

機関番号：16201

研究種目：基盤研究(C)

研究期間：2008～2010

課題番号：20570222

研究課題名(和文) マイクロRNA遺伝子族の分子進化

研究課題名(英文) Molecular evolution of microRNA gene families

研究代表者

岩間 久和 (IWAMA HISAKAZU)

香川大学・総合生命科学研究センター・准教授

研究者番号：20398035

研究成果の概要(和文)：マイクロRNA遺伝子族は、進化の過程で、そのターゲットとなる遺伝子制御の冗長性を高める一方、顕著にターゲットとなる遺伝子数が減少したことを明らかにした。マイクロRNA遺伝子族は、既に確立した転写因子による遺伝子制御網を攪乱しないという主に消極的な役割において結果として遺伝子制御網の冗長性が増加したと考えられ、この遺伝子制御網の冗長性の増加は、頑健性、微細調節などの適応的役割に必ずしも帰着しないことが示唆された。

研究成果の概要(英文)：MicroRNA (miRNA) gene families are shown to have evolved to alter their network configurations so that they increase the network redundancy and, in contrast, reduce their target gene repertoire. These findings suggest that a large part of the redundancy-adding role of miRNA gene families is attributable to non-adaptive processes of avoiding conflicts with the established transcription factor regulatory networks. This further indicates that the high redundancy in the miRNA regulatory networks is not necessarily a hallmark of adaptive roles such as robustness, failsafe or fine-tuning.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	700,000	210,000	910,000
2009年度	500,000	150,000	650,000
2010年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	2,200,000	660,000	2,860,000

研究分野：生物学

科研費の分科・細目：生物科学・進化生物学

キーワード：マイクロRNA、遺伝子族、ターゲット配列、遺伝子制御網、オルソログ、冗長性、協調的遺伝子制御、フィードフォワード回路

## 1. 研究開始当初の背景

成熟マイクロRNA (mature microRNA) はタンパクをコードしない約22塩基の1本鎖RNA分子で、メッセンジャーRNA (mRNA) の3' UTR配列との相補的結合により抑制的な転写後制御をする。マイクロRNAは、脊椎動物において全遺伝子の2～4%を占める最大の多重

遺伝子族の一つと考えられ、同時に最大の遺伝子制御因子群を構成する。ヒト・ゲノムに約1000のマイクロRNA遺伝子がコードされていると推測されていたが2008年当時miRBase (Release 10) ではヒトで533エントリーが登録されていた。(2011年4月現在Release 17では1424エントリー。)

一つのマイクロ RNA は約 100~200 の遺伝子を制御することが、種間で保存されたマイクロ RNA ターゲットの解析から推定され、さらに、保存されていないマイクロ RNA ターゲットはその 10 倍にのぼると推測される。これらの知見から、マイクロ RNA 遺伝子族は密な遺伝子制御網を構成すると想定されていた。

このように多くの遺伝子群によって構成されるマイクロ RNA ではあるが、遺伝子族としての進化については、究明が進んでおらず、遺伝子族として捉えた時、新規機能の獲得 (gain of function) の機序などの知見はほとんど示されていない。また、どのような機能が、その進化的変遷の対象となっているかが解明されるべき重要な点である。

上述の点を考える上で、マイクロ RNA 遺伝子族と同様に、多数のメンバー遺伝子を擁し、遺伝子制御にかかわる転写因子の遺伝子族との比較検討を行うことは、マイクロ RNA 遺伝子族の進化の特質を明らかにする上で有用と考えられた。

## 2. 研究の目的

マイクロ RNA 遺伝子を遺伝子族と捉え、分子進化の過程の特質を解明する。特に、マイクロ RNA 遺伝子族がターゲットとする 3' UTR 中の配列を網羅的に予測し、それらの進化過程 (つまりは、マイクロ RNA による遺伝子制御網の進化過程を意味する) の特質と機能の関係を明らかにする。

