカニズムは基本的にtRNA遺伝子と蛋白質遺伝子及びrRNA遺伝子とは異なっているのであらう。

脊椎動物とショウジョウバエの蛋白質とrRNAの遺伝子配置の位置関係の相違は3カ所の遺伝子反転によって説明できる¹⁰。またヒトデとウニの相違点はtRNA遺伝子まで含めて一回の遺伝子反転だけで生じている。それに対してキョク皮動物と脊椎動物の遺伝子配置の相違は転座を必要とする。遺伝子反転はミトコンドリアゲノムに限らずしばしば見られる現象である。その生成のメカニズムはDNA中に相同な配列が向かいあって存在し、それらが部位特異的に交差することで説明できる。(図7) それに対して転座のメカニズムは複雑である。転座は複数回の遺伝子反転による場合と、遺伝子の挿入と欠落による場合とが考えられる。

遺伝子の欠落としては複製過程において相同部位のslippage⁵⁹による場合と遺伝子反転と逆に同一方向の相同部位における部位特異的交差によって欠落する場合とが考えられ、それらは一挙に長い領域が欠落する。また偽遺伝子化した断片が徐々にslippageあるいは化学的、物理的損傷によって欠落する場合もあるであろう。

遺伝子の挿入についてはカビや植物のミトコンドリアの例が参考になるかもしれない。カビや植物のミトコンドリアゲノム中には"mobile element"が存在すると報告されている^{60,61}。さらにtRNA遺伝子はその"mobile element"の認識に何等かの関連があると報告されている。動物のミトコンドリアについては下等なものまで含めてこのような"mobile element"は見つかっていない。しかしかつて動物ミトコンドリアゲノムにもmobile elementが存在して遺伝子の転座を行い、その後ゲノムの縮小に伴いゲノム中から欠落した可能性があるとの報告もある⁶²。

他の考えられるメカニズムとしてはtRNAがDNA複製のプライマーとして働きそのままゲノム中に組み込まれたと言うものである²¹。(図8) 例えば脊椎動物においてそのD-loop中にtRNA-likeな構造が見られると言う

報告がある^{63,64}。また脊椎動物ミトコンドリアゲノム中の最も大きいクラスターはL鎖の複製開始部位に存在していることも示唆的である。さらにウシガエルのミトコンドリアゲノムにおいてはtRNA Leu遺伝子がD-loop内に転座しており、またtRNA Thr遺伝子およびtRNA Pro遺伝子がD-loop中で転座している¹⁹。さらにトカゲのゲノムではD-loop中あるいはD-loopの両側に0.8 kbp~8.0 kbpの何重もの遺伝子重複がありいくつかの蛋白質遺伝子あるいはrRNA遺伝子を含んでいる。そしてそれぞれの重複の境界にはたいていtRNA遺伝子が存在している。トリにおいてもND6遺伝子の転座はD-loopに隣接しておりtRNA Glu遺伝子をとまっている。従ってtRNAは単にプライマーとして自ら組み込まれるだけでなく、それに隣接してポリシストロニックに転写される遺伝子をも組み込む先導になっているのかもしれない。

ではキョク皮動物にたいする脊椎動物、ショウジョウバエの遺伝子配置の差異について、これらの可能性を検討してみたい。形態学的な研究によりこの3者においてまずショウジョウバエが分岐し、その後には脊椎動物とキョク皮動物が分岐したと考えられている。ショウジョウバエと脊椎動物を比べると12SrRNA遺伝子、16SrRNA遺伝子、ND1遺伝子およびD-loopの位置関係が保存されている。従ってキョク皮動物と脊椎動物のこの部分での遺伝子配置の差異はキョク皮動物側で生じた可能性が高い。またND4L遺伝子、ND4遺伝子、ND5遺伝子の位置関係およびCOI遺伝子とCOII遺伝子の位置関係もショウジョウバエと脊椎動物で保存されている。従ってND4L遺伝子の転座もキョク皮動物側の現象であるだろう。さらにヒトデとウニのどちらが先祖型かと言う問題もある。まずこの点についてはtRNAのプライマー化という仮説からするとウニのほうが先祖型とするほうが都合が良い。つまりウニでは15個のtRNAがD-loopの両側に集中しているが、各々のtRNAがプライマーとなり、

つきつぎと組み込まれていき、その後ヒトデとウニの遺伝子反転が起こり現在に至ったのであろうとするものである。さらに12S、16Sの両rRNA遺伝子およびND1、ND2遺伝子はこのD-loopを囲んで存在しており、トカゲやトリのようにtRNAを伴いD-loopの回りに転座したと言う仮説と同様のメカニズムが働いたのかもしれない。この場合においては転座した遺伝子は重複して存在していたであろう。その後、転座した遺伝子の方が遺伝情報を保持し、元の遺伝子は消失したものと考えられる。特にキョク皮動物においてはこれらのシナリオのうち、前半の部分を強く支持するtRNA Leuの重複、消失の証拠があり、これについては7章でのべる。ND4L遺伝子の転座の場合、この遺伝子はD-loopから離れており状況が異なる。これを遺伝子反転で説明しようとすれば延べ3回の反転によらなくては説明できない。それはまず脊椎動物の遺伝子配置で見てCOII遺伝子からND4L遺伝子までが反転し、その後ND4L遺伝子とCOII遺伝子からND3遺伝子がそれぞれ反転するというものである。しかし、このND4L遺伝子の転座以外は良く保存されていることからすると、ND4L遺伝子の挿入されたとする方が適当と思われる。このND4L遺伝子の転座はtRNA Arg遺伝子を伴っており、“mobile element”などによる転座かもしれない。あるいはCOII遺伝子とCOIII遺伝子間にかつて複製関連の配列が存在したか、またはそれと相同性の高い配列があったのかもしれない。

次に回虫²²の遺伝子配置についてはキョク皮動物、脊椎動物、ショウジョウバエの遺伝子配置との関連性はほとんどない。回虫についてはおそらくゲノム内ゲノム間でたびたび交差が起こり大幅な再編成があったのだろう。

これらの遺伝子配置の変化の歴史を現在において調べる事は難しい。例えばヒトデとウニの分岐は5億年以上さかのぼれる。その後この両者は遺伝子反転を除いてよく保存されていた事になる。また脊椎動物においても両性類、鳥類、哺乳

類などが分岐してから2～3億年になるが一部の例外を除いて良く保存されている。これらの遺伝子配置の変化はおもに数億年以上前において生じたことでその後はかなり保存されている。そしてその間にゲノムの縮小などにより重複した遺伝子や遺伝子反転などに必要な相同配列などが消失したのであろう。この遺伝子配置の変化をもっと詳しく調べるには数多くの生物種をサーチして様々な中間段階にあるものを見つけ出し、比較する事が必要となるであろう。

遺伝子配置が変化するメカニズムについて実験的なアプローチも困難である。それはこれらの遺伝子置換はおそらく生体のシステムに組み込まれた合目的な過程というよりも、数億年の時間のなかで数回生じた偶然によるものであると考えられるからである。これを分子的に説明するのは現在実際に遺伝子編成の変化をしている系からの推論に頼らざるを得ない。

最後にrRNA等が遺伝子に組み込まれるという仮定にはRNAを鋳型にDNAが複製されるという過程が必要となる。一つの考えとしては細胞中にもぐりこんだレトロウイルスの逆転写酵素がミトコンドリア中に入りこみ、逆転写をおこなったというものである。もう一つの考えはミトコンドリアのDNAポリメラーゼに逆転写の活性があるというものである。DNAポリメラーゼがRNAを鋳型にしてDNAの伸長をおこなうことは一般に報告されていない。しかし最近Taq DNAポリメラーゼによってRNAを鋳型にしてPCRを行い目的のプロダクトを得る事に成功した例が報告された。このことはRNAを鋳型にDNAポリメラーゼが伸長反応をした事を示している。この例からすると、従来DNAポリメラーゼの逆転写活性が報告されないのは、その活性が低いためであり、PCRのような非常に高感度の検出法により、初めてその活性を検出できたのかもしれない。このようなケースがミトコンドリアのDNAポリメラーゼに当てはまるかどうかは今後の課題である。

イトマキヒトデミトコンドリアゲノムの
遺伝子配置図

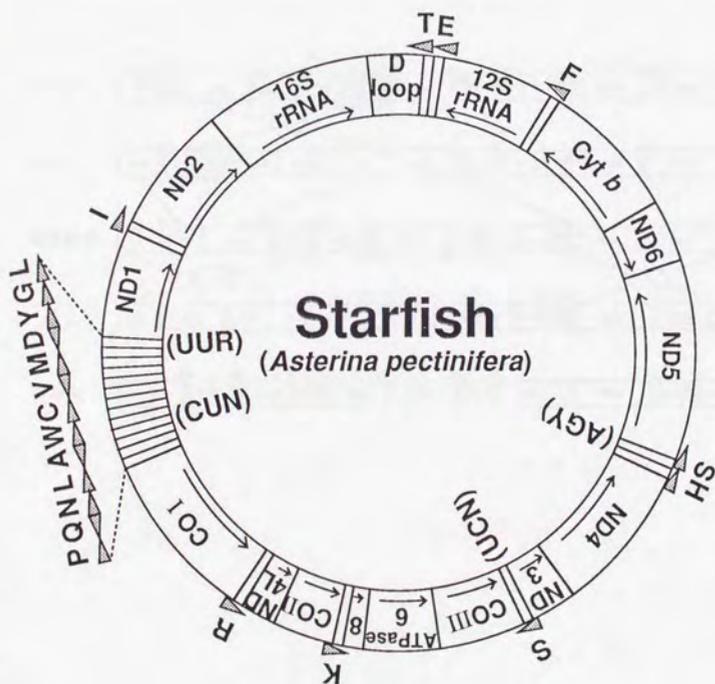


図5 ヒトデミトコンドリアゲノムの遺伝子配置図

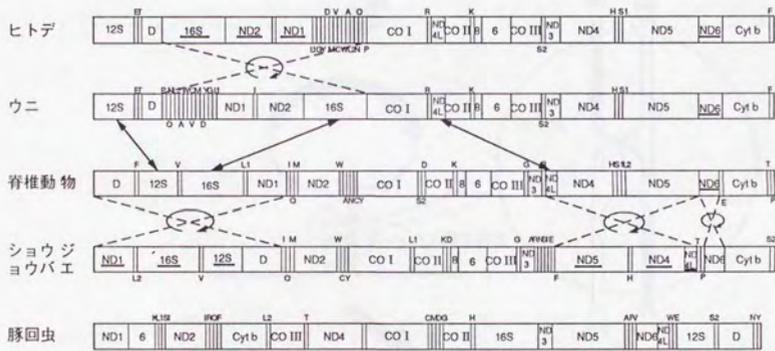


図6 種々の生物における遺伝子配置の比較

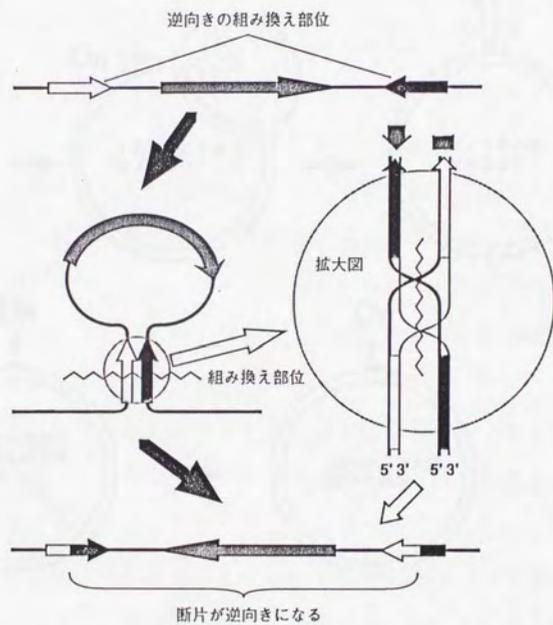


図7 遺伝子反転の概念図

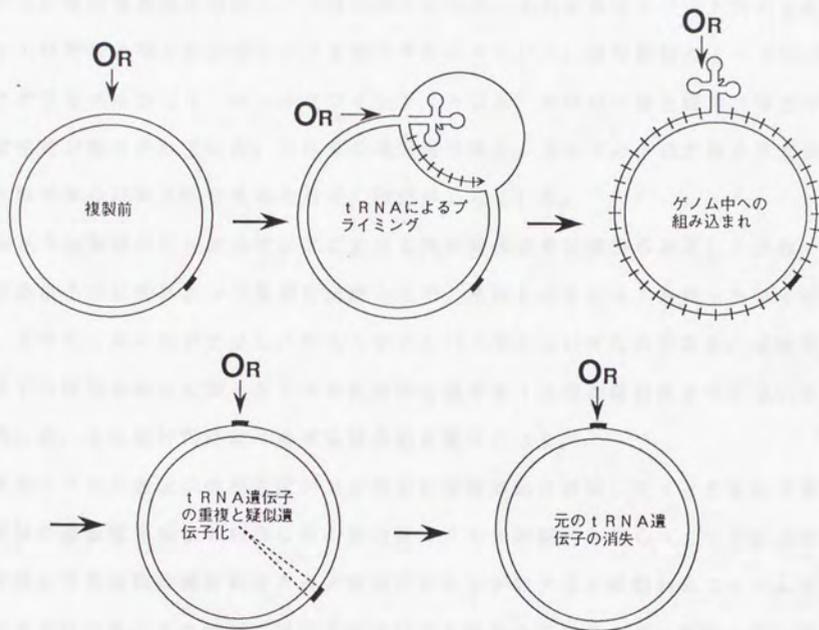


図8 tRNA遺伝子移動の模式図

5. 遺伝暗号

序で述べたようにミトコンドリアの蛋白質合成系にはいくつかの異常な遺伝暗号が見つかっている。これらの異常暗号のうちTGA—トリプトファンは酵母、下等動物、高等動物などの広い生物種でよく保存されている。(表6)この点についてはヒトデミトコンドリアにおいても例外ではなかった。TGAコドンは全ゲノムの蛋白質遺伝子中のべ78個出現しており、それが示すトリプトファンのサイトは脊椎動物との比較において良く保存されていた。脊椎動物のミトコンドリアではそれ以外にATA—イソロイシン、AGA、AGG—停止暗号、などの異常暗号が報告されている。これらの異常暗号はウシやヒトの一部の蛋白質でアミノ酸分析の結果と照合されており、確認がとれている。

キョク皮動物のミトコンドリアにおける異常遺伝暗号は姫野らが本ヒトデの一部の遺伝子の解析によって最初に指摘した²⁹。それによるとATAがイソロイシン、AGA、AGGがセリン、AAAがアスパラギンというものである。本研究ではその報告をもとに彼らが行った統計的な調査を13個の蛋白質全てについて追試した。その統計的処理の簡単な概念図を図9に示す。

まずヒトデの遺伝子の塩基配列およびそれを暫定的に翻訳した(とりあえず脊椎動物の遺伝暗号などを利用した)蛋白質のアミノ酸配列をならべ、つぎに比較の対照とする生物の蛋白質のアミノ酸配列をヒトデのアミノ酸配列とフレームを合わせて比べる。この処理は遺伝子解析ソフトのMAXMHを使って行った。次にそれぞれのコドンについて、そのコドンのでてくるアミノ酸のサイトが比較の対照となっている生物では何のアミノ酸になっているかを調べ、各コドンごとに集計する。その結果を表5に示す。

表5によるとヒトデの蛋白質遺伝子にはATAが延べ178個出現したが、例えばヒトのミトコンドリアと比べたところ、その178個のサイトの内70箇所にはイソロイシンが出現している。それに対してメチオニンは18個しか出現して

いない。つまりヒトでの A T A のサイトにおける相同性（アミノ酸の保存性）はイソロイシンとした場合約 40% であるが、メチオニンとした場合わずかに 10% となる。メチオニンがランダムに存在したとしても約 5% の相同性が期待できること、13 個の蛋白質における全体の相同性は約 50% であること、イソロイシンのコドンからメチオニンのコドンへの移行は 1 b p 置換で済むことなどを考慮するとやはり A T A - イソロイシンと考える方が適当である。同様に他の脊椎動物やショウジョウバエとの比較においても A T A がイソロイシンのコドンであることを支持する結果が得られている。脊椎動物間ではアミノ酸配列の相同性が高いため、ヒトで得られる結果に準ずる結果が他の脊椎動物においても得られることは予想されることであるが、特にショウジョウバエとの比較においても同様の結果が得られることの意味は大きい。A T A - イソロイシンは普遍暗号と同一であるが、他の酵母などの下等生物からショウジョウバエや脊椎動物にいたるまでミトコンドリアの系では全て、A T A - メチオニンであることからすると、ミトコンドリアの中で例外的であり、おそらく祖先生物で A T A - メチオニンであったものが、再び普遍暗号側に変化したものであろう。

