

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 5 日現在

機関番号：12601

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25830128

研究課題名(和文)全ゲノム重複におけるエピゲノムの進化と制御の解明

研究課題名(英文)Epigenetic evolution and regulation during Whole Genome Duplication

研究代表者

曲 薇 (QU, WEI)

東京大学・新領域創成科学研究科・特任講師

研究者番号：00631566

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：エピゲノム情報が進化を推進するプライマリーな要素だと考えられる。本研究では、メダカを使ってエピゲノム情報を収集・解析し、それらの進化を詳しく調べた。その結果、パラログ遺伝子領域に関して、ゲノム配列の保存度と配列のDNAメチル化の保存度とは相関していないことを見出した。さらに、ロングリードを用いて細胞間のDNAメチル化の多様性をより効率的に捉える手法を提案し、細胞集団におけるDNAメチル化の不均一性を解明した。具体的には、ゲノム全域にわたって低メチル化領域のDNAメチル化の不均一性が抑えられたことを明らかにし、高・低メチル化領域の境界における漸次的に変化するメチル化パターンを最初に発見した。

研究成果の概要(英文)：Epigenetic information has been considered to be the primary driver for evolution. In this research, we collected epigenetic information of medaka fish and studied epigenetic evolution in depth. One of our findings is methylation level conservation rates of paralogue gene pairs seem to be uncorrelated with DNA sequence conservation rates of them. Another one is that we designed a new efficient approach to study cell-to-cell variability of cytosine methylation, which is essential for deeply understanding inherent cellular perturbation. Comparing methylation status of coexisting CpG sites on long sequencing reads, we observed repressed methylation variability in hypomethylated regions across the entire genome, and a first gradational change of methylation status on boundaries of hypomethylated regions. This approach allows a concise and comprehensive assessment of cell-to-cell DNA methylation variability.

研究分野：バイオインフォマティクス

キーワード：DNAメチル化 エピゲノム進化 次世代シーケンサ

1. 研究開始当初の背景

- (1) 進化の中心要素であるゲノム配列より、遙かに環境要因に柔軟に適応するエピゲノムのほうが、進化を推進するプライマリーな要素だと考えられる。しかし、エピジェネティック情報と生命現象の制御との関係は未解決な問題が多く、特に進化上の重要プロセスである全ゲノム重複について、網羅的なエピゲノム研究はほとんど行われておらず、その保存パターンと制御の詳細についてまだ解明されていない。
- (2) 技術革新によって、次世代シーケンサを使えば全ゲノム領域のエピゲノム情報の収集を容易にできた。また、以前の研究によって、申請者らは一塩基解像度 DNA メチロームを同定する生物実験手法とインフォマティクス技術の確立をした。(Qu *et al.* 2012)

2. 研究の目的

本研究では、インフォマティクスの技術を使い、メダカを使って DNA メチル化や、クロマチン修飾因子などのエピジェネティック情報を収集解析し、エピジェネティック情報の進化と制御を解明することを目指していた。

- (1) DNA メチル化をはじめとする多数のクロマチン修飾因子などのエピジェネティック情報が、全ゲノム重複によって形成されたパラログ領域において、どのような進化をたどり、制御機能を獲得したかを明らかにする。
- (2) エピゲノミクスレベルの細胞摂動とその分子機構の理解するために、細胞集団における DNA メチル化の不均一性・多様性を明らかにする。

3. 研究の方法

- (1) バイオインフォマティクスの手法を使って、収集したエピジェネティック情報をゲノムへ高精度にマッピングした。Ka/Ks 解析によって全ゲノム重複に由来する染色体ブロックの同定と、パラログ遺伝子のゲノム配列の保存度の比較と、進化的に保存された低メチル化領域のエピゲノムの特徴を分析した。
- (2) 我々はまず、ロングリードにおける連鎖するシトシンメチル化 (CpG サイト) を利用して、細胞間の DNA メチル化の多様性をより効率的に捉える手法を開発した。この手法のもとに、メチル化の連鎖情報を得られるための Illumina 社 MiSeq Systems を使う全ゲノムバイサルファイトシーケンシング実験を実行し、具体的にはメダカの肝臓細胞を使った実験と情報解析を行った。