## 3. 研究の方法

### (1) 研究材料

#### ① ゲノム配列

NCBI (ftp.ncbi.nih.gov/genome/) からリファレンスゲノム配列及びアノテーション情報を取得。

#### ② マイクロ RNA 遺伝子族の情報

miRBase Release 10 を www.mirbase.org から取得

#### ③ Gene Ontology 情報

NCBI

ftp.ncbi.nlm.nih.gov/gene/ DATA から gene2go ファイルを取得。

#### ④ マイクロ RNA 遺伝子ターゲットサイトの予測

PITA プログラムを

http://genie.weizmann.ac.il/pubs/mir07 より取得。

### (2) 種間保存情報に依拠しないマイクロ RNA ターゲットサイトの予測

ゲノム配列情報及びアノテーション情報を用いて 3' UTR 配列を全タンパク・コード遺伝子について抽出し、PITA プログラムにより網羅的に予測した。予測条件

は、シードマッチはミスマッチ許さない 8-mer、または 7-mer ; wobble マッチを許さず ;  $\Delta \Delta G < 0$ 、とした。

### (3) 種間保存情報を加味したマイクロ RNA 遺伝子ターゲットサイトの予測

#### ① オルソログ遺伝子の同定

翻訳開始点上流 8kb を含めて他の遺伝子とのオーバーラップのないタンパクコード遺伝子をアノテーション情報より網羅的に抜き出し、ヒトとマウスゲノム配列から、それらの 3' UTR 配列と上流 8kb の配列を抽出。次に、それらの中で Official Gene Symbol が一致するものをオルソログとした。この方法は配列の相同性のみによるオルソログ同定と比べて、実験的検証のある確実なオルソログのみ選択できる利点がある。

#### ② 種間保存された 3' UTR 配列の同定

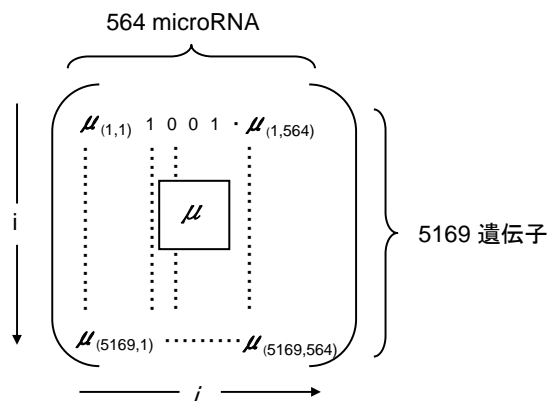
本研究のために、3' UTR 配列のアライメントに適したプログラム ReAlignerVR を作成し、3' UTR 配列中の種間保存配列ブロックを同定した。

#### ③ 種間保存マイクロ RNA ターゲット配列

3' UTR 配列中、種間で 100% 保存された予測マイクロ RNA ターゲット配列を保存されたターゲット配列とした。

### (4) マイクロ RNA 遺伝子族による遺伝子制御網の行列 ( $\mu$ ) での表示

種間保存されたマイクロ RNA ターゲット配列の網羅的情報を以下の行列  $\mu$  にまとめた。行列  $\mu$  において、 $\mu_{ij}$  はマイクロ RNA  $j$  が遺伝子  $i$  を制御の対象としている時  $\mu_{ij} = 1$ 、そうでないときは、 $\mu_{ij} = 0$  と表わされる。



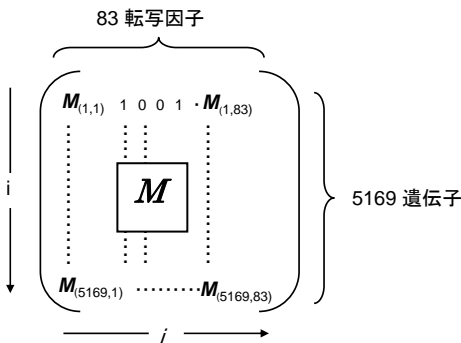
### (5) マイクロ RNA 遺伝子制御網との比較のための転写因子による遺伝子制御網の表示

#### ① 転写因子による保存された遺伝子制御網の同定

マイクロ RNA の転写制御網の同定に用いたものと同じオルソログ遺伝子セ

ットについて、それらの上流 8kb の配列中の転写因子結合配列を TRANSFAC MATCH プログラムで網羅的に予測した。ReAlignerV で網羅的に上流配列の保存領域を同定し、その保存領域内で、同方向かつ identity が 90% 以上の転写因子結合配列を同定し、以下用いた。予測に用いた転写因子は、オルソログ遺伝子セットに含まれ、かつ、十分な実験的検証のある 83 種の転写因子である。

- ② 転写因子による遺伝子制御網の表示  
 マイクロ RNA 遺伝子族と同様に、83 種の転写因子について、転写因子とそれが制御するオルソログ遺伝子の制御網を行列  $M$  を用いて次図のように表示する。

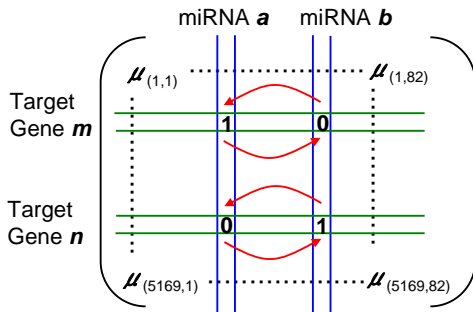


- (6) 結合度を保った行列のランダム化

- ① 結合度を保ったランダム化

(degree-preserving randomization ; 以降、単にランダム化と記載)

網羅的に予測されたマイクロ RNA 遺伝子の遺伝子制御網と、比較対象のために予測された転写因子による遺伝子制御網のランダム化を行う。この際、行列要素の結合度はランダム化の過程を通じて等しくなるように行う。結合度を保つとは、一つの遺伝子に着目して、その遺伝子を制御する遺伝子の数を一定にし、かつ、その遺伝子が制御する遺伝子の数も同じに保つことを意味する。



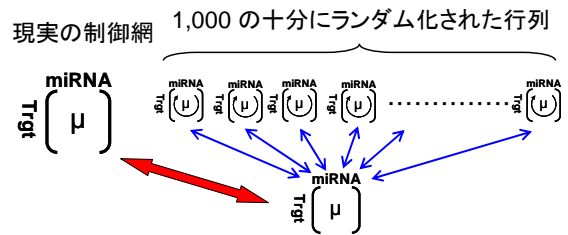
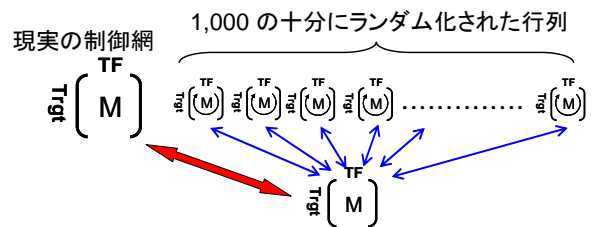
具体的には、上図のごとく、ランダムに選んだ一組の列 a と b ( $a \neq b$ )、と一組の行 m と n ( $m \neq n$ ) について、 $\mu_{ma} = 1 \cap \mu_{nb} = 1 \cap \mu_{na} = 0 \cap \mu_{mb} = 0$  となるときのみ、 $\mu_{ma} = 0$ ,

$\mu_{nb} = 0$ ,  $\mu_{na} = 1$ ,  $\mu_{mb} = 1$  と変換することにより、結合度を保ったランダム化を行う。

- ② 現実の遺伝子制御網とランダム化された制御網との統計学的比較

上述のランダム化を現実の制御網を表す行列  $M$  と  $\mu$  について、50 万ステップ行うことで十分にランダム化された行列を生成する。さらに、そのようなランダム行列を各々 1000 個、コンピュータ上で生成する。模式的に示すと下図のようになる。その 1000 個の行列と現実の行列を比較することにより、現実の遺伝子制御網に統計学的に有意に頻出するマイクロ RNA が関与する遺伝子制御の回路を見いだす。模式的に手続きを下図に示す。

図中、 $\text{Target} \begin{pmatrix} \text{TF} \\ (M) \end{pmatrix}$   $\text{Target} \begin{pmatrix} \text{miRNA} \\ (\mu) \end{pmatrix}$  は、それぞれの行列のランダム化を意味している。



- (7) マイクロ RNA のかかわる遺伝子制御網のモチーフの検出

現実の遺伝子制御網 (行列) とランダム化した行列中に観察されるマイクロ RNA がかかわるモチーフをフィードフォワード回路に焦点を当て、網羅的にプログラムにより検出し、統計的解析に供する。

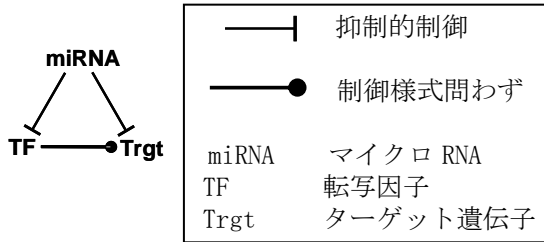
- (8) 部分的ランダム化による、マイクロ RNA 遺伝子族の役割の探索

ランダム化する際に、マイクロ RNA 遺伝子族による制御網の行列のみをランダム化し、転写因子によるものは現実の行列に固定する、また、逆にマイクロ RNA 遺伝子族の制御網の行列は現実のものに固定し、転写因子による制御網の行列のみをランダム化する。この部分的ランダム化により、マイクロ RNA 遺伝子族の制御網の成り立ちの特質を明確化する。

#### 4. 研究成果

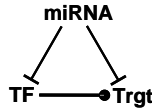
(1) マイクロ RNA 遺伝子族のかかわる遺伝子制御網（特にフィードフォワード回路：Feedforward Circuit (FFC)）について、進化過程における自然選択の影響

具体的には、定量的に次の3つの指標によって評価した。ここでは、マイクロ RNA 遺伝子と転写因子のいずれもがかかわる回路を対象としているので、composite FFC (cFFC) と表記する。  
模式的に示すと下図のようになる。



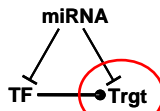
① フィードフォワード回路 (cFFC) の数

右図の cFFC の出現数をカウントし現実の制御網とランダム化された制御網で比較する。



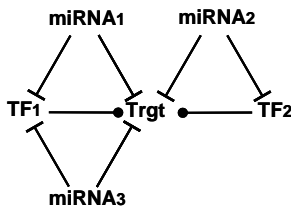
② cFFC によって制御される遺伝子数

右図の回路で制御されるターゲット遺伝子 (Trgt) の実数 (重複して制御される場合もその都度カウントする。)



③ cFFC の冗長性

次図のような例では、ターゲット遺伝子は3つの cFFC によって制御されているので、冗長性は3となる。



①②③の指標の結果を右欄上の表にまとめる。

表より、マイクロ RNA 遺伝子族は遺伝子制御網に対して cFFC のターゲットとなる遺伝子の数を減少させるように、また同時に、その冗長性を増すように selection の影響ははたらいっていることがわかる。一方、転写因子の遺伝子制御網においては、影響はほとんどないことがわかる。

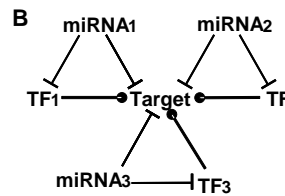
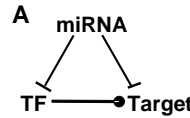


	miRNA の制御網のみランダム化		転写因子の制御網のみランダム化		いずれもランダム化	
	Z	P	Z	P	Z	P
cFFC の数	0.28	0.390	-0.64	0.262	0.09	0.463
cFFC が制御する遺伝子数	-13.94	$1.8 \times 10^{-44}$	-1.13	0.128	-13.44	$1.8 \times 10^{-41}$
cFFC の冗長性	4.47	$3.9 \times 10^{-6}$	0.30	0.383	4.15	$1.5 \times 10^{-5}$

ここで Z 値は、生成した 1000 のランダム行列と比較して、現実の制御網で多く出現していれば正の値、少なければ負の値となる。その P 値は標準正規分布による。

(2) 冗長性はマイクロ RNA の関与による

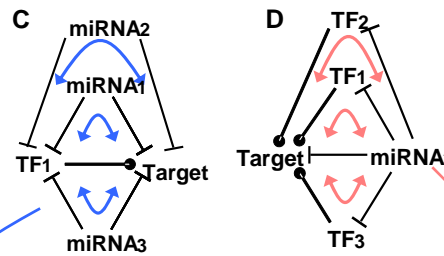
確実にマイクロ RNA によって冗長性が生じていることを確かめるために、次にしめす、制御網のトポロジーについて考える。



単位 cFFC (上図 A) の組み合わせり方により、左図 B のように、1 つのターゲット遺伝子が、同数のマイクロ RNA と転写因子とによって冗長性が生じている場合がある。

一方、下図 C, D の場合は状況が異なる。

つまり、図 C の場合はマイクロ RNA 遺伝子により冗長性が生じ、図 D の場合は、転写因子によって冗長性が生じている。この違いに着目して、冗長性の由来を分別する。



マイクロ RNA 由来の冗長性

$$= \frac{\text{cFFC の総数}}{\text{miRNA-target connection の総数}}$$

転写因子由来の冗長性

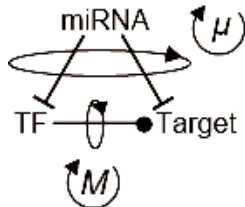
$$= \frac{\text{cFFC の総数}}{\text{TF-target connection の総数}}$$

上記2式を用いて、それぞれの由来による冗長性を定量化できる。この方法による結果を下の表にまとめる。

	miRNAの制御網のみランダム化		転写因子の盛業網のみランダム化		いずれもランダム化	
	Target (U)		Target (M)		Target (U)	
	Z	P	Z	P	Z	P
マイクロRNA由来の冗長性	35.10	$2.6 \times 10^{-270}$	0.08	0.468	34.24	$3.5 \times 10^{-257}$
転写因子由来の冗長性	0.98	0.164	0.00	0.500	0.96	0.168
	$P=5.8 \times 10^{-4}$		NS		$P=0.010$	

cFFC にかかわる遺伝子制御網の冗長性は、マイクロ RNA 遺伝子に多く由来することが高い有意性を持って示された。

(3) ランダム化される結合 (edge) の数に結果は影響されない



左図で示すように、2つのランダム化の方法は、ランダム化される結合 (edge) の数が異なることが分かる。この違いが、マイクロ RNA 遺伝子族による遺伝子制御網と転写因子による遺伝子制御網の違いの原因となるような artifact を生じさせる可能性を否定するために、次のような操作を行った。

下図のように、マイクロ RNA からの2つの edge を個別にランダム化する手法をおこなった。



その結果、次の表に示すように、

	miRNA-target 結合のみランダム化		miRNA-TF 結合のみランダム化	
	Z	P	Z	P
cFFC の数	-4.39	$5.8 \times 10^{-6}$	1.73	0.042
cFFC によって制御される遺伝子数	-15.04	$2.0 \times 10^{-51}$	-6.62	$1.8 \times 10^{-11}$
cFFC の冗長性	11.84	$1.2 \times 10^{-32}$	4.48	$3.8 \times 10^{-6}$
miRNA 由来の冗長性	89.16	$2.4 \times 10^{-1729}$	27.70	$3.2 \times 10^{-169}$
転写因子由来の冗長性	-4.57	$2.4 \times 10^{-6}$	0.99	0.162
	$P \approx 1.5 \times 10^{-8}$		$P \approx 6.0 \times 10^{-8}$	

マイクロ RNA からの2つの edge を片側のみランダム化した場合、いずれをランダム化しても、結果は、同様の傾向を示し、ランダム化する edge の数による artifact である可能性は否定された。

#### (4) 結果のまとめ

- ① マイクロ RNA 遺伝子族は、進化の過程でターゲットとなる遺伝子制御の冗長性を高めるよう、その制御網の構成を変化させてきた。
- ② 一方、マイクロ RNA 遺伝子族が転写因子とともに協調的に制御するターゲット遺伝子の数は顕著に減少した。
- ③ 比較として、転写因子による制御網を見ると、マイクロ RNA 遺伝子族の制御網の端的な変化と比べて、その変化は極めて少なく、安定的である。
- ④ マイクロ RNA 遺伝子族の遺伝子制御網での役割は、安定した転写因子の制御網に合わせるように進化の過程を経ていると考えられる
- ⑤ このような過程で、冗長性を増していることは、しばしば適応的に (adaptive に) 解釈される、つまり、robustness, failsafe, fine-tuning 等々である。
- ⑥ しかし、今回の研究で示されたように、マイクロ RNA 遺伝子族の制御の対象となる遺伝子は顕著に減少しているので、多くの場合、適応的な解釈は困難であることを示唆している。
- ⑦ マイクロ RNA 遺伝子族は遺伝子制御網の進化の過程で、既に確立した制御網を攪乱しないという消極的な役割で結果として冗長性を増加させていることが示唆された。

#### 5. 主な発表論文等

[雑誌論文] (計2件)

(\*は Corresponding Author)

- ① Iwama H\*, Murao K, Imachi H, Ishida T, MicroRNA networks alter to conform to transcription factor networks adding redundancy and reducing the repertoire of target genes for coordinated regulation. 査読有、Molecular Biology and Evolution. Vol 28(1), 2011, 639-646.
- ② Iwama H\*, Murao K, Imachi H, Ishida T, Transcriptional double-autorepression feedforward circuits act for multicellularity and nervous system development. 査読有、BMC Genomics. 2011. 12(1):228.

〔学会発表〕（計 5 件）

（○は発表者）

- ①岩間久和<sup>○</sup>，他、マイクロRNA遺伝子族による遺伝子制御網の変化の特性、日本遺伝学会第82回大会（口演、査読有、2010年9月20日、北海道）
- ②岩間久和<sup>○</sup>，他、マイクロRNA遺伝子族による遺伝子制御網構成の特質、（口演、査読有、日本進化学会第12回大会、2010年8月4日、東京）
- ③岩間久和<sup>○</sup>，他、マイクロRNA及び転写因子により構成される自己制御組み込み型フィードフォワード回路のヒトゲノムにおける *in silico* 解析、日本人類遺伝学会第54回大会（口演、査読有、2009年9月26日、東京）
- ④岩間久和<sup>○</sup>，他、マイクロRNAと転写因子の遺伝子制御ネットワークにおける役割の差異に関する比較ゲノム解析、第81回日本遺伝学会大会（口演、査読有、2009年9月17日、長野）
- ⑤岩間久和<sup>○</sup>，他、マイクロRNAと転写因子による協調的制御のヒトゲノムにおける網羅的 *in silico* 解析、日本人類遺伝学会第53回大会（口演、査読有、2008年9月28日、神奈川）

〔産業財産権〕

○取得状況（計 1 件）

名称：ReAlignerVR（プログラム）  
発明者：Hisakazu Iwama  
権利者：Hisakazu Iwama  
種類：GNU General Public License  
番号：-  
取得年月日：2010年6月20日  
国内外の別：国内外問わず

〔その他〕

ホームページ等  
（開発したプログラムの公開）  
<http://genet.med.kagawa-u.ac.jp/>  
（結果データの公開）  
[http://genet.med.kagawa-u.ac.jp/pub/HM0\\_5169/List.htm](http://genet.med.kagawa-u.ac.jp/pub/HM0_5169/List.htm)

## 6. 研究組織

### (1) 研究代表者

岩間 久和（HISAKAZU IWAMA）  
香川大学・総合生命科学研究センター・准教授  
研究者番号：20398035