同様の議論で A A A についてもリシンのコドンと考えるよりもアスパラギンのコドンとする方が適当であると考えられる。この場合についても A A A - アスパラギンとした場合このサイトにおける相同性は約 40% であるが、リシンとした場合は約 5% である。一方、対照として調べた A A G についてはリシンとした場合が約 75% であるのに対し、アスパラギンとすると約 4% になり、A A A の結果と対照的な結果である。この結果はこれらの統計的処理が十分に有意なものであることを示している。

A A G におけるリシンが 75% とかなり全体の相同性から比べて高いのは、従来機能的に重要なサイトのリシンの情報が A A A と A A G の 2 個のコドンにシェアされていたものが、遺伝暗号の変化により A A G 一個によって担われるため

あろう。さらに9章でふれるBase-bias⁸⁴により、さほど重要でないリシンのサイトにおいて、AAGコドンがAAA、AACなどに置換されたことが、相同性の高さによつて拍車をかけたものと考えられる。AAAがアスパラギンであるケースはミトコンドリアにおいても、ミトコンドリア以外でも初めてのケースである。

次にヒトデでは表5に示すようにORF中にAGAが54回、AGGが19回出現しており、何かの暗号であることは確かである。これらのコドンは普通暗号や酵母などの下等生物のミトコンドリアにおいてはアルギニンのコドンであるが、脊椎動物のミトコンドリアゲノムにおいては終始コドンとされる。またショウジョウバエのミトコンドリアゲノムではAGAはセリンのコドンであり、AGGはゲノム中に存在せずナンセンスコドンとなっている。本研究ではヒトデにおいて、AGA、AGGコドンはセリンのコドンと判断した訳であるが、ATA、AAAに比べるとこのサイトの相同性は低く30%前後であった。

その後、荒木、スミスらによって報告された他種のヒトデにおいてもこの結果を支持するデータが得られている。さらにウニにおいて得られた結果も、これらの遺伝暗号の変化を支持するものであった。従つてこの遺伝暗号の変化はキョク皮動物に共通しており、5億年以上前に生じた変化であることが予想された。

また最近ヘン形動物⁸⁵ミトコンドリアにおいてAAAがアスパラギンのコドンとするキョク皮動物と同じ遺伝暗号の変化を持つものが報告されている。形態学的な進化の系統からするとこれらがキョク皮動物と同一の祖先生物から派生したものと考えるににくい。おそらく互いに独立して同じ遺伝暗号の変化をしたのであろう。表6には様々な生物の異常暗号と比較しながら、ヒトデで得られた結論をまとめた。

最後に、遺伝暗号の変化について考察するには、それに対応するtRNA構造の変化を考慮する必要がある。このことについては7章で考察する。

— ATGAGCGTAAATAACGAACAGAACTAGCGCAACCTGGATCC — ...

M	S	V	?	?	R	T	E	L	A	Q	P	G	S
L	S	L	L	I	R	A	E	L	G	Q	P	G	N
L	S	L	L	I	R	A	E	L	G	Q	P	G	T
L	S	I	L	I	R	A	E	L	G	Q	P	G	A
L	S	L	L	I	R	A	E	L	S	Q	P	G	T
L	S	I	L	I	R	A	E	L	G	H	P	G	A

... — TTACTGCAAGACGACCAATAATACAAAGTATAGTTACTGCA —

L	L	Q	D	D	Q	?	Y	?	V	?	V	T	A
L	L	G	N	D	H	I	Y	N	V	I	V	T	A
L	L	G	D	D	Q	I	Y	N	V	V	V	T	A
L	L	G	D	D	Q	Y	I	N	V	I	V	T	A
L	L	G	D	D	Q	I	Y	N	V	I	V	T	A
L	I	G	D	D	Q	I	Y	N	V	I	V	T	A

図9 生物間の相同部位比較による変則遺伝暗号推定の概念図

様々な生物で比較した、ヒトデの各コドンの存在位置に対応するサイトに出現するアミノ酸の出現頻度

	Human	Bovine	Mouse	Xenopus	Drosophila	
ATA /178	Ile	70	69	60	66	64
	Leu	47	46	38	48	48
	Met	18	19	19	13	8
	Val	12	11	7	14	14
	Pro	1	1	2	5	2
	Thr	9	7	6	6	1
	other	21	25	26	26	41
AAA /109	Asn	43	47	43	39	42
	Lys	5	7	5	6	2
	Ser	9	6	8	8	16
	Thr	11	7	8	9	1
	Ile	1	4	6	4	2
	Leu	6	6	7	10	7
	other	34	32	32	33	39
AAG /48	Lys	36	35	35	36	32
	Asn	2	1	2	1	4
	Leu	0	1	0	0	2
	other	10	11	11	11	10
AGA /54	Ser	17	19	16	13	21
	Arg	0	0	0	0	0
	Thr	7	4	3	4	3
	Gly	5	6	4	4	1
	other	25	25	31	33	29
AGG /19	Ser	4	5	5	8	5
	Arg	0	0	0	0	0
	His	1	0	1	0	0
	other	14	14	13	11	14

表5

表6

生物種	UGA (終止)	AUA (Ile)	AAA (Lys)	AGA/G (Arg)	CUN (Leu)	CGG (Arg)
棘皮動物	Trp	Ile	Asn	Ser	+	+
脊椎動物	Trp	Met	+	終止	+	+
節足動物	Trp	Met	+	Ser/?	+	+
軟体動物	Trp	Met	+	Ser	+	+
線形動物	Trp	Met	-	Ser	+	+
扁形動物	Trp	Met	Asn?	Ser	+	+
腔腸動物	(Trp)	(Met)	+	+	+	+
原生動物	Trp	Met	+	+	+	+
高等植物	+	+	+	+	+	+
緑藻植物	?	+	+	+	+	+
コケ植物	?	+	+	+	+	+
酵母菌	Trp	Met	+	+	+	+
真正子囊菌	Trp	+	+	+	+	+
細胞性粘菌	?	+	+	+	+	+

+: 普遍暗号 -: 不確定な暗号

?: 未同定または遺伝子中に見出されていない暗号

6. 蛋白質遺伝子

3. すでに述べたようにヒトデミトコンドリアゲノムには13個の蛋白質遺伝子がコードされており、その構成は脊椎動物、ウニ、シヨウジョウバエと同一であった。これらの蛋白質の同定は脊椎動物やウニの対応する蛋白質との相同性をその根拠とした。特に現在、いくつかの脊椎動物についてミトコンドリアDNAにコードされている遺伝子と蛋白質との対応関係がとれており、その後の様々な生物において解析されたミトコンドリアDNA中のオープンリーディングフレーム(ORF)の同定の"原盤"となっている。図10には遺伝子から推定されるCOIのアミノ酸配列を示し、それに相当するヒト、ウニのアミノ酸配列との比較を行っている。また表7にはそれぞれの蛋白質のアミノ酸残基数と分子量を示した。さらに表8は様々な生物間でアミノ酸の相同性の比率を求めた。ヒトデとウニの間で各遺伝子の相同性を比較するとCOI > COII > COIII > Cytb > ATPase6、ND1、ND3 > ND4L、ND4 > ND5 > ND6 > ND2 > ATPase8の順に保存されていた。この順序は基本的には脊椎動物間の相同性の序列と同じ傾向を示しており、相対的な各蛋白質の進化速度はほぼ同じ順序になっていたが、一部においては大きく異なっていた。特に序列に顕著な差が見られるのはATPase6である。ATPase6は哺乳類とアフリカツメガエルの間では9番目に保存されているがヒトデとウニ間では5番目であった。またND4Lについても同様のことがいえた。逆にND4はキョク皮動物間では9番目であったが哺乳類とアフリカツメガエルの間では6番目に保存されていた。また単なる序列関係ではなく序列間の数値的な内容に注目すると、キョク皮動物間と脊椎動物間ではかなり異なっていた。まずキョク皮動物間と哺乳類-アフリカツメガエル間のCOIについてはどちらも87%であり、これを基準に他の蛋白質を見たところ、例えばCOIIとCOIIIの順序はキョク皮動物間と脊椎動物間で逆転しており、それぞれキョク皮動物間が80%、75%であるのに対し、哺乳

類-アフリカツメガエル間は73%、80%となっており、その差は単なる統計誤差以上のものであった。また比較的相同性の低い蛋白質であるND4L、ATPase8、ND6についてはキョク皮動物間の保存性は脊椎動物間よりかなり高かった。例えばATPase8とND6は哺乳類とアフリカツメガエルの間でそれぞれ32%、35%であるがキョク皮動物間では46%、48%であった。この結果は特に脊椎動物で置換速度の速い蛋白質について、相対的にキョク皮動物では比較的緩やかに進化してきていることを示している。

次にキョク皮動物と脊椎動物との間の絶対的な置換速度について検討したところ、13個の蛋白質のうちND3、ND4L、ND6、COII、ATPase8、ATPase6、についてキョク皮動物間のほうが哺乳類-アフリカツメガエル間より相同性が高くなっている。蛋白質の進化速度の情報を得るにあたり、その蛋白質の置換速度が飽和している生物間については良い情報を得ることができない。しかし上にあげた6個の蛋白質も含めて13個の蛋白質全てについて、キョク皮動物-脊椎動物間の相同性はキョク皮動物間および脊椎動物間より明確に低くなっている。従って脊椎動物間、キョク皮動物間についてはまだ飽和していないだろう。ヒトデとウニは少なくとも5億年以前に分岐されたといわれている。それに対してアフリカツメガエルなどの両性類が初めて出現したのはデボン紀末期(約3億数千年前)であり、最初の哺乳類が見つかったのはそれよりさらに後であるので、両性類と哺乳類が分岐したのは、長く見積もっても3億年以前程度である。このことからすると相同性の値に多少の誤差を伴ったとしても、最低、これらの6つの蛋白質の進化は両性類以後の脊椎動物のほうがかかなり速いことになる。

各蛋白質の開始コドンはATGが8個、ATTが3個、GTGが2個であった。(表7) 開始コドンがATGの場合は特に問題はないが、ATTとGTGについては本当に開始コドンかどうかという疑問が生じる。そこでこれらのATG以

外が開始コドンであると思われる蛋白質遺伝子について、A T Gがどこに出てくるかを調べた所、N D 3 遺伝子についてはその上流ではA T Gが出現する前に終始コドンがあり、下流にはA T Gが遺伝子中に存在しなかった。またN D 5 遺伝子では約4 0 0 b p下流、C y t b 遺伝子ではすぐ隣に、N D 1 遺伝子では1 0 3 b p下流に、そしてN D 4 L 遺伝子では2 3 2 b p下流に最初のA T Gが存在していた。また遺伝子の上流域を調べるとN D 5 遺伝子、C y t b 遺伝子、N D 1 遺伝子、N D 4 L 遺伝子の全てにおいてA T Gコドンが出現する前に終始コドンが存在していた。これらの蛋白質のうちC y t bを除いて初めてA T Gの出でくるサイトは他の生物との相同性や蛋白質のサイズから見て、適当とは考えられない。この結果はC y t bを除くこれらの蛋白質において、A T Gが開始コドンである可能性を否定するものであった。

現在までに報告されている高等動物のミトコンドリアにおいても、開始コドンは必ずしもA T Gではないことが報告されている。脊椎動物のミトコンドリアではA T A、A T T、A T Cが遺伝子の5'端部分に存在する蛋白質が例外なく数個存在し、蛋白質の相同性などを根拠に開始コドンと考えられていたが、それらはウニやショウジョウバエでも1 3個の蛋白質中、数例ずつ報告されている。それに対してG T GはA T Nに比べてあまり見つかっていない。しかし従来、G T Gが開始コドンであると推定されるケースは脊椎動物には見つかってないとされていたが、G a d a l e t aらはラットのミトコンドリアゲノムを解析し、さらにすでに報告されたマウスなどのケースを再検討を行い、G T Gが開始コドンである可能性を指摘している¹⁷。またショウジョウバエ¹⁸やウニ²⁸などの非脊椎動物においても、G T Gが開始コドンであると推定されるケースが一例ずつ報告されている。

G a l a l e t aの指摘によると、G T Gが開始コドンとなるケースでは、となりにコードされるt R N A 遺伝子との間にスパーサーが1 b pも存在しない、

という共通性がみられる。この場合プロセッシング後はGが5'端に露出することが予想される。ヒトデにおいてGTGが開始コドンと思われるのはND1遺伝子とND5遺伝子であるが、どちらの場合も隣接するtRNA遺伝子との間にはスパーサーが存在しておらず、Gagdaletaraの指摘と合致する。(図11)

しかし最近トリで報告されたCOI遺伝子の場合、5'端のGTGとなりのtRNA Tyr遺伝子の間には1bpのスパーサーが存在している。この場合tRNA Tyr遺伝子は相補鎖にコードされておりアンチセンスtRNAがどのようにプロセッシングされるかが問題であろう。この場合、アンチセンスtRNAにおいてスパーサーの1bpがセンスtRNAのディスクリミネーター塩基様にヌクレアーゼに認識されるのかもしれない。

GTGが開始コドンになる例は、まれではあるが原核生物において存在していることや、我々の結果を含めて、高等生物のミトコンドリアにおいてのべ7、8例、発見されたことなどから、本研究においてはこれらのGTGを開始コドンと推定している。

終始コドンについて見ると13個の蛋白質遺伝子中11個に存在し、その内分けはTAAが9個、TAGが2個であった。残りのCOII遺伝子とCytb遺伝子については終始コドンが出現する前にもtRNA遺伝子がコードされていた。これを無視して終始コドンを求めるとtRNA遺伝子との境界から数えて、COIでは12bp、Cytb遺伝子では108bp下流に終始コドンが出現した。しかし特にCytb遺伝子については蛋白質のサイズから見てこの終始コドンは適当でないであろう。

脊椎動物においても終始コドンのない蛋白質はしばしば存在するが、これらに対して良い説明がなされている³⁹。それはこの遺伝子を含めゲノム全体がポリストロニックに転写され、次にこの遺伝子の3'側に隣接するtRNAが切断された後、露出した3'末端がポリアデニル化され終始コドンが出現するというも

のである。そのためには3'末端がT、TA、AG (AGAは脊椎動物において終始コドン)である必要があり、さらにフレーム枠が一致しなければならない。

ヒトデの終始コドンのない二つの蛋白質遺伝子について、このことを検討した所、どちらの場合についてもtRNAが切断された後に存在する3'末端はTであり、これがポリアデニル化された後はTAAがインフレームで出現することが解った。(図12) 従ってヒトデミトコンドリアのシステムにおいても、脊椎動物と同じく、tRNAのプロセッシングやポリアデニル化などが本来以外の多様な機能を持っていることが推察される。

蛋白質遺伝子の配置について見るとND3遺伝子はtRNA^{Ser}(UCN)遺伝子とATPase8遺伝子はATPase8遺伝子と重複が見られた。このうちND3遺伝子の重複は相補鎖間似よるものであり特に問題とは考えられない。ATPase8遺伝子とATPase6遺伝子の間の重複は16bpにわたっており、mRNAのプロセッシングあるいは翻訳上のメカニズムが問題となる。この両者の重複は生物間で良く保存されている。その重複の長さは様々でヒト、マウス、ウシなどは長くそれぞれ46bp、43bp、40bpである^{65,66}。またトリでは10bp、ウニでは4bpである。ヒト、ウシ、トリにおいてRNA転写物の解析がなされたところ、この両者のmRNAは分離されておらず1本であることが報告された。このことは翻訳レベルでこの2個の遺伝子を発現させるための何等かのシステムがあることを示している。これらの重複はいずれも3N+1となっており、ヒトデにおいても例外ではなかった。(図13) ATPase8遺伝子は進化速度が速く3'末端付近においてもさほど保存されていない。しかし広い生物間でこの重複が保存されていることから、脊椎動物に働いている未知のシステムが広く保存されているのであろう。

次にヒトデミトコンドリアのコドンの使用頻度を蛋白質のコードされている鎖ごとに分けて求めたものを表9に示す。表9を見るとそれぞれコードされている

鎖によりコドンの使用頻度の傾向が異なっていることが解る。本研究ではこのことからミトコンドリアゲノムに働いている *Base-bias* を導いたがそれについては後章で詳しく述べる。

表9をではミトコンドリアでは特にコドンの2文字目にTが使われていることが多く全体の40%以上を占めている。これらのコドンが指定するアミノ酸は疎水性アミノ酸であり、これに他の疎水性アミノ酸であるプロリン、グリシン、アラニンを加えると全体の約60%が疎水性アミノ酸によって占められることになるが、このことはミトコンドリアゲノムにコードされる蛋白質が全て膜蛋白質であり、疎水性が強いことを反映しているのだろう。

後に示すコドン-*bias*以外にコドンの使われ方には若干の方言が見られる。まずCあるいはTの"ならび"が良く選択されている。コドンの2文字目がCであるコドンを見るとアラニンを除きNCCが最も良く使用されている。アラニンについてもGCCとGCAはほぼ等しい。これは他の4コドンファミリーにおいてコドンの3文字目にAが使用される頻度がCが使われる頻度の2~3倍になっていることと対照的である。またコドンの2文字目にTが使われているコドンにおいてはNTTがNTCの約1.5倍~2倍である。これも他のコドンファミリーにおいてNNTがNNCに比べ~1倍程度であることと対照的である。ミトコンドリアでは後に示す*bias*などの要素も関係するため、このようなコドンの方言はあまりはっきりしたものではない。

次にヒトデのCGNのコドンボックスを見るとCGAが全体の60%以上を占めている。他の生物について同じコドンボックスを調べたところ、ヒトをのぞく脊椎動物、ウニ、ショウジョウバエでこの傾向が保存されていた。またヒトにおいてもCGAが最も良く使用されていることが解った。これらの生物は*bias*については共通の傾向をもっているものの、他のコドンボックスにおいてNNAの存在比は様々であり、このようにコドンボックスの傾向が一致しているものは

他に存在しない。特に脊椎動物やキョク皮動物とは全く異なるコドンの使用頻度の傾向をしているにもかかわらず、このコドンボックスにおいては共通していることからして、おそらく蛋白質合成上の共通した調節機構が存在するのかもしれない。

GENETYX : Amino Acid Sequence Homology Data]

1st Amino Acid Sequence
 File Name : B:apcolaa.NUC
 Sequence Size : 517

2nd Amino Acid Sequence
 File Name : B:SPCO1AA.NUC
 Sequence Size : 517

Unit Size to compare = 2

[87.2% / 517aa]

```

1' MQLSRWFPSTNIHKDIGTLYLIFGAWAGMAGTAMSVIIRTELAQPGSLIQDDQIYNVIVTA
*****
1" MQLSRWLPSTNIHKDIGTLYLIFGAWAGMAGTAMSVIIRTELAQPGSLIQDDQIYNVVVTA

61' HALVMIFFMVMPIMIGGFGNWLTPLMIGAPDMAFPRMNNMSFWLIPPSFLLLLASAGVES
.,*****
61" QSLVMIFFMVMPIMIGGFGNWLTPLMIGAPDMAFPRMNNMSFWLIPPSFLLLLASAGVEN

121' GAGTGWTIYPPLSSGLAHAGGSVDLAI FSLHLGASSILASINPITTVINMRTPGISFDL
*****
121" GAGTGWTIYPPLSSNITHAGSSVDLAI FSLHLGASSILGLINPITTVINMRTPGMSLDL

181' LPLFVWSVFVTAFLLLLSLPVLAGAITMLLTDNRNVTTFPDPAGGGDPIILFQHLFPWFFGH
*****
181" LPLFVWSVFVTAFLLLLSLPVLAGAITMLLTDNRNINTTFPDPAGGGDPIILFQHLFPWLFGH

241' PEVYILILPGFGMISHVIAHYAGKSEPPGYLGMVVAIVSIGILGFLVWAHHMFTVGMDDVD
*****
241" PEVYILILPGFGMISHVIAHYSKREPPGYLGLVYAMIAIGVLGFLVWAHHMFTVGMDDVD

301' TRAYFTAATMIITAVPTGIVKFSWMA TLQGSNLRWDTPLLWALGFVFLPTIGGLTGIVLAN
*****
301" TRAYFTAATMIITAVPTGIVKFSWMAKIQGSNLQWSLPLLWTLGIVFLPTIGGLTGIVLAN

361' SSIDFVLIHDTYVVVAHFHYVLSMGAVFAIFAGFTHWFPLPSGVGLHPLWGNHHTLMFTG
***
361" SSIDFVLIHDTYVVVAHFHYVLSMGAVFAIFAGFTHWFPLPSGYSLHPLWGVHHTLMPVGV

421' VNLTFPPQHFLGLAGMPRRYSYDPDAYTLWNTVSSIGSTISLIATLVFLFLLWEAFPSQR
*****
421" VNLTFPPQHFLGLAGMPRRYSYDPDAYTLWNTISSIGSTISVVAMLFFLPLIWEAFASQR

481' TALQPEPSTSSLEWQYSSFPSSHHTFDEIPSTVYLIK
...
481" EGITPEPSSHASLEWQYTSFPSSHHTFDETPSTIIVK
  
```

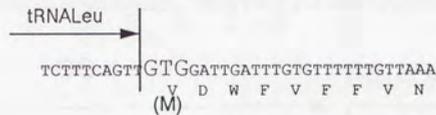
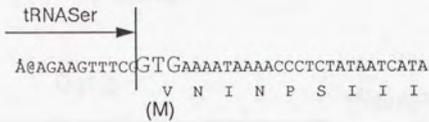
図10 COIの相同性比較図 ヒトデーウニ

Gene	Size		開始コドン	終止コドン
	アミノ酸	分子量		
ND1	326	36091	GTG	TAG
ND2	354	38653	ATG	TAA
ND3	110	12603	ATT	TAG
ND4	460	50767	ATG	TAA
ND4L	98	16747	ATT	TAA
ND5	641	71054	GTG	TAA
ND6	162	17414	ATG	TAA
CO I	517	57034	ATG	TAA
CO II	229	26074	ATG	T
CO III	260	29587	ATG	TAA
ATPase6	230	25506	ATG	TAA
ATPase8	54	6241	ATG	TAA
Cyt <i>b</i>	379	42646	ATT	T

表7 各蛋白質遺伝子の開始コドンと終止コドン

	脊椎-棘皮		棘皮-棘皮	脊椎-脊椎		
	ヒト ヒトデ	ヒト ウニ	ウニ ヒトデ	カエル トリ	トリ 哺乳類	カエル 哺乳類
ND 1	52	55	64	70	70	67
ND 2	36	39	47	57	45	51
ND 3	49	48	63	67	57	58
ND 4	43	44	56	61	59	59
ND 4L	36	37	59	41	47	44
ND 5	44	42	52	62	56	57
ND 6	30	29	48	42	27	35
CO I	75	75	87	86	86	87
CO II	59	61	80	70	67	73
CO III	69	62	75	80	75	80
ATPase 6	43	37	65	68	55	53
ATPase 8	25	19	46	47	27	32
Cyt b	64	63	71	75	74	74

表 8 各蛋白質における脊椎動物-棘皮動物、棘皮動物間、脊椎動物間のアミノ酸配列の相同性



他にスパーサーがないのはCOIIだけである

他の生物にてでくるGTG		隣接するtRNA遺伝子との間の塩基数
マウス	ND1	0
ラット	ND1	0
ショウジョウバエ	ND5	0
ウニ(P.I.)	ATPase8	0
ウニ(S.p.)	ATPase8	0
トリ	COI	1 アンチセンス アンチセンスtRNAにおける 識別塩基として切断される？

図11 GTGで始まる蛋白質遺伝子の特徴

停止コドンのない遺伝子における終止コドンの生成

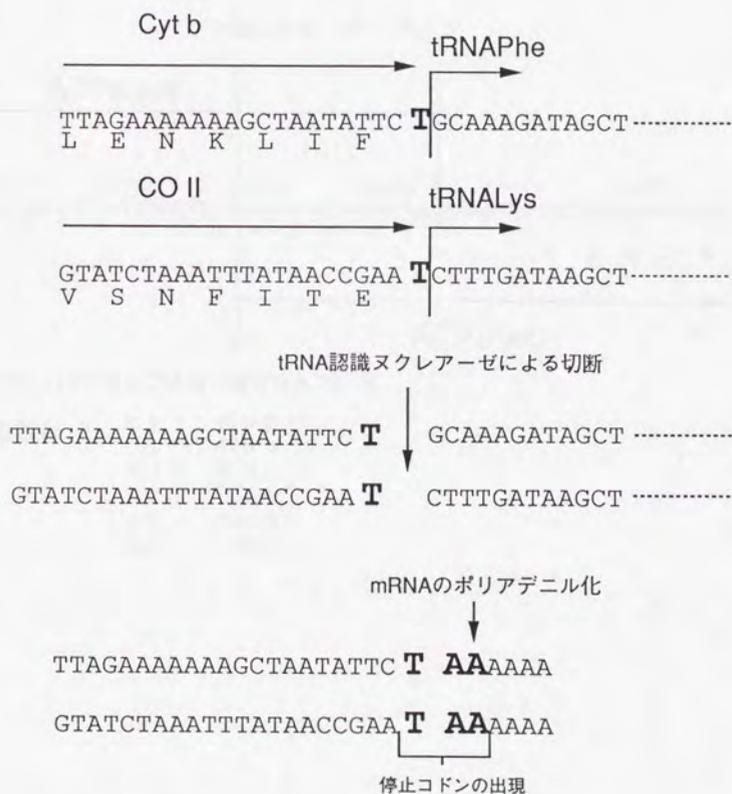
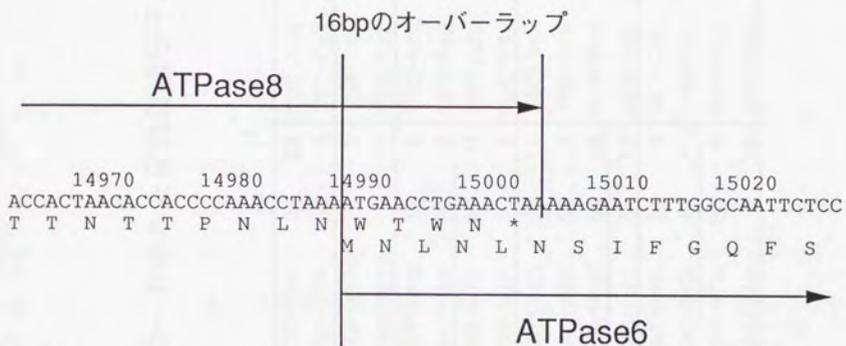


図12 終止コドンのない蛋白質遺伝子とその終止コドンの生成



この部分のオーバーラップは良く保存されている

3N+1bpの保存	ヒト	46 bp
	ウシ	43 bp
	マウス	40 bp
	カエル	10 bp
	トリ	10 bp
	ウニ	4 bp

図13 ATPase8遺伝子とATPase6遺伝子のオーバーラップ

表9 ヒトデミトコンドリアゲノムにコードされた蛋白質遺伝子の鎖別に求めた
コドンの使用頻度

	L H			L H			L H			L H		
TTT-Phe	145	58	TCT-Ser	51	43	TAT-Tyr	27	25	TGT-Cys	8	6	
TTC-Phe	105	8	TCC-Ser	82	4	TAC-Tyr	58	9	TGC-Cys	17	4	
TTA-Leu	93	53	TCA-Ser	64	3	TAA-stop	7	2	TGA-Trp	72	6	
TTG-Leu	14	51	TCG-Ser	13	10	TAG-stop	1	1	TGG-Trp	11	13	
CTT-Leu	104	25	CCT-Pro	43	16	CAT-His	18	5	CGT-Arg	3	5	
CTC-Leu	70	2	CCC-Pro	58	3	CAC-His	59	2	CGC-Arg	4	1	
CTA-Leu	161	8	CCA-Pro	41	6	CAA-Gln	57	7	CGA-Arg	40	5	
CTG-Leu	22	12	CCG-Pro	3	5	CAG-Gln	10	9	CGG-Arg	10	4	
ATT-Ile	104	25	ACT-Thr	57	15	AAT-Asn	26	15	AGT-Ser	10	13	
ATC-Ile	61	5	ACC-Thr	101	4	AAC-Asn	56	2	AGC-Ser	21	2	
ATA-Ile	152	25	ACA-Thr	66	3	AAA-Asn	94	18	AGA-Ser	40	14	
ATG-Met	59	19	ACG-Thr	13	2	AAG-Lys	31	17	AGG-Ser	7	12	
GTT-Val	33	51	GCT-Ala	33	25	GAT-Asp	12	9	GGT-Gly	21	17	
GTC-Val	34	4	GCC-Ala	71	5	GAC-Asp	47	1	GGC-Gly	26	6	
GTA-Val	62	19	GCA-Ala	71	7	GAA-Glu	60	5	GGA-Gly	96	16	
GTG-Val	10	27	GCG-Ala	14	6	GAG-Glu	6	16	GGG-Gly	21	29	

7. tRNA 遺伝子

ヒトデミトコンドリアゲノムで見つかった22種類のtRNA遺伝子から推定されるtRNAの二次構造を図14、15に示す。その内訳はセリンを指定するUCN、AGNのコドンボックスに対するtRNAがそれぞれ1種類、ロイシンを指定するCUN、UURに対応するものが各1種類づつとなっており、その他の18種類のアミノ酸に対応するものが各1種類である。その内訳は他の高等動物と同じである。このうち11種類がH鎖にコードされており、残りがL鎖にコードされている。

通常ミトコンドリア以外の系では蛋白質合成に関与するtRNAの種類数は40~50種類であるとされるがミトコンドリアの系では酵母では24種類、高等生物では22種類であり、tRNAの大幅な省略化が生じている。

一般のtRNAには安定な高次構造を組むために共通する構造や特徴を保持している。それらの特徴は、まず3つのステム、2つのループなどの塩基数が保存されている。ステム部分は水素結合を強くするため、一般にG-Cペアが50%以上である。ほとんどG-Uペアやミスマッチなどが存在しない。またtRNAはほとんど例外なくD-loop中にGG配列、TΨC-loop中にTΨCA配列が保存されている。X線結晶解析からtRNAはL字型の三次元構造を取っていることが知られているが⁶⁷、その三次元構造はこれらのD-loopとTΨC-loop中の保存塩基が水素結合を形成することにより、保持されていることがわかっている。また8番目の塩基はUとなっておりD-loop中のAと水素結合を形成して、構造の安定化に寄与している。またTΨC-loopは7bpとなっている。さらにこれらは転写後、様々な修飾を受けている。これらの要素やそれによるtRNAの構造の安定化はアミノアシル合成酵素、延長因子、リボソームなどに対する基質特異性を高めるとともに、mRNAとのコドンとアンチコドンによる会合においても、正確性を高める効果があるのであろう。しい

ては正確な蛋白質合成を行うために必要なものであろう。それに対して脊椎動物などで報告されたミトコンドリアのtRNAは、遺伝子が決定されたもの、RNAレベルで解析されたものなど様々であるが、それらはミトコンドリア以外のものと比べて対照的な特徴を持っている。それはG-C含量が低く、G-Uペア、ミスマッチが多く、保存配列がほとんどなく、loopの数が様々で、かつ修飾塩基がほとんどないというものであった。一方、セリンのtRNA AGYを除き、これらのミトコンドリアのtRNAにおいても、アクセプターシステム、アンチコドンシステム、アンチコドンループなどの塩基数、アクセプターシステムとDシステム間の2塩基長、Dシステムとアンチコドンシステム間の1塩基長などは良く保存されている。本研究においてもこれらを基準にtRNA遺伝子を同定した。

以上に述べた観点からイトマキヒトデのミトコンドリアのtRNAを調べたところ、基本的には脊椎動物のtRNAの特徴と同じであった。しかしGG配列、TΨC配列の見られるtRNAはイトマキヒトデの方がかなり多かった。まずイトマキヒトデではTΨCRA配列が保存されているtRNAが7種類あった。またTΨCRA配列のうち4bpまで保存されているものを含めると14種類であった。さらにGG配列については11種類のtRNAに保存されていた。TΨCRA配列が完全に保存されている7種類のtRNAについては全てGG配列が存在していた。(図16) またGG配列の保存されている11種類のtRNA側から見ると、前の7種類のtRNAを含めて10種類までにおいてTΨCRA配列のうち4bpが保存されていた。脊椎動物においてはこのようにGG配列、TΨCRA配列が共に保存されているのは、tRNA Ser(UCN)、tRNA Leu(UUR)、tRNA Glnの3種類だけである。ウニについて調べると、イトマキヒトデより少なく2種類のウニのどちらにおいても4種類だけであった。またGG配列が保存されているtRNA、TΨCRA配列のうち4bpが保存されているtRNAを別々に調べたところ、*P. lividus*ではTΨCRA配

列の内4bpが保存されているものが14種類あったが、5塩基が完全に保存されているものは5種類であり、またGG配列は4種類でしか見られなかった。さらに*J. purpuratus*ではどちらも8種類づつで、やはりイトマキヒトデよりかなり少ない結果であった。しかしこれらのウニにおいてもTΨCRA配列が完全に保存されている場合はほとんどGG配列を伴っており、2種類のウニで延べ9種類中8種類という結果であった。荒木によって解析されたキヒトデにおいてはウニと同じくイトマキヒトデよりかなり保存塩基の存在は少なかった。従ってイトマキヒトデで共通塩基が良く保存されていることは、ヒトデの特徴とは言えない。