4. 研究成果

- (1) 収集したメダカの DNA メチル化とクロマチン修飾マークなどのエピジェネティック情報とメダカゲノムの進化速度との関係を調べた。パラログ遺伝子領域に関して、ゲノム配列の保存度と配列の DNA メチル化の保存度とは相関していないことを見出した(図 1)。これによって、DNA メチル化の進化にはトランスエレメントが関与することが示唆された。

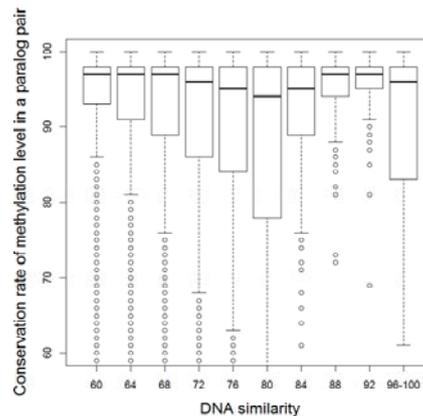


図 1. パラログ遺伝子における DNA メチル化の保存度と DNA 配列の保存度との関係。

- (2) 細胞集団における DNA メチル化の不均一性・多様性の解明。以下の成果をまとめ、現在投稿中である。

- ① 細胞間の DNA メチル化の多様性をより効率的に捉える新しい手法を提案した。具体的には、ロングリード (Illumina 社 MiSeq シーケンサ、2x300bp ペアエンド) における連鎖するシトシンメチル化 (CpG サイト) を解析した(図 2)。我々は、3~21 個の隣接 CpG サイトのメチル化の連鎖情報の抽出に成功し、隣接 5 つの CpG サイトのメチル化パターンに注目した場合、全 CpG サイトの 38% をカバーできた。これは以前の手法 (Landan *et al.* 2012) より 100 倍もカバー率を上げた。

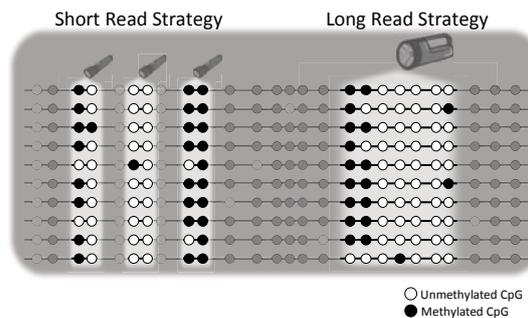


図 2. ロングリードからより長い DNA メチル化の連鎖情報が得られる。

② ゲノム全域にわたって低メチル化領域の DNA メチル化の不均一性が抑えられたことを明らかにした。細胞間のメチル化パターンの多様性を測る指標として、われわれはジニ係数 G ($G = 1 - \sum (P_i^2)$) を使った。 P_i はメチル化パターン i の頻度である。メチル化パターンが均一の場合は $G = 0$ 、逆にメチル化パターンが非常に不均一で、多様な場合は G が 1 に近づく (図 3)。興味深い発見は、メチル化パターンの多様性は低メチル化領域に抑えられることである。これは DNA メチル化の複製に関わるメチルトランスフェラーゼの酵素の精度が不完全であることが示唆された。図 3 は隣接 5 つの CpG サイトのメチル化パターンの分布を示しているが、同様の傾向は、全ての 3~21 個の隣接 CpG サイトのメチル化パターンにおいて観察された。

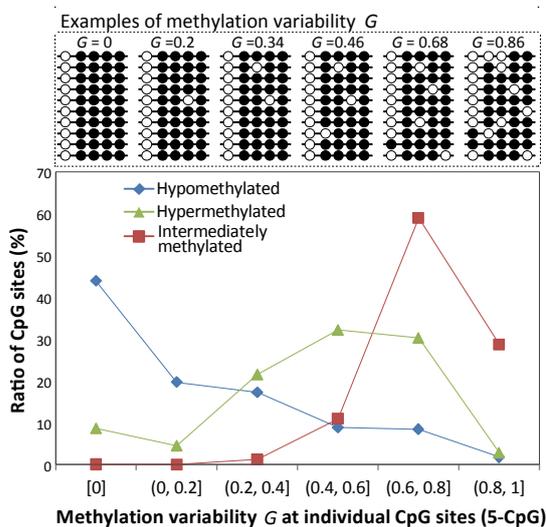


図 3. DNA メチル化パターンの多様性は高メチル化、低メチル化と中間的にメチル化された領域に大きく異なる。

③ 高・低メチル化領域の境界における漸次的に変化するメチル化パターンを最初に発見した。今までは高・低メチル化領域の境界に、メチル化レベルの平均値が大きく変わることがよく知られるが、その変化の詳細は不明であった。我々の手法を使い、ローグリードにおけるメチル化のパターンの頻度に比率の差の検定を行い、トップに相違であったパターンのほとんどが漸次的なメチル化の変化をサポートすることが明らかであった (図 4)。