TΨCRA配列とGG配列は水素結合を組んで高次構造を形成するという意味で関係しており、独立に存在しても、さほど意味があるとはいえない。また原核生物においてはTΨC配列は16SrRNAと相互作用をしているとの報告があるが、これについても、GG配列と水素結合を形成して、L字型構造を取ることが前提であろう。先の結果のようにTΨCRA配列とGG配列の高い相関関係はこれらを持つミトコンドリアのtRNAにおいても、これらの配列が相互作用をして、高次構造をとるのに役立っていることを示すものである。

イトマキヒトデ、キヒトデと2種類のウニの全てにおいてTΨCRA配列とGG配列が比較的保存されているのはtRNA Ser (AGN)と2種類のロイシンのtRNAであった。(表10) このうちtRNA Ser (AGN)は高等動物ミトコンドリアに共通して異常構造を持つtRNAであり、生物種によって、塩基置換などというレベルでなく、Dアームのサイズが縮小するという過激に進化した分子種である。このtRNAについては後に詳しく述べるが、遺伝暗号の変化を説明する上で立体構造が特に重要な要素であると思われる、高次構造の解明が望まれるものであるが、このように共通塩基がキョク皮動物で広く保存されていることから、それが機能的な要素である可能性は高い。

またロイシンの2種類のtRNAも後に詳しく述べるが、tRNAの進化上、極めて注目される分子種である。この2種類のtRNAはおそらく1種類のアミノアシル合成酵素(ARS)によって認識されるという強い状況証拠があり、アンチコドンはARSの認識部位ではないか、あるいは弱い認識部位であると思われる。

ウシミトコンドリアにおける数種類のARSの研究によると、それらのARSは様々なミトコンドリア以外のそのアミノ酸に対応するtRNAを認識し、アミノアシル化を行うということが報告されている。それらのヘテロなtRNAに共通する部位はアンチコドンだけであることが多く、ARSの主な認識部位はアンチコドンである場合が多いと思われる。これらのことを総括して考えると、たいていのtRNAはアンチコドンがARSの主な認識部位であり、他の部分が少々ルーズでも問題がないことが多いであろう。

しかしロイシンについては、1種類のARSが2種類のアンチコドンの異なるtRNAを認識すると考えると、アンチコドンがARSの主要な認識部位でない可能性が高いと思われる。従ってロイシンのARSはその他の部分に情報を求める必要があり、さらにそれは2種類のtRNAに共通している必要があるであろう。そのためにこの2種類のtRNAはアンチコドン以外の部分が進化上大きな制約を受け、共通塩基が保存されたのかもしれない。

一方、他のtRNAにおいてはアンチコドンは言うまでもないが、アンチコドンループ全体を見てもtRNAの他の部分に比べて広い種間で最も保存されており、アンチコドンの周囲の環境をなるべく変化させないような制約があると思われる。このことはアンチコドンがARSの主な認識部位であることを反映しているのかもしれない。

次にキョク皮動物において異常な遺伝暗号となっているコドンに対応するtRNAの構造を調べた。

まずイソロイシンのコドンである A T T、A T C、A T A に対応する t R N A I l e ではそのアンチコドンが G A T となっていた。(図 17) さらにメチオニンのコドンである A T G に対応する t R N A M e t ではアンチコドンが C A T であった。脊椎動物やショウジョウバエの場合イソロイシンのコドンは A T T、A T C でありメチオニンのコドンは A T A、A T G である。しかし遺伝子上それらに対応する t R N A のアンチコドンはそれぞれ G A T と C A T であり、ヒトデとの差異は見られなかった。

まずメチオニンのコドンについてであるが、高等動物ミトコンドリアの場合、N N A と N N G の 2 コドンファミリーに対応する他の t R N A では例外なく遺伝子上でアンチコドンの 1 文字目は T であることからすると、この脊椎動物、ショウジョウバエの t R N A M e t の結果は例外的であり、遺伝子上の解析だけでは理解できない。ウシや蚊のミトコンドリアの t R N A M e t の解析の結果によるとアンチコドン 1 文字目の C は未修飾で存在していることが示されているが、それが A T G だけではなく A T A を認識することは通常では考えにくい。

従来ミトコンドリアには細胞質からの t R N A の移入はないとされていたが、最近になって植物ミトコンドリアにおいては細胞質から一部の t R N A が移入されていることが判明した³¹。さらに動物ミトコンドリアにおいても t R N A の移入があるという情報もありはっきりしていない。その点ヒトデの結果は遺伝子上で良くつじつまが合っており、アンチコドンが未修飾の C A T であればそれに対応するコドンは A T G だけであろう。従ってこれらの差異を説明するためには、脊椎動物側のアプローチが望まれる。

最近ウシの t R N A M e t のアンチコドン 1 文字目の C は修飾されているかもしれないという、従来とは異なる可能性が上田、横川、森谷らによって示唆されており、この点については彼らの解析中の結果が期待される。

次にヒトデのイソロイシンの t R N A については遺伝子中でアンチコドン G A

TがATC、ATT、ATAの三つのコドンに対応している。このうち転写物が未修飾であってもATCとATTには対応できるが、ATAには対応できるとは考えられない。一方、キョク皮動物以外の生物種においても遺伝子上のアンチコドンはGATであるが、それに対応するコドンがATTとATCの2種類であり、他の2コドンファミリーにおけるコドンとアンチコドンの対号のシステムと統一性のとれた結果である。これらのようにヒトデについては遺伝子解析だけでは他の高等動物との遺伝暗号の差異の説明ができないため、tRNAの解析をする必要があるであろう。この点については以下で説明するtRNA Asnも同様の結果であった。

次にAANのコドンボックスにおけるコドンとアンチコドンとの関係について考察する。このコドンボックスについてもヒトデは他の植物、下等生物、高等動物などほとんどの生物種のミトコンドリアとは読み分けが異なっている。ヒトデについてリシンのコドンがAAGだけであり、それに対するtRNA LysのアンチコドンはCTTであった。(図17) 一方ミトコンドリアにおいてもほとんどの生物種ではリシンのコドンはAAAとAAGであり、普遍暗号と同じであるが、それに対応するtRNAのアンチコドンはショウジョウバエを除いてTTTになっている。従ってキョク皮動物とその他の生物種における遺伝暗号の差異は、アンチコドンがTTTについてはAAAとAAGの2種類のコドンに対応できるが、CTTの場合対応するコドンはAAGだけであるというtRNAのアンチコドンの差異によって説明できるであろう。一方ヒトデのアスパラギンにコドンはAAT、AAC、AAAの三つであるがこれに対応するtRNA AsnのアンチコドンはGTTであった。それに対して他の生物種のアスパラギンのコドンはAAT、AACの2種類であるが、tRNA AsnのアンチコドンはやはりGTTであり、遺伝子上でアンチコドンに差異はなかった。これらの状況はイソロイシンのものと同一であり、ヒトデにおいてGTTが三つのコドンに対応する

ことを説明するためにはtRNA^{Asn}の解析が必要と思われた。残念ながらヒトデ卵巣のミトコンドリアにおけるミトコンドリアtRNAの量は極めて少なく、現時点でtRNAの解析の結果は得られていない。

ヒトデミトコンドリアのtRNA^{Ile}やtRNA^{Asn}のように1種類のtRNAがNN₁T、NN₁C、NN₁Aの三つのコドンに対応するのは、ミトコンドリア以外のシステムでしばしば見られる。

普遍暗号では唯一イソロイシンのコドンだけが、3コドンファミリーであり、それはATT、ATC、ATAであるが原核生物ではこれに対応するtRNAは2種類になっており、ATT及びATCとATAとを読み分けている。それに対し真核生物ではイソロイシンのtRNAは1種類しか見つかっていないが、tRNAの解析の結果によるとそのアンチコドンの1文字目はイノシンであることがほとんどである。それらの遺伝子は例外なくアンチコドンの1文字目がAであり、イノシンはAから誘導されたことが解っている。また真核生物には4コドンファミリーをNN(T、C、A)とNNGの3:1に読み分けているものがしばしば存在しているが、その3種類のコドンに対応するtRNAのアンチコドンはやはりIであることが多い。IはU、C、Aとwobble塩基対を形成することができるため、このように3種類のコドンに対応できる訳である。

さてヒトデミトコンドリアのイソロイシンとアスパラギンはこのように3種類のコドンに対応する必要があるわけで、そのアンチコドンの1文字目がIであれば良いわけだが、遺伝子上ではこの部分はどちらもGである。しかしアンチコドンがIになっているtRNAの遺伝子はその部分が例外なくAとなっている。従ってこのアンチコドンが修飾されるとしても、従来とは異なった代謝経路でIに変換されるか、あるいはI以外の修飾塩基であるかもしれない。

ヒトデのセリンのtRNA(AGN)についてはアンチコドンがGCTであるが、それは下等生物から脊椎動物に至るまでほとんど保存されている。しかしそ

れに対応するコドンは進化の過程で変遷している。下等生物ではAGYがセリンでありAGRはアルギニンであるが、非脊椎高等動物ではAGNがセリンになっている。そしてAGRに対応するはずであったアルギニンのtRNAは存在していない。さらに脊椎動物ではAGYがセリンのコドンでありAGRは停止コドンとなっている。非脊椎高等動物ではアンチコドンGCTが4種類のコドンAGNに対応しているわけである。このように1種類のtRNAが4種類のコドンに対応するのは、ほとんどミトコンドリアだけに見られる特徴ではあるが、このケース以外は例外なくアンチコドンの1文字目がTとなっている。従ってこれらのコドンとアンチコドンの対合関係は従来の描像では説明できない。

高等生物のtRNA Ser (AGN)は共通してきちんとしたDアームに欠く構造を持っており、立体的にはかなり不安定な構造であろう。なかでも脊椎動物のtRNAはほとんどDアームの痕跡が残っていない。それにたいして先に述べたようにキョク皮動物ではステムアンドループにはなっていないが、GG配列とTTCRA配列を持っており脊椎動物より安定な構造をしている可能性が強い。これらの差異が立体化学的な意味でコドンの読み分けの差異になっているのかも知れない。(図18)

最後にロイシンのtRNAの進化について考察する。ヒトデヤウニではND5遺伝子が脊椎動物と比較して約100bp程長くなっている。この周囲の遺伝子配置は脊椎動物ではtRNA His、tRNA Ser (AGN)、tRNA Leu (CUN)、ND5となっているが、キョク皮動物ではtRNA His、tRNA Ser (AGN)、ND5となっておりtRNA Leu (CUN)が欠落している。その代わりに約100bpほど長くなっているのである。(図19)ウニのグループがこの部分を詳しく調べたところ、tRNA-likeな配列がこの部分に存在し、それは脊椎動物などのtRNA Leu (CUN)遺伝子と相溶性が高いとの結果を得たと報告した⁹⁸。つまりこの約100bpにはもともと

tRNA Leu (CUN) 遺伝子が存在してのだが、何等かの理由により疑似遺伝子化した可能性が強いというものであった。さらに現存している tRNA Leu (CUN) 遺伝子を tRNA Leu (UUR) 遺伝子と比較したところ、83.1%の相同性があることが解った。その報告を受けて我々もヒトデについて検討してみたところ、まず ND5 遺伝子の 5' 端の約 100 bp の部分にやはり脊椎動物の tRNA Leu (CUN) 遺伝子と約 50% 程度の相同性があることが解った。またヒトデの tRNA Leu (CUN) 遺伝子と tRNA Leu (UUR) 遺伝子の間には 76.1% の相同性が見られた。(図 19) それに対して脊椎動物の 2 種類のロイシンの tRNA 遺伝子の間にはせいぜい 40% ~ 50% 程度の相同性しか見られない。これらの結果を総括するとミトコンドリアの遺伝子配置の変化の様相の一端を伺うことができる。(図 20)

まずキョク皮動物において 2 種類の tRNA 遺伝子の相同性が高く、一方脊椎動物ではそうではないことから、キョク皮動物については現存する tRNA Leu (CUN) は tRNA Leu (UUR) からの派生物であるだろう。ミトコンドリアの大抵の tRNA ではアミノアシル合成酵素の主な認識部位はアンチコドンであると思われるが、ミトコンドリア以外の蛋白質合成系については、アンチコドン以外の部分にARSの主な認識部位が移行しているケースも多い。キョク皮動物の 2 種類のロイシンの場合もこのようなケースであると考えられる。さらにアンチコドン以外の部分にARSが情報を求めるため、tRNAの進化速度が制約を受け、キョク皮動物では例外なく 2 種類の tRNA 中に GG 配列 TΨCR A 配列が保存されているのかもしれない。

tRNA Leu (CUN) が tRNA Leu (UUR) の派生物であるということは、過去に tRNA Leu (UUR) が組み込まれて、重複してゲノム中に存在したことを示している。また先祖型の tRNA Leu (CUN) の痕跡があることも、遺伝子配置の変化を考察するうえで、良い情報であった。不要な配列

を廃出して縮小化しているミトコンドリアゲノムにおいて、このように疑似遺伝子が残っていることは注目されることであるが、おそらく、偶然に元の tRNA^{Leu} (CUN) 遺伝子中に開始コドンがフレームに沿うように存在し、終始コドンが存在せず、さらに余分に翻訳される N 末のポリペプチドが蛋白質の機能にネガティブな影響を与えないなど、の偶然があったのであろう。

図14 ヒトデミトコンドリアゲノムのL鎖にコードされるtRNA遺伝子の2次構造

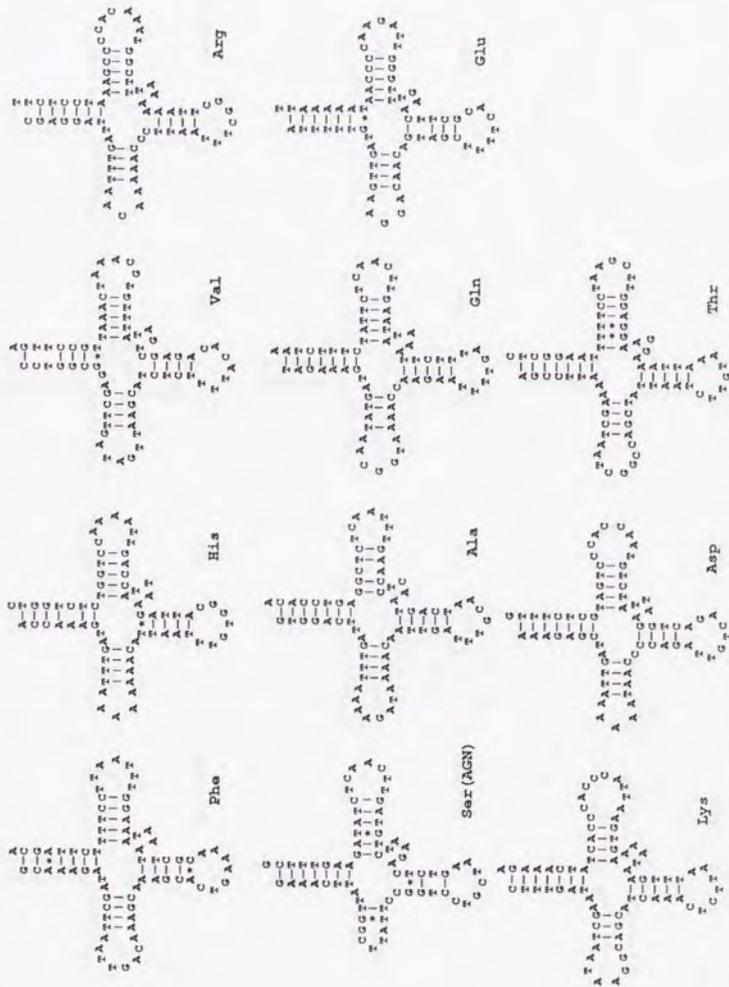
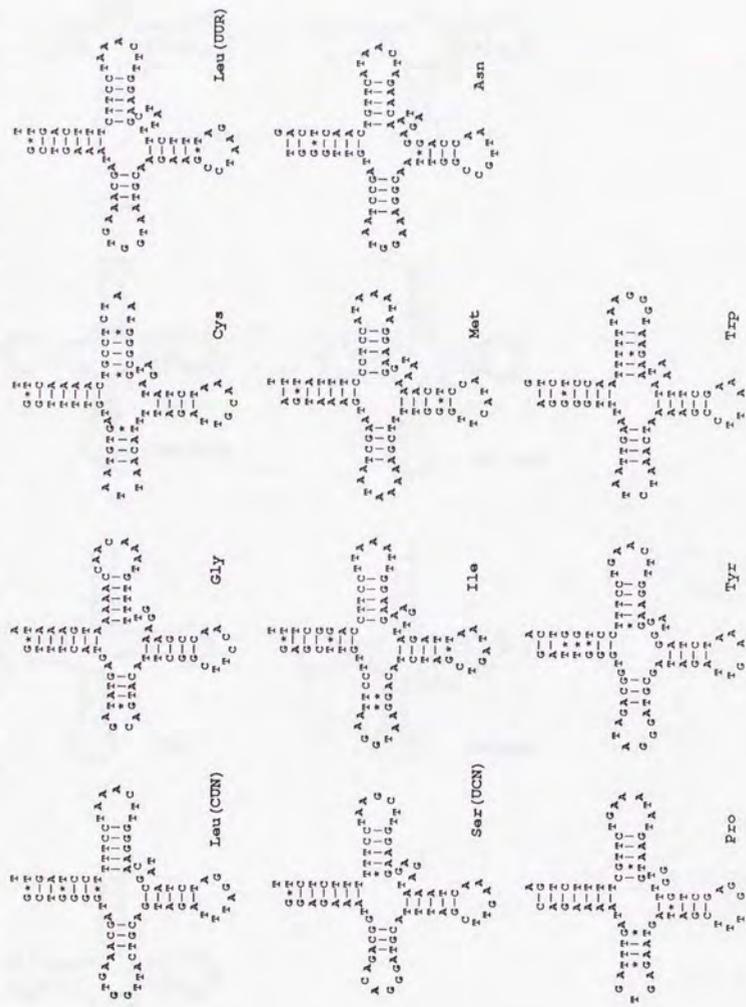


図15 ヒトデミトコンドリアゲノムのH鎖にコードされるtRNA遺伝子の2次構造



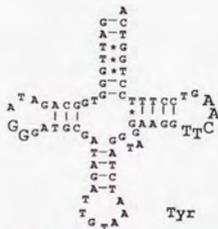
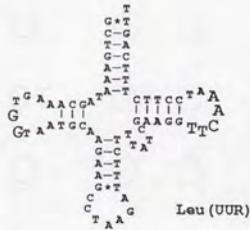
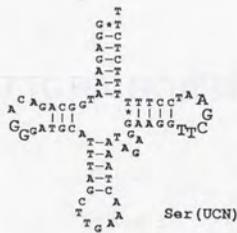
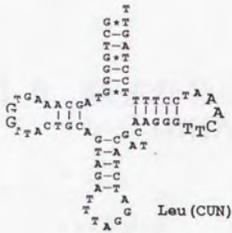
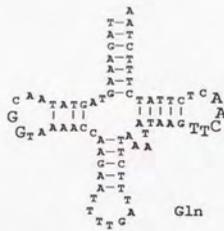
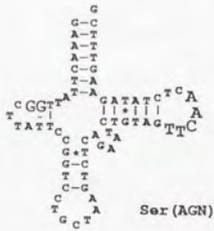


図16 TTCRA-GG配列をもつtRNA

表10 棘皮動物においてTTCRA-GG配列の保存されているtRNA

	イトマキヒトデ		キヒトデ		バフンウニ		ウニ(<i>P. lividus</i>)	
	TTCRA	GG	TTCRA	GG	TTCRA	GG	TTCRA	GG
tRNA ^{Leu} _{CUN}	○	○	○	○	○	○	○	○
tRNA ^{Leu} _{UUR}	○	○	○	GT	○	○	○	○
tRNA ^{Ser} _{AGN}	○	○	○	○	○	○	○	○
tRNA ^{Ser} _{UCN}	○	○	○	○	X	○	○	GA

AUN box におけるコドンとアンチコドンの関係 AAN box におけるコドンとアンチコドンの関係

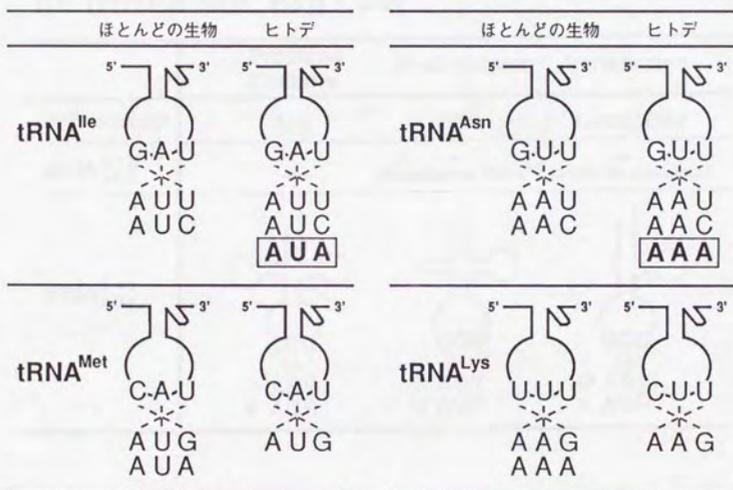


図17

Change of Codons AGR with Change of tRNAs for Arg / Ser

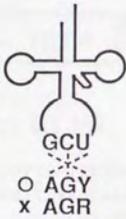
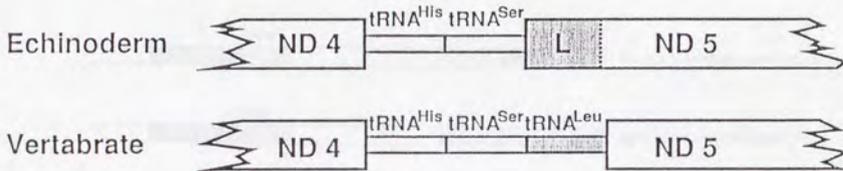
	Protozoa Cnidaria	Invertebrates	Vertebrates
AGR codons	Arg	Ser	Terminator
tRNA ^{Arg} _{AGR}	+	[Deletion of tRNA ^{Arg} (AGR) in Genomes]	
tRNA ^{Ser} _{AGY}	 <p>GCU ○ AGY x AGR</p>	 <p>GCU ○ AGY ○ AGR</p>	 <p>GCU ○ AGY x AGR</p>

図18

Comparison of 5' - adjacent region of ND 5 gene between Echinoderm and Vertebrate mitochondrial genomes



Sea urchin (*P. lividus*): 83.1%

Leu(CUN) GCTGAAATAGCAAAGTGGTTAATGCAAAAAGCTTAGGA|TTTTTC-ATCGAAAGGTTCAACTCCTTCTTTCA
 Leu(UUR) GCTAAAATA6CAAAGTGGTTAATGCA6AAGGCTAAGA|CCTTCCTATCAAAGGTTCAACTCCCTTTTTTA
 (CTTAGAA)
 Anticodon loop of the tRNA^{Leu}(CUN) gene buried in ND5 gene

Starfish (*A. pectinifera*): 76.1%

Leu(CUN) GCTG666TAGCAAAGTGGTTACTGCAGTAGA|TTTAGGA|CTA-CGATCAAAGGTTCAAATCCTTCTTTCA
 Leu(UUR) GCTGAAATAGCAAAGTGGT-AATGCAAGAAGCTAAGA|TTCTTTATCGAAAGGTTCAAATCCTTCTTTCA
 (TATAAAA)

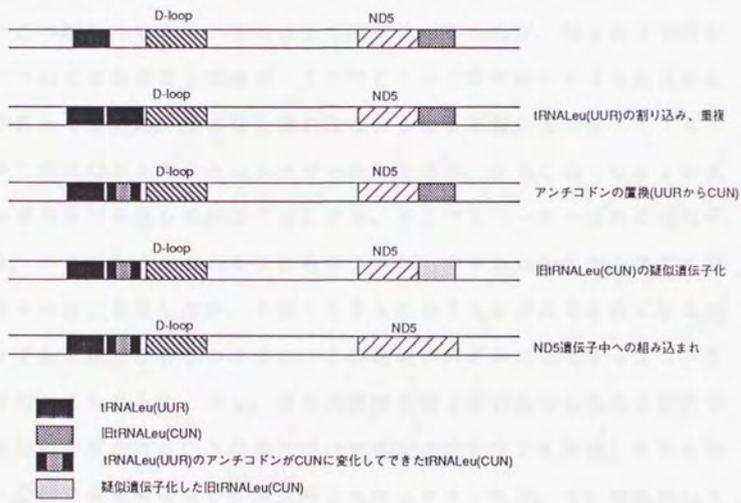


図20 tRNA^{Leu}の進化と遺伝子配置の変化

8. プロモーター配列

2. で述べたようにヒトデミトコンドリアゲノム中には約800bpのスペースー配列が存在する。その内大きなスペースーは三カ所でそれぞれ約450bp、140bp、51bpでありスペースー配列の約80%を占めている。さらに10bp~30bpのスペースーが5カ所に存在している。これらのスペースー配列を調べた所この5カ所の10bp~30bpのスペースーに、共通する”TTATAATAA”というパリンドローム配列が存在していた。(図21) このうち完全にこの配列になっているのは2カ所だけであったが、残りの3カ所の内、2カ所については末端の1塩基が、1カ所については末端から2塩基目において1塩基が異なっていた。つまり正確にはこれらの5カ所の配列は”TTATAATAATAA”の10bp中9bpまでが一致しており、またこの10bpの真ん中の6bpが完全に一致していることになる。そこでスペースー以外の遺伝子領域、約15、500bpについてこれらの10bp中9bpまでが一致する部分を探した所4カ所に存在したが、そのうち真ん中の6bpが保存されているのは1カ所だけであった。これらのように、この配列がわずかに存在するスペースー領域に選択的に見られることから、何かの機能を持っているであろうと判断できたが、この配列の置かれている位置について状況(図21)を考慮したところプロモーター配列である可能性が非常に高いものと考えられた。その根拠は以下の通りである。

まず3章で述べたようにtRNAはrRNAやmRNAをしきるように存在し、tRNAが切断されることによって、rRNAやmRNAがプロセシングされる。しかしヒトではtRNAがあまりゲノム中に散在しておらず、脊椎動物と異なり、16SrRNAやND6、CytbのmRNAのプロセシングについてはtRNA punctuationモデルでは説明できない。しかしヒトデミトコンドリアゲノムではtRNAの代わりに、この配列が16SrRNA遺伝子の5'端、ND6遺伝

子の5'端とCytb遺伝子の5'端に挟まれて存在しており、tRNAの切断によりプロセシングされる代わりに、プロモーターからの転写によって5'端が生じると考えると都合が良い。さらにこの配列はパリンδροームであるためプロモーターとして働くとなると両方向に転写する可能性が高い。その意味において見てもND6遺伝子とCytb遺伝子の間にあることは都合が良い。また同様にND3遺伝子とtRNAserUCN遺伝子、COI遺伝子とtRNAPro遺伝子も5'末端が接しているが、その間にもこの配列が存在していることなどがあげられた。

さらにウニやキヒトデのゲノム中にもこの配列は保存されており、その位置もほぼ同じ部分に存在していた。(図22)従ってこの配列はキョク皮動物に共通するプロモーター配列であることは間違いないであろう。

脊椎動物のプロモーターは良く研究されており^{69,70}、インビトロでの転写なども行われている。それはD-loopに局在しており、RNAポリメラーゼはH鎖、L鎖それぞれの鎖を鋳型にゲノムを一周する。その配列はあまり保存されておらず、ヒトとアフリカツメガエルでは全く共通性がない。ヒトのプロモーターはパリンδροーム構造になっていない。またアフリカツメガエルは不完全なパリンδροーム構造になっている。それにも関わらず、それらについて転写実験の結果、脊椎動物のプロモーターでも両方向に転写されることが解っている。

一方酵母のミトコンドリアのプロモーターは各遺伝子ごとに存在し、配列も良く保存されている⁷¹。それはキョク皮動物や脊椎動物と異なり一方向のみの転写である。

このようにRNAポリメラーゼがプロモーターから両方向に転写するのは、高等動物ミトコンドリアだけでみられる特徴であり、進化の過程で獲得した性質であろう。そして最初はキョク皮動物のようにパリンδροーム配列を認識することから始まり、その後脊椎動物のように進化したのであろう。また現在の高等生物

の祖先においては、やはり現在の下等生物のように、ゲノム中にプロモーターが散在していたであろう。その後脊椎動物とキョク皮動物のように、プロモーターの配置状況に差が生じたのは、tRNA遺伝子の遺伝子配置の差に大きく影響されたのであろう。脊椎動物ではtRNA遺伝子を他の遺伝子間へ配置させることにより、rRNAやmRNAをきちんとプロセシングするための機能を省略化できたのに対して、キョク皮動物では何等かの原因でtRNA遺伝子がクラスターを形成したため、プロモーターを脊椎動物ほど集約化できなかったのであろう。

最後に、この推定プロモーターを検証するため16S rRNAのマッピングを行ったのでそれについては10章で結果を示す。

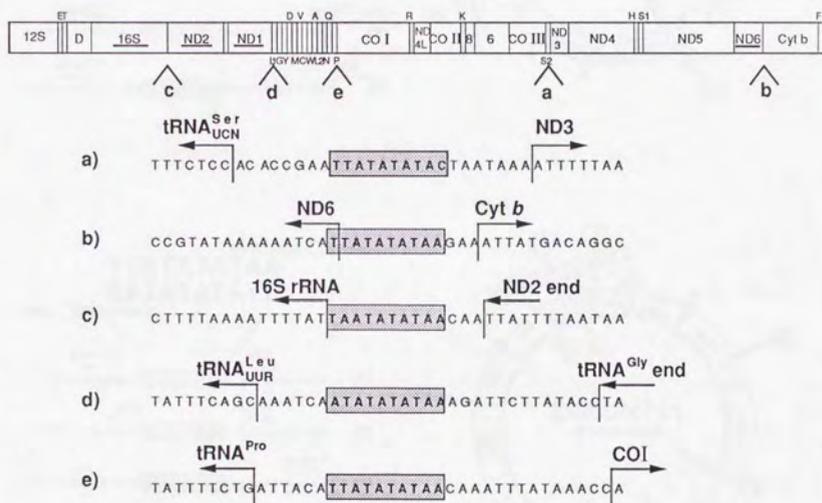
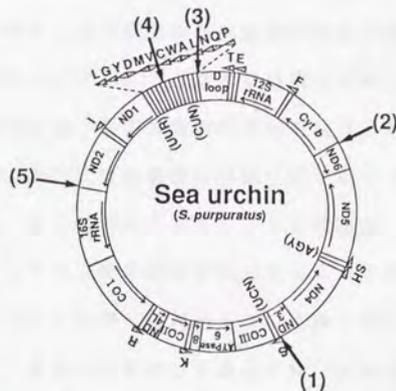
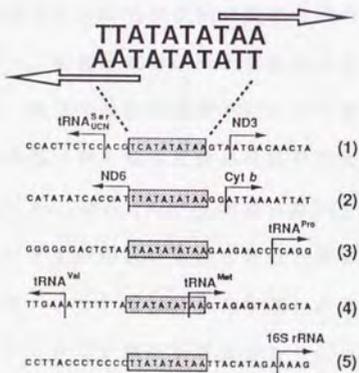
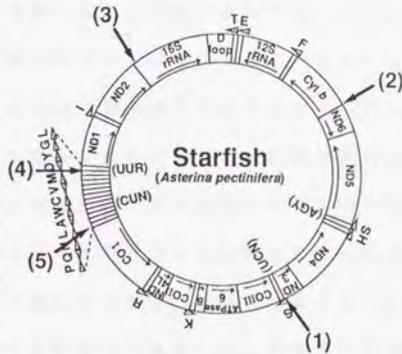
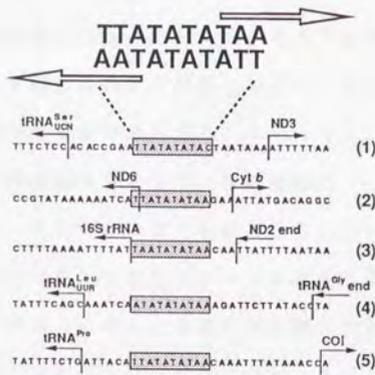


図21 推定プロモーター配列



H.T. Jacobs et al. (1988). J. Mol. Biol. 202, 185-217

図22 棘皮動物ミトコンドリアDNA中で見いだされた推定プロモーター部位

9. 鎖 (Strand) に特異的な塩基組成の偏り

従来、様々な生物のゲノムについて、最も簡単なキャラクタリゼーションは塩基組成分析により、 $G+C/A+T$ 含量を調べることであった⁷²⁻⁷⁵。GとC、AとTを組み合わせる理由としては、2本鎖DNAにおいては $A=T$ 、 $G=C$ という厳格な関係が成り立つことであった。またサイズの極めて大きいゲノムについて、相補鎖を分離して、その各鎖について塩基分析を行うことは実験上不可能であることもその一因であるだろう。DNAは2本鎖である限り $A+T \gg G+C$ 、 $A+T \ll G+C$ などのケースは存在しても、 $A=T$ 、 $G=C$ は厳密に守られることである。しかし、それぞれの鎖について議論する場合、原理的には $A \gg T$ 、 $A \ll T$ 、 $G \gg C$ 、 $G \ll C$ などのケースがあっても良い。実際には、少なくとも遺伝子のコードされている部分については遺伝情報を保持しているため塩基置換の自由度が制約され、あまり極端な塩基の偏向は無いかもしれない。

近年DNAの塩基配列の決定の技術が進歩し各遺伝子の塩基配列が決定されるに従い、塩基組成の偏向や塩基置換の傾向などについての緻密な議論が可能になった。様々な蛋白質遺伝子において解析の結果、蛋白質遺伝子の中ではコドンの3文字目に最も塩基組成の偏向が見られ、さらに塩基置換が頻繁に生じることが判明した。コドンの3文字目においては、例えば4コドンファミリーの場合、3文字目がどの塩基に置換されても指定するアミノ酸が変化せず、また2コドンファミリーについてもプリン塩基、あるいはピリミジン塩基どうしの置換の場合、指定するアミノ酸配列が変化しないなど、最も塩基置換の自由度の高い部分である。従って塩基置換が頻繁に生じる（生じた塩基置換が高い確率で定着する）のであろうが、もし塩基置換の傾向に塩基特異的な偏向がある場合はその塩基の偏向が最もコドン3文字目に現れるわけである。逆に言えばコドンの3文字目を見ると塩基置換の傾向が推定できるわけである。

ところで以上に述べた塩基置換の原因というのは、DNAポリメラーゼのキャ

ラクター、修復系酵素のキャラクター、塩基の代謝量、化学的、物理的な攻撃によるダメージなどによる、あるいは影響されるものであるが、コドンの3文字目の塩基組成に影響するもう1つの重要なファクターが池村により報告された^{76,77}。

それは細胞中のtRNAの存在量とそれに対応するコドンの使用頻度に正の相関関係が存在するというものであった。つまり存在量の多いtRNAに対応するコドンは良く使われ、逆に存在量の少ないtRNAに対応するコドンの使用頻度は少ないというものである。さらに1種類のtRNAが2種類以上のコドンに対応する場合はそのコドンとアンチコドンとの対合能がコドンの使用頻度に関係すると説明された。これらの理由としては、様々な見解があり統一されているわけではない。しかし発現量の多いtRNAに対するコドンを多く使用した場合、コドンに対するアンチコドンの会合確率が高くなるなどの理由で、効率よく蛋白質合成が進行する言うことは事実であり、見解の相違はそれらの事実のコドンの偏りに対する関連が積極的、あるいは消極的であるという所にあると思われる。対合能が問題とされる場合はブルーフリーディングなど効率の都合の良し悪しが合成効率と関連あるのであろう。

遺伝子配置の章で述べたようにヒトデではND1遺伝子、ND2遺伝子がH鎖にコードされている。この蛋白質遺伝子のコドンの使用頻度を調べたところ、ND6遺伝子を除く他の蛋白質遺伝子のものとは全く傾向が異なっていた。表9にヒトデの蛋白質遺伝子のコドンの使用頻度をH鎖、L鎖別に示した。この傾向を端的に言えば塩基置換の自由度の最も高いコドンの3文字目においてL鎖にコードされている蛋白質遺伝子ではA>T、C>Gとなっており、逆にH鎖のものについてはT>A、G>Cとなっているということである。また従来、ミトコンドリアのゲノムはA+T richであると言われていたが、L鎖についてみるとコドンの2文字目がTであるコドンボックスを除いてCの頻度がTより多かった。

H鎖についてはAとGが比較の対照となる14のコドンボックスの内、7のコドンボックスにおいてG>Aであった。さらにこの結果は全ての蛋白質遺伝子に共通にみられることであった。

表11にヒトデの各蛋白質についてのコドンの1、2、3文字目の塩基組成を示した。これで示されるように全ての蛋白質遺伝子についてL鎖にコードされるものはA>T、C>Gとなっており、H鎖のものはT>A、G>Cとなっていた。

次に図23にヒトデ、ウニ、脊椎動物、シロウジヨウバエの蛋白質遺伝子についてコドンの1+2文字目、3文字目における、塩基組成をH鎖にコードされているもの、L鎖にコードされているものを分けて示す。この図で示されるようにシロウジヨウバエを除く全ての動物について程度の差はあるものの、コドンの3文字目においてL鎖にコードされるものはA>T、C>Gとなっており、H鎖のものはT>A、G>Cとなっていた。特にウニを除いてはCとGの組成の差は大きく、ヒトデ、脊椎動物で4~15倍ほどの差になっていた。これらのことからこのコドンの使用頻度の偏りはキョク皮動物、脊椎動物に共通していることが解った。さらにこれらのL鎖においてコドンの3文字目がA>T、C>Gという傾向は脊椎動物やウニにおいて、ヒトデではH鎖にコードされているND1、ND2遺伝子についても全く同様に見られた。これらのことによりコドンの3文字目における塩基の偏りはコードされている鎖によって決まっていることと考えられる。

次にコドンの使用頻度とtRNAの存在比との関連であるが、ミトコンドリアのシステム場合22種類のtRNAで全てのコドンを読み分けており、例えば4コドンファミリーに1種類のtRNAが対応しているなどの事情から、tRNAの存在量が同義コドンの使用頻度に影響することは本質的にない。ただしコドンとアンチコドンの対合能との関連の可能性は依然存在している。しかし図23を見る限り、個別のコドンボックスについては議論できないが、総じて議論すると

L鎖とH鎖のコドン3文字目の塩基組成はほぼ完全な対照性を示しており、tRNAとの対合能などの影響はないと思われる。

それに対してショウジョウバエは全く異なった傾向を示している。そのコドンの3文字目は極端にA+T richであり、そのAとTの存在頻度はほぼ等しく45%~50%である。わずかに数%存在するCとGについては若干ヒトデと同様の偏りがみられた。しかしこのショウジョウバエについては基本的に従来から良く知られているA+T richという描像に従っていると思われる。

コドンの3文字目がにこのようなコドンの偏りは従来から脊椎動物のミトコンドリアにおいても報告されていた現象であった。それは蛋白質遺伝子のコドンの使用頻度について、その3文字目はA、Cに富み、Gが大変少ないと言うものであった。さらにND6だけはそれとは傾向が異なっていると報告されていた。しかし脊椎動物においては、ND6のコドンの使用頻度を含めて全ての蛋白質遺伝子のコドンの使用頻度をまとめて報告しており、ND6の差異については簡単にふれられているだけであった。

ヒトデにおいて得られたデータがそれらに対して有意義な点の1つは、かなり多くのH鎖にコードされた蛋白質のコドンの使用頻度のデータが得られたことである。高等動物のミトコンドリアにコードされている蛋白質の総アミノ酸残基数は約3800残基であり、その内ND6は160~170残基である。それに対してヒトデのND1、ND2はそれぞれ322、354残基であり、ND6を加えると839残基のアミノ酸がH鎖にコードされていることになる。本研究ではこれらのヒトデについて得られた豊富なデータをもとに、従来ほとんど調べられることのなかったH鎖にコードされる蛋白質におけるコドンの偏りを解析することができた。

さらに最も有意義である点はヒトデとウニにおける遺伝子の反転部分(特にND1、ND2遺伝子)を比較する事により、従来塩基の偏りをストックとしてで

しか捉えることができなかつたのに対して、フローの概念を持ち込めることであった。つまり近視眼的に（分岐したのは5億年以上前であるが、その後遺伝子反転を除いて他の要素はほとんど変わっていない。）同一起源に由来するヒトデとウニにおいて、ND 1、ND 2 遺伝子だけにコドンの3文字目の塩基組成の傾向の差が見られることは、反転によってこれらの遺伝子の塩基置換の傾向が変わつたことを明確に示している。つまりストックだけでは塩基が偏っていることしか解らないが、反転領域の比較により初めて塩基が偏るメカニズムが働いたと言うことが言える。

以上のことによりコドンの3文字目に偏りが見られることは、蛋白質合成効率などの影響でなくコードされている鎖によるであろうことが解つたわけである。さらにそれは従来から用いられていた、 $A+T/G+C$ という偏りの捉え方とは本質的に異なつたものであつた。

次にこのような鎖による塩基の偏りがtRNA遺伝子にどのように影響しているかを調べた。tRNAはコドンの3文字目のように必ず塩基置換の自由度が高いといえる所は存在しない。そこでまずそれぞれの鎖にコードされているtRNAについてトータルの塩基組成を調べたところ、各鎖間で塩基組成に大きな差は存在しなかつた。またループ部分について調べた結果も大きな差異は見られなかつた。しかしミトコンドリアに多く見られるステムのミスマッチ塩基対について調べたところ、コドンの3文字目に見られた塩基の偏りと矛盾しない結果を得た。

表12はヒトデのtRNAに見られるG-U塩基対、ミスマッチペアの個数を数えたものである。またイトマキヒトデとともに荒木によって調べられたキヒトデのデータを付記した。この表から明かなことはH鎖とL鎖のtRNAの間でG-T塩基対の存在数に著しい差が見られたことであつた。H鎖にコードされるtRNAは11種類存在するが、それらのtRNAに各1~5個のG-T塩基対が存在し、11種類の合計でイトマキヒトデには32個、キヒトデには24個のG

—T塩基対があった。それに対してL鎖にコードされるtRNAについては11種類の合計でそれぞれ7個、8個でありH鎖のものに比べて1/3~1/4となっていた。G—T塩基対のつぎに良く存在しているのはA—Cペアであるが、これについてはL鎖にコードされているtRNAのほうに若干多く見られた。この結果はtRNAについても、そのステム部分においてH鎖はG、TにL鎖はA、Cに置換されていくという塩基の偏りに影響されていることを示している。

そこでつぎに脊椎動物、ウニについても同様の解析をした。図24は各生物のtRNA一種類当たりのG—T塩基対、A—Cペアの個数を示している。この図から示されるように、ヒトデと同様にG—T塩基対、A—Cペアに差が存在していた。

以上の結果によりこれらの塩基の偏りは蛋白質遺伝子だけでなく、tRNA遺伝子にも見られることから、全ゲノム的に働いていることが示された。

このような鎖によってA>T、C>Gなどの塩基の偏りは一般には見られない現象である。その理由としては次のように考えられる。

DNAに塩基置換を引き起こす原因は様々存在する。それは例えば、複製系、修復系などDNAポリメラーゼのミス、それらを増長させるような塩基代謝の偏り、あるいは他の修復系酵素のキャラクター、化学的試薬の攻撃、酵素的攻撃、UV光、放射線等による攻撃、などである。しかしこれらは対称的な構造を持つDNAの2本鎖に対して鎖に対する特異性を持っているとは考えられない。あるいはDNAポリメラーゼには配列特異的に誤った塩基を導入する可能性があることも否定できない⁷⁸。それは例えば蛋白質遺伝子の章でふれたように、ここでのべたbiasとは別のコドンの方言を形成する要因かもしれない。さらにそれがルーチン化して、特異的な配列の自己増殖を引き起こす可能性もある。しかし蛋白質遺伝子もtRNA遺伝子も数十%以上の配列が機能的な必要上変化できない部分であり、その部分が逆にゲノムの塩基配列のある程度のランダムさを保証し

てくれる。従って、仮にDNAポリメラーゼに配列特異的に誤った塩基を導入するというキャラクターが存在したとしても、それは保証されたランダムさによって打ち消されるであろう。

これらの非対称的な塩基の偏りはミトコンドリアゲノムの非対称的な複製機構に関連があるのかもしれない。脊椎動物において研究されたミトコンドリアの複製はまずD-loop領域よりH鎖の複製から始まる。そのH鎖の複製が約60%~70%終了した時点で初めてL鎖の複製開始点が露出して、L鎖は複製開始の権利を得る⁷⁹⁻⁸¹。(図25) 電子顕微鏡による解析によると、H鎖鎖の複製が終了した時点でまだL鎖の複製が始まっていないゲノムが数十%存在していることが報告されている^{82,83}。これらのことから塩基の偏りの原因となる2通りの可能性があると思われる。

まず1番目として、ミトコンドリアDNAポリメラーゼに塩基特異的に誤った塩基を導入するキャラクターがあるとすると。実際報告されているほとんどのDNAポリメラーゼについてA、C、G、Tに平等なアフィニティが存在することはない。このようなキャラクターが存在しても、通常のDNA複製の場合は2本鎖DNA各々のどちらにも作用して、非対称性を生む原因にはならないであろう。しかしミトコンドリアゲノムの場合、H鎖の複製が終了したにもかかわらずL鎖の複製は完了しない状況で、例えばヌクレアーゼ、化学的、物理的な攻撃等でL鎖複製のテンプレートとなる親H鎖DNAが分解されることなどがあるであろう。あるいはL鎖の合成が終了する前に、次のH鎖の合成が終了するかもしれない。いずれにせよ、H鎖の合成がL鎖の合成の条件になるということから、H鎖のL鎖に対する優位が存在する事は確かである。2本鎖の完全な合成が必須である核のゲノムと異なり、ミトコンドリアゲノムは細胞中、あるいはミトコンドリア中に複数存在しているため、L鎖の合成がが必ずしも完了しなくてもそれは問題としないであろう。このような場合塩基特異的なポリメラーゼのキャラクターが

鎖に一方的に影響するかも知れない。

もう一つの可能性としては、複製過程ではL鎖はかなり長い時間1本鎖状態で存在している。1本鎖DNAは2本鎖DNAに比べて、化学的、酵素的な攻撃を受け易いという報告は様々存在している。従ってこのような攻撃による塩基置換がL鎖に一方的に蓄積し、鎖による塩基の偏りが生じたのかもしれない。

Table with 4 columns and 10 rows, containing faint text and numbers, likely a data table or experimental results.

表 1 1 ヒトデミトコンドリアゲノムにコードされた蛋白質
遺伝子の塩基組成

	1st				2nd				3rd			
	A	C	G	T	A	C	G	T	A	C	G	T
ND 1	19.0	16.9	29.8	34.4	17.2	21.2	15.3	46.3	19.0	6.4	28.8	45.7
ND 2	25.1	13.6	26.6	34.7	16.4	17.8	20.1	45.8	26.3	7.3	26.0	40.4
ND 3	20.0	27.3	22.7	30.0	16.4	25.5	10.0	48.2	33.6	38.2	4.5	23.6
ND 4	34.6	26.7	15.9	22.8	15.7	30.2	12.4	41.7	38.9	30.2	11.3	19.6
ND 4L	34.7	27.6	15.3	22.4	15.3	25.5	10.2	49.0	39.8	27.6	12.2	20.4
ND 5	35.3	20.6	17.5	26.7	20.1	25.6	12.8	41.5	39.0	25.0	6.7	29.3
ND 6	24.7	7.4	32.1	35.8	16.0	15.4	19.8	48.8	24.7	9.3	28.4	37.7
CO I	25.1	22.4	28.4	24.0	17.0	26.7	15.9	40.4	36.4	27.9	11.0	24.8
CO II	25.3	23.1	25.3	26.2	25.8	24.5	12.7	37.1	43.2	31.4	7.9	17.5
CO III	25.4	24.2	24.2	26.2	18.8	23.1	19.2	38.8	38.8	32.3	5.0	23.8
ATPase 6	29.1	25.2	18.3	27.4	16.5	27.0	11.3	45.2	39.1	36.5	6.1	18.3
ATPase 8	51.9	20.4	7.4	20.4	20.4	29.6	14.8	35.2	48.1	27.8	5.6	18.5
Cyt b	26.9	24.0	20.8	28.2	20.6	24.5	13.2	41.7	42.2	27.2	6.9	23.7

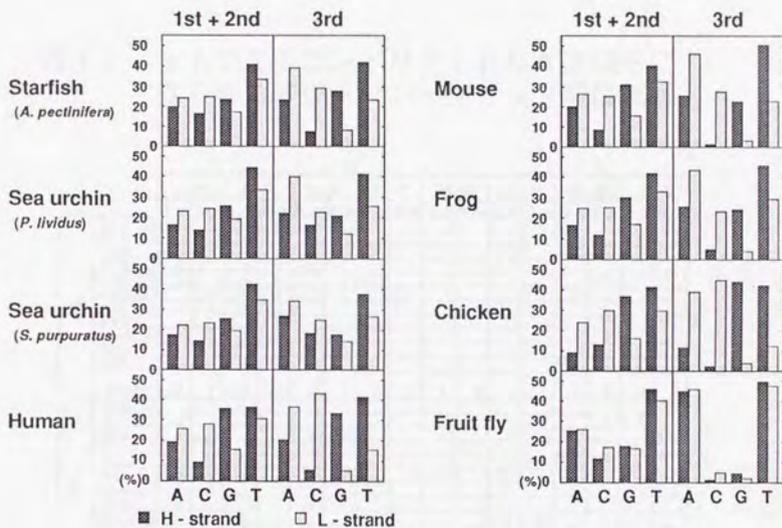


図23 様々な生物におけるミトコンドリアDNAにコードされる蛋白質遺伝子の塩基組成

表12 ヒトデミトコンドリア tRNA 遺伝子に
見られる非ワトソン-クリック型塩基対

	A-C		G-T		A-A		C-C		G-G		T-T		A-G		C-T	
	A.P	A.s	A.P	A.s	A.P	A.s	A.P	A.s	A.P	A.s	A.P	A.s	A.P	A.s	A.P	A.s
H	Asn		2	2									1			
	Cys		4	2							1	1				1
	Gly		2	2												
	Ile	1		5	3	1		1								
	Leu(CUN)	1		4	1		1					1				
	Leu(UUR)		1	2	1		1					1				
	Met	1	1	2	3											
	Pro			3	5											
	Ser(UCN)			2	1										1	
	Trp			2	2											
	Tyr			4	2									2	1	1
Total	3	2	32	24	1	2	1	0	0	0	1	3	3	2	0	2
L	Ala	1	2	1												
	Arg	1												1	1	
	Asp				1	1	1				1					
	Gln					1										
	Glu		1	1							1					
	His	1	1	1			1									
	Lys	2	1													1
	Phe	1				1	1						1			
	Ser(AGN)				2			1			1	2				
	Val			2	3						1		2			2
	Thr			2	2		1									1
Total	6	5	7	8	3	4	1	0	0	0	4	2	3	1	1	4

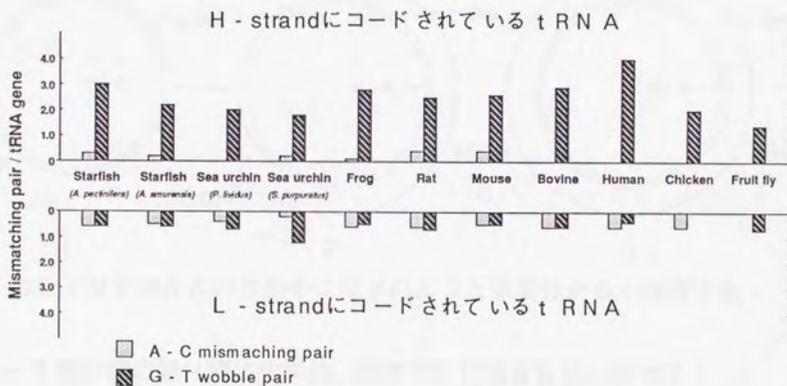
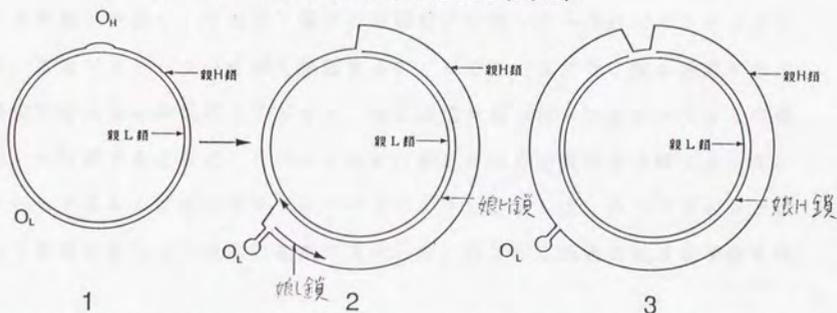


図24 tRNA遺伝子一個当たりのA-C,及びG-Tpairの存在数の平均値

なぜ鎖特異的に塩基置換が生じるのか？

ミトコンドリアDNAの特異な複製系



ミトコンドリアDNAの複製中には3のような中間体が多く存在する

1. 一本鎖状態の親H鎖は化学的、酵素的に攻撃され易いのでは？
2. 二つの娘DNAは平等な数生じるのか？

図25 なぜ鎖特異的に塩基置換が生じるのか？

10. 12S、16S rRNA 遺伝子

16S rRNA 遺伝子はその両末端には脊椎動物と異なり tRNA 遺伝子が存在しないためその末端がはっきりしない。その 5' 端側には ND2 遺伝子があるため、5' 端の限界については一応の予想はつくが、特にその 3' 末端は非コーディング領域と接しており、さらにウニなどとの相同性が低いためその末端をサイズのしか予想できないという、DNA の解析からだけではその両末端がはっきりしない遺伝子である。しかしその 5' 端については 16S rRNA 遺伝子とその 5' 側に存在する ND2 遺伝子との間にある推定プロモーター配列から転写している可能性が強く、その 5' 端がこの推定プロモーター付近にマッピングできれば、推定プロモーターを強く検証することができ、また 3' 端を決定することで遺伝子の完全な同定だけではなく、転写終結あるいはプロセッシングなどの構造について推察できるなど、RNA レベルの解析が特に必要な分子種であった。この 16S rRNA に対してプライマーエクステンション法、S1 マッピング法によって末端を数 bp の誤差の範囲で決定した。図 26 に両法の簡単な手順を示す。

図 27 は得られたオートラジオグラムである。それによって得られた結果をまとめると図 28 のようになる。

得られた 5' 端は 9019 の T に決定された。さらに 9044 の T にわずかなバンドが認められた。9019 の T は推定プロモーター配列の 2 塩基下流にあり、この T T A T A T A T A A がプロモーター配列であることを強く支持した。また 9044 の T のすぐ上流の配列は T T A T A T C A T となっており、若干プロモーターとして働いているのかもしれない。またこれに対する 9044 の T の位置は 9019 の T の位置と比較して 2 bp 上流であった。従ってプロモーター配列に対する転写開始点の位置は厳密でないのかもしれない。

3' 端は S1 プローブの 3' 端 (7722) から 228 ± 5 bp にマッピング

された。この3'端はウニなどとのサイズ的な予想からは30~40bp大きかった。この付近にターミネーションとなるようなステムアンドループ構造などを存在を調べたが、せいぜい数bp程度のステムアンドループ構造しか存在せず、またそのような数bpの構造はかなり多く存在するため、はっきりしたことは解っていない。

それに対して12SrRNA遺伝子は両末端にtRNA遺伝子が存在しているため、tRNA punctuationモデルによるプロセッシングを受けるものと考えられる。

Primer extension による
16S rRNA の 5' 末端の決定

S1 mapping 法による
3' 末端の決定

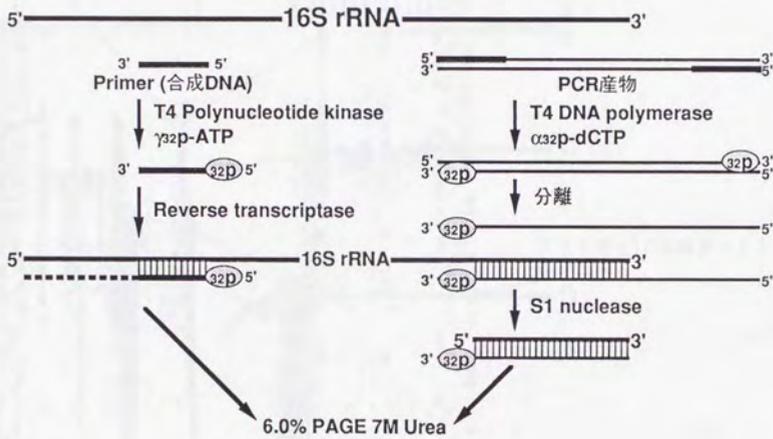
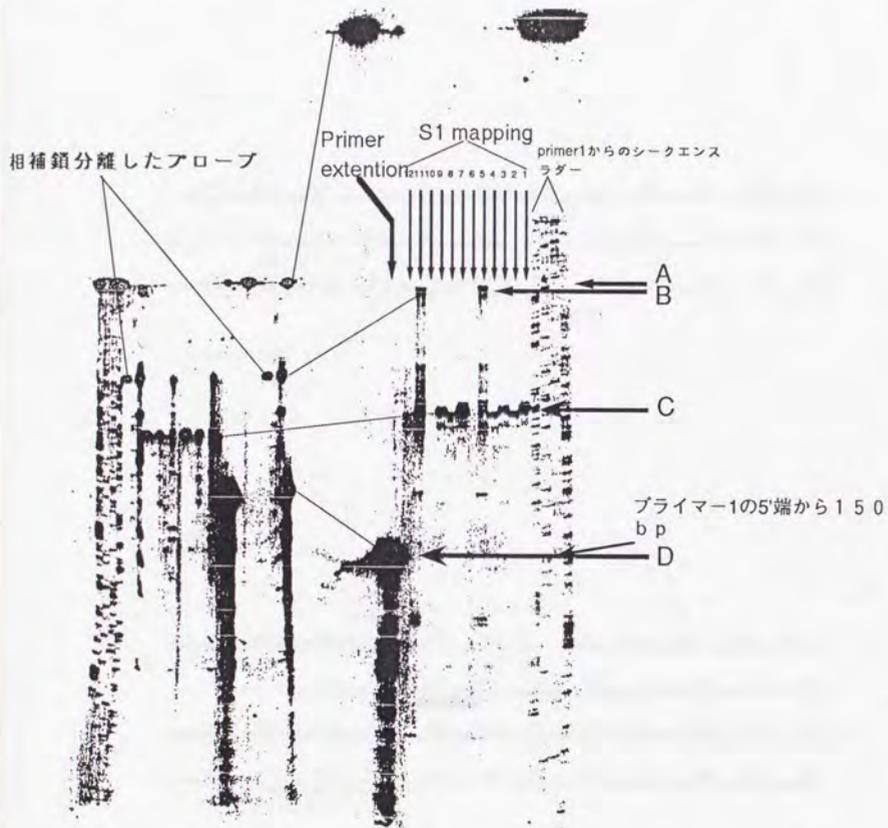
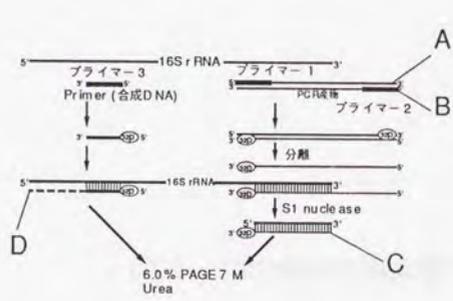


図 26

16Sr RNAのPrimer extension, S1 mappingのオートラジオグラム



S1 mapping



ハイブリダイズ温度	反応時間	酵素量	プローブ
1	46°C 30分	1 unit	相補鎖分離精製
2	46°C 30分	10 unit	相補鎖分離精製
3	46°C 60分	1 unit	相補鎖分離精製
4	46°C 60分	10 unit	相補鎖分離精製
5	46°C 30分	1 unit	相補鎖未分離未精製
6	46°C 30分	10 unit	相補鎖未分離未精製
7	48°C 30分	1 unit	相補鎖分離精製
8	48°C 30分	10 unit	相補鎖分離精製
9	48°C 60分	1 unit	相補鎖分離精製
10	48°C 60分	10 unit	相補鎖分離精製
11	48°C 30分	1 unit	相補鎖未分離未精製
12	48°C 30分	10 unit	相補鎖未分離未精製

図 27

IV. まとめ

1.

イトマキヒトデミトコンドリアの内 *Bam*HI (12, 286) - *Xba*I (14, 480) フラグメント、*Eco*RI (14, 424) - *Sal*I (15, 369) フラグメントを大腸菌ベクター (pUC18) 中にクローニングした。

2.

*Sal*I (15, 369) - *Bam*HI (1; 16, 261) フラグメントを含む断片をPCR法により増幅し、得た。

3.

イトマキヒトデミトコンドリアゲノムの約 $3/4$ (12, 411 bp) の塩基配列を新たに決定し、全塩基配列 (16, 260 bp) の決定を完成させた。

4.

その塩基配列の解析により次のことが解った。

①. ゲノムには13個の蛋白質遺伝子、22個のtRNA遺伝子、2個のrRNA遺伝子がコードされていた。その構成は脊椎動物、節足動物ミトコンドリアゲノムのものと同一であった。

②. その遺伝子配置は脊椎動物のものとはかなり異なっていた。ウニのものとの間には4.6 Kbpの遺伝子反転があったが、反転領域内、反転領域外の遺伝子配置は完全に一致していた。ヒトデの特徴は13個のtRNA遺伝子がクラスターを形成していること、2個のrRNA遺伝子が互いに相補鎖にコードされていることであった。前者は脊椎動物にみられるようなtRNA punctuation modelを補完するシステムの必要を示唆し、プロモーター配列の推定につながった。後者は脊椎動物と異なり、2個のrRNAの発現が同一regulation下でないことを示した。

③. 蛋白質遺伝子の解析により、ND1、ND5 遺伝子についてはGTGが開始コドンであると推定された。またATPase8とATPase6 遺伝子には16塩基長のオーバーラップが存在した。Cytb、COI I 遺伝子には終始コドンが存在せず、mRNA前駆体のポリアデニル化が終始コドン生成の条件となっていた。キョク皮動物の蛋白質の進化速度は脊椎動物のものより遅いと考えられた。

④. 脊椎動物及び節足動物のミトコンドリアゲノムの蛋白質遺伝子との比較により、ATA-Ile、AAA-Asn、AGA、AGG-Serという異常遺伝暗号が、ヒトデ（キョク皮動物）ミトコンドリアに使用されていることを支持する統計的データを得た。

⑤. tRNA 遺伝子を解析したところ、リシンのtRNAについては、他の生物とのアンチコドン部分の相違により、AAAはリシンの暗号ではなく、AAGだけがリシンの暗号に相当する事が確認できたが、それ以外のヒトデに特異的な異常遺伝暗号については、tRNA 遺伝子のアンチコドン部分について、他の生物種との差がなかった。GG-TTCRAの共通塩基の保存されているものは7種類あり、なかでもセリンとロイシンのそれぞれ2種類のtRNAについてはこれらの共通塩基がキョク皮動物間でよく保存されていた。またロイシン(CUN)のtRNAはウニと同様ロイシン(UUR)のtRNAからの派生物であることが解った。

⑥. TTATATATAAからなる配列が非コーディング領域に5カ所存在し、プロモーター配列と推定した。この配列はウニにも保存されており、キョク皮動物共通のプロモーター配列であると考えられた。これによって、tRNA punctuation modelによって説明できない16SrRNA、ND6 mRNA、Cytb mRNAの5'末端プロセッシングについて、説明がついた。

⑦. ヒトデとウニの反転領域にコードされているND1、ND2 遺伝子のコード

ンの使用頻度をヒトデの他の蛋白質遺伝子及びウニのND1, ND2 遺伝子と比べることによって、これらのゲノムには鎖によって塩基組成を偏らせる機構があることが解った。この機構はしRNA 遺伝子にも作用しており、全ゲノム的に働いていた。

5.

遺伝子からはその末端が同定できない16S rRNAの両末端を決定した。決定された5'末端からTTATATATAA配列がプロモーター配列であることがほぼ確定した。

V. 参考文献

- (1) Nirenberg, M.W. and Matthaei, J.H., Proc. Natl. Acad. Sci. USA, 47, 1588 (1961)
- (2) Leder, P. and Nirenberg, M.W., Proc. Natl. Acad. Sci. USA, 52, 420 (1964)
- (3) Sprinzl, M and Gauss, D.H. Nucl. Acids Res., Supplement 12. r1-r131
- (4) 大島、井上、湯浅、渡辺編：現代の遺伝学，2，遺伝情報，(1974)，朝倉書店
- (5) Majar, J., Biochem. Biophys. Acta, 38. 150 (1960)
- (6) Brost, P. and Grivell, L.A., FEBS Lett. 12, 129 (1971)
- (7) 藤井義明, et al., 細胞生物学, 2. 156-180 (1981)
- (8) Rendi, R., Exp. Cell Res., 17. 585 (1959)
- (9) Sanger, F., Science, 214, 1205-1210 (1981)
- (10) Clary, D.D. and Wolstenholme, D.R., J. Mol. Evol., 22, 252-271, (1985)
- (11) Batuecas, B., et al., Nucl. Acids Res., 16, 6515-6530, (1988)
- (12) Anderson, S., et al., Nature, 290, 457-464, (1981)
- (13) Anderson, S., de Bruijn, M.H.L., Coulson, A.R., Eperon, I.C., Sanger, F. and Young, I.G., J. Mol. Biol., 156, 683-717, (1982)
- (14) Bibb, M.J., Van Etten, R.A., Wright, C.T., Walberg, M.W. and Clayton, D.A., Cell, 26, 167-186, (1981)
- (15) Roe, B.A., Ma, D.-P., Wilson, R.K. and Wong, J.F.-H., J. Biol. Chem., 260, 9759-9774, (1985)
- (16) Fujii, H., et al., J. Biochem., 103(3), 474-481, (1988)
- (17) Gadaleta, G., Pepe, G., De Candia, G., QuagliarIELLO, C., Sbisà, E. and Caccone, C., J. Mol. Evol., 28, 497-516, (1989)

- (18) Desjardins, P. and Morais, R., *J. Mol. Biol.*, 212, 599-634, (1990)
- (19) 米山芳雄、日大医誌、54巻、429-440、(1986)
- (20) Himeno, H., Masaki, H., Kawai, T., Ohta, T., Kumagai, I., Miura, K. and Watanabe, K., *Gene*, 56, 219-230, (1987)
- (21) Jacobs, H.T., Asakawa, S., Araki, T., Miura, K., Smith, M.J. and Watanabe, K., *Curr. Genet.*, 15, 193-206
- (22) Wolstenholme, D.R., *Proc. Natl. Acad. Sci. USA*, 84, 1324-1328
- (23) Ma, D.P., et al., *Nucl. Acids Res.*, 16, 11373- (1988)
- (24) Goddard, J.M., et al., *J. Mol. Biol.*, 97, 593-609, (1975)
- (25) Benne, R., et al., *Nucl. Acids Res.*, 6925-6941, (1983)
- (26) Lambert, A.M., et al., *Nucl. Acids Res.*, 12, 7327-7343, (1984)
- (27) Wallence, D.C., *Microviol Rev.*, 46, 208-240, (1982)
- (28) Jacobs, H.T., Elliott, D.J., Math, V.B. and Farquharson, A., *J. Mol. Biol.*, 202, 185-217, (1988)
- (29) Cantatore, P., Roberti, M., Rainaldi, G., Gadaleta, M.N. and Saccone, C., *J. Biol. Chem.*, 264, 10965-10975, (1989)
- (30) Barrell, B.G., Bankier, A.T. and Drouin, J., *Nature*, 282, 189-194 (1979)
- (31) Marechal-Drouard, L., Weil, J-H. and Guillemaut, P., *Nucl. Acids Res.*, 16, 4777-4788 (1988)
- (32) Roe, B.A., Wong, J.F.H., Chen, E.Y. and Armstrong, P.W., in *Recombinant DNA, Proceeding of the 3rd Cleveland Symposium of Macromolecules*, ed. Walton, S.G., pp. 167-176, Elsevier Publishing Co., Amsterdam
- (33) Crick, F.H.C., *J. Mol. Biol.*, 19, 548-555 (1966)

- (34) Sakore, T.D. and Sobell, H.M., J. Mol. Biol., 46, 77-78 (1969)
- (35) Soll, D., et al., J. Mol. Biol., 29, 97-112 (1967)
- (36) Heckman, J.E., Sarnoff, J., Alzner-Deweerd, B., Yin, A. and Rajibha
ndary, U.L., Pro. Natl. Acad. Sci. USA, 77, 3159-3163 (1980)
- (37) Barrell, B.G., et al., Pro. Natl. Acad. Sci. USA, 77, 3164-3166 (19
80)
- (38) Lagerkvist, U., Pro. Natl. Acad. Sci. USA, 75, 1759-1762 (1978)
- (39) Ojala, D., Montoya, J. and Attardi, G., Nature, 290, 470-474 (1981)
- (40) Baldwin, A.N. and Berg, P., J. Biol. Chem., 241, 839-845 (1967)
- (41) Kumazawa, Y., Yokogawa, T., Hasegawa, E., Miura, K. and Watanabe, K.,
J. Biol. Chem., 264, 13005-13011 (1989)
- (42) Messing, J., Method in Enzymology, 101, 20-78 (1983)
- (43) Feinberg, A.P. and Vogelstein, B., Anal. Biochem, 32, 6-13 (1983)
- (44) Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erl
ich, H.A. and Arnheim, N., Science, 230, 1350-1354 (1985)
- (45) Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R.,
Horn, G.T., Mullis, K.B. and Erlich, H.A., Science, 239, 487-491 (198
8)
- (46) Henikoff, S., Gene, 28, 351-359 (1984)
- (47) Yanisch-Perron, C., Vieira, J. and Messing, J., Gene, 33, 103-119 (1
985)
- (48) Mizusawa, S., Nishimura, S. and Seela, F., Nucl. Acids Res., 14, 131
9-1324 (1986)
- (49) Vogt, V.M., Eur. J. Biochem., 33, 199-200 (1973)
- (50) Ghosh, P. K., Reddy, V.B., Piatak, M., Lebowitz, P. and Weissman, S.

- M., Method in Enzymology, 65, 580-595 (1980)
- (51) Lee, D.C., and Roeder, R.G., Mol. Cell Biol., 1, 635-651 (1981)
- (52) Aloni, Y. and Attardi, G., J. Mol. Biol., 55, 251-270, (1971)
- (53) Brown, M.W., Ann. NY. Acad. Sci., 361, 119-134, (1981)
- (54) Deley, J., J. Bacteriol., 101, 708- (1962)
- (55) Schildkraut, C.L., et al., J. Mol. Biol., 4, 430- (1962)
- (56) Wells, R.D., et al., J. Mol. Biol., 54, 465 (1979)
- (57) Wong, J.F.H., Ma, D.P., Wilson, R.K., and Roe, B.A., Nucl. Acids Res., 11, 4977-4995 (1983)
- (58) Chang, D.D., Fisher, R.P. and Clayton, D., Biochem. Biophys. Acta., 909, 85-91 (1987)
- (59) Streisinger, G., Okada, Y., Emrich, J., Newton, J., Tsugita, A., Terzaghi, E and Inoue, M., Cold Spring Harbor Symp. Quant. Biol., 31, 77-84 (1966)
- (60) Clark, D.J., Bilanchone, V.W., Haywood, L.J., Dildine, S.L and Sandmeyer, S.B., J. Biol. Chem., 263, 1413-1423 (1988)
- (61) Marschalec, R., Brechner, T., Amon-Bohm, E. and Dingermann, T., Science, 244, 1493-1496 (1989)
- (62) Jacobs, H.T., Elliott, D.J., Math, V.B and Farquharson, A., J. Mol. Biol., 202, 185-217 (1988)
- (63) Saccone, C., Attimonelli, M and Sbisà, E., J. Mol. Evol., 26, 205-211 (1987)
- (64) Brown, G.G., Gadaleta, G., Pepe, G., Saccone, C. and Sbisà, E., J. Mol. Biol., 192, 503-511 (1986)
- (65) Ojala, D., Merkel, C., Gelfand, R. and Attardi, G., Cell, 22, 393-4

03, (1980)

- (66) Desjardins, P., L'Abbe, D., Lang, B.F., and Morais, R., J. Mol. Biol., 207, 625-629 (1989)
- (67) Kim, S.H., Suddath, F.L., Quigley, G.J., McPherson, A., Sussman, J.L., Wang, H.J., Seeman, N.C. and Rich, A., Science, 185, 435-440 (1974)
- (68) Cantatore, P., Gadaleta, M.N., Roberti, M., Saccone, C., and Wilson, A.C., Nature, 329, 853-855 (1987)
- (69) Chang, D.D., Hixson, J.E. and Clayton, D.A., Mol. Cell Biol., 6, 294-301 (1986)
- (70) Bogenhagen, D.F. and Yoza, B.K., Mol. Cell Biol., 6, 2543-2550 (1986)
- (71) Schwegen, R.J., et al., Mitochondria, 24-45 (1983)
- (72) Sueoka, N., Proc. Natl. Acad. Sci. USA, 48, 582-592 (1962)
- (73) Jukes, T.H. and Bhushan, V., J. Mol. Evol., 24, 39-44 (1986)
- (74) Muto, A. and Osawa, S., Proc. Natl. Acad. Sci. USA, 84, 166-169 (1987)
- (75) Ohama, T., Muto, A. and Osawa, S., Nucl. Acids Res., 18, 1565-1569
- (76) Ikemura, T., J. Mol. Biol., 151, 389-409 (1981)
- (77) Ikemura, T., J. Mol. Biol., 158, 573-597 (1982)
- (78) Kunkel, T.A., J. Biol. Chem., 260, 12866-12874 (1985)
- (79) Bolden, A., Noy, G.P. and Weissbach, A., J. Biol. Chem., 252, 3351-3356
- (80) Bogenhagen, D. and Clayton, D.A., J. Mol. Biol., 119, 49-68, (1978)
- (81) Martens, P.A. and Clayton, D.A., J. Mol. Biol., 135, 327-351, (1979)

- (82) Rpberson, D.L., Kasamatsu, H. and Vinograd, J., 69, 737-741, (1972)
- (83) Goddard, J.M. and Wolsteinholme, D.R, Nucl. Acids Res., 8, 741-757,
(1980)
- (84) Asakawa, S., Kumazawa, Y., Araki, T., Himeno, H., Miura, K. and Wat
anabe, K., J. Mol. Evol., in press
- (85) Garey, J.R., and Wolstenholme, D.R., J. Mol. Evol., 28, 374, (1989)

VI. 謝辞

本研究を行う機会を与えていただき、絶えず暖かい御指導を賜りました、東京大学工学部工業化学科の三浦謙一郎教授に深く感謝致します。

生命科学の初歩から御教示賜り、教育効果の著しく低い私を見放すことなく叱咤激励、御指導して下さいました、東京工業大学生命理工学部生体機構学科の渡辺公綱教授に大変感謝致します。

東京大学工学部の熊谷泉助教授には、様々な貴重な御指導、御教示をしていただき大変感謝致します。

東京大学工学部助手の平尾一郎博士には、厳しい御指導とともに良く学び良く遊べを伝授していただき大変感謝致します。

東京工業大学の西川一八助教授にはRNAについて様々な御教示をして頂き大変感謝いたします。

東京工業大学助手の上田卓也博士にはミトコンドリアについて様々な御教示をして頂き大変感謝致します。

宇宙科学研究所助手の姫野俵太博士には本研究におけるあらゆる先導をして頂いただけでなく、様々な御指導、御援助をして頂き大変感謝致します。

東京大学農学部の正木春彦助教授にはdeletion mutantの作成方法を御教授して頂き大変感謝致します。

東京大学医科学研究所の水本教授にはキャッピング酵素を頂きまして大変感謝致します。

東京大学工学部の氏平祐輔教授には、本学において最初に研究の取り組みかたを御教示して頂き大変感謝いたします。

東京大学工学部の熊沢慶伯博士（現カルフォルニア大学）には、研究の初歩から御指導いただき、様々な御教示、御議論を賜り大変感謝致します。

東京大学工学部の堀弘幸先生（現三菱化成生命科学研究所）には研究だけでなく、様々なことについて御指導、御援助して頂き大変感謝いたします。

東京大学工学部の小畑充生博士（現トヨタ自動車）には、寛大なる御指導を賜り大変感謝致します。

東京大学工学部助手の小島修一博士には様々な御面倒を見ていただき大変感謝いたします。

東京大学工学部の荒木武義氏（現旭化成）には実験、データ整理、書類管理、力仕事などあらゆることを御援助していただき、また互いに良い議論ができたことを大変感謝致します。

東京工業大学生命理工学部（元三浦研）の横川隆志氏には実験、データ整理、製図、宿泊などあらゆることを御援助していただき、せっぱ詰まった時に活路を切り開いて頂き大変感謝いたします。

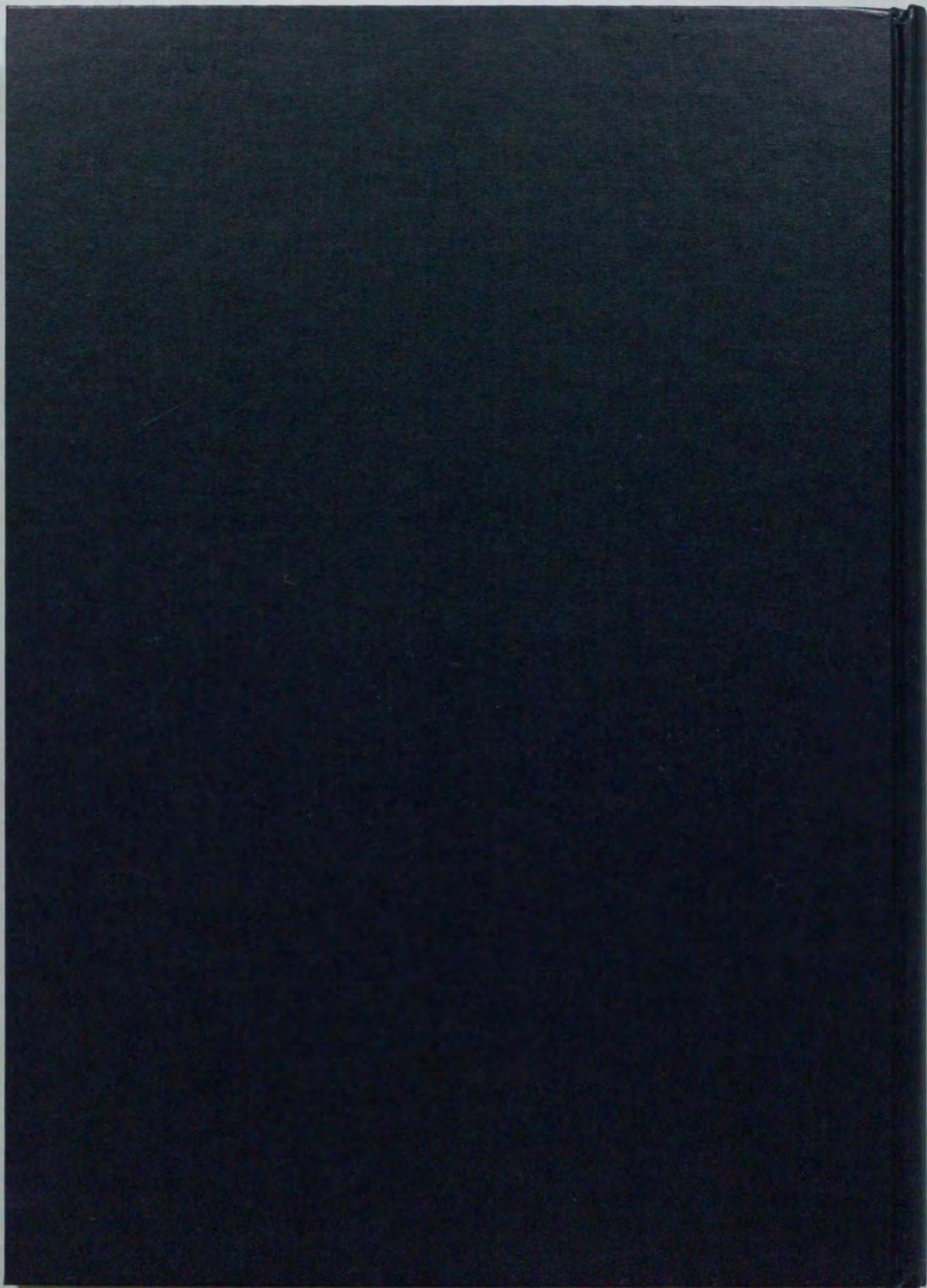
東京大学工学部の北村昌也氏には、実験、データ整理、書類管理、入浴などあらゆることを御援助していただき、また耳の痛いお叱りをいただいたことを大変感謝致します。

また実験など様々なことにおいて怠惰な私を御援助していただいた、奈良岡哲志氏、田口精一氏、三好洋氏、小山淳氏をはじめとする先輩の皆さん、お互いに良き刺激を与え合った玉木英輔氏、宇都宮龍馬氏、三沢悟氏、関誠氏、武田茂樹氏、駒山素明氏らの同期の皆さん、研究やそれ以外様々なことで大変御世話になった、古久保進氏、大久保光晴氏、嶋山隆氏、杉山亨氏を初めとする後輩の皆さん、男性の多い研究室を華やかにし、様々なことで御援助いただいた舛井裕子氏、並沢輝美氏を初めとする日本女子大学、東京理科大学、東京都立大学、北里大学などの皆さん、鎌田美奈氏、吉田恵子氏を初めとする秘書の皆さんに深く感謝いたします。おもしろい研究生活を送らせてくれた高仲康弘氏、森谷二郎氏に深く感謝します。

日本育英会には経済的な援助をいただき深く感謝いたします。

私を励ましてくれた多くの友人に深く感謝いたします。

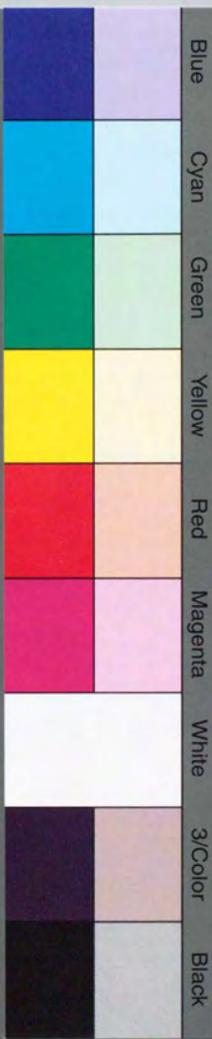
最後に暖かく私の成長を見守ってくれた、両親、家族に深く感謝します。



inches
0 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

© Kodak, 2007 TM, Kodak



Kodak Gray Scale



© Kodak, 2007 TM, Kodak

A 1 2 3 4 5 6 **M** 8 9 10 11 12 13 14 15 **B** 17 18 19