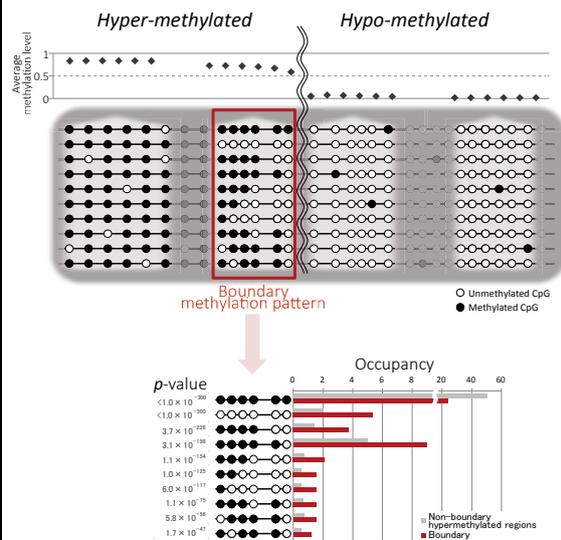


図 4. 高・低メチル化領域の境界における DNA メチル化パターンの変化。従来の手法によれば、平均化されたメチル化レベルが得られる (上図) ことに対して、ロングリードにおけるメチル化の連鎖を使えば、より詳細な変化の様子が捉える (中図)。比率の差の検定を行った場合のトップ 10 個のパターン (頻度 $\geq 1\%$, 下図) がリストされたが、トップの 7 個は漸次的変化を構成している。

<引用文献>

- ① Qu *et al*、Genome-wide genetic variations are highly correlated with proximal DNA methylation patterns、Genome Research、2012、22:1419-1425
- ② Landan. et al、Epigenetic polymorphism and the stochastic formation of differentially methylated regions in normal and cancerous tissues、Nature Genetics、2012、44:1207-1214

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 2 件)

- ① Ryohei Nakamura, Tatsuya Tsukahara, Wei Qu, Kazuki Ichikawa, Takayoshi Otsuka, Katsumi Ogoshi, Taro L Saito, Kouji Matsushima, Sumio Sugano, Shinichi Hashimoto, Yutaka Suzuki, Shinichi Morishita and Hiroyuki Takeda, Large hypomethylated domain serves as strong repressive machinery for key developmental genes in vertebrate、Development、査読有、2014、141(13):2568-80、DOI:10.1242/dev.108548
- ② Yoichiro Nakatani, Wei Qu and Shinichi Morishita, Comparing the Human and Fish

Genomes、eLS、査読有、2013、
DOI: 10.1002/9780470015902.a0021004

(3)連携研究者
()

[学会発表] (計 3 件)

- ① Wei Qu、Assessing cell-to-cell DNA methylation variability on individual long reads、第11回国際ゲノム会議、2015年5月20日-2015年5月22日、学術総合センター(東京都千代田区)
- ② Wei Qu、Assessing cell-to-cell DNA methylation variability on individual long reads. The Biology of Genome、2015年5月5日-2015年5月9日、Cold Spring Harbor Laboratory (米国ニューヨーク州 Cold Spring Harbor)
- ③ 曲 薇、メダカのエピゲノムから分かった進化と制御、NGS 現場の会第三回研究会(招待講演)、2013年09月04日、兵庫県神戸市神戸国際会議場

研究者番号:

[図書] (計 件)

[産業財産権]

○出願状況 (計 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

○取得状況 (計 件)

名称:
発明者:
権利者:
種類:
番号:
出願年月日:
取得年月日:
国内外の別:

[その他]

ホームページ等

Medaka Methylome Browser

<http://utgenome.org/methylome/>

6. 研究組織

(1) 研究代表者

曲 薇 (QU, Wei)
東京大学・新領域創成科学研究科・特任講師
研究者番号: 00631566

(2) 研究分担者

()

研究者番号: