

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 11 日現在

機関番号：17102

研究種目：挑戦的萌芽研究

研究期間：2013～2014

課題番号：25640078

研究課題名(和文)腫瘍ゲノム解析として転座・逆位を含む包括的構造解析を可能とする新規システムの開発

研究課題名(英文)Development of a novel method for comprehensive analysis of structural changes of cancer genome

研究代表者

水野 晋一 (MIZUNO, SHINICHI)

九州大学・先端医療イノベーションセンター・准教授

研究者番号：40569430

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：腫瘍のゲノム解析は、原因遺伝子探索をはじめ腫瘍診断に重要な検査である。本研究では、包括的なゲノム構造解析を高精度に可能とする新規システムを開発し腫瘍解析へ応用した。本システムは環状化DNAの結合両末端(メイトペア)解析を基本とし、ゲノム・ライブラリーの各DNA断片へ固有塩基配列を導入し、シンプルな2段階DNA環状化を行うことで、「単分子DNA環状化」由来の正確なメイトペア配列のみをコンピューター(in silico)解析で選別可能とした新規法である。本システムは、腫瘍ゲノムの転座・逆位の解析も可能であり、経済性にも優れた方法であることから、腫瘍の診断・研究に役立つと期待される。

研究成果の概要(英文)：We developed a novel method for comprehensive analysis of structural changes of cancer genome. Our method is based on a mate-pair analysis, which enables us to collect information of both ends of DNA fragments from genome. In conventional mate-pair analysis, there must contain circularized DNAs derived from multiple DNA fragments and those mate-pair can be a cause of false positive data. To solve this problem we developed a novel method, which allows us to extract "right" information of mate-pairs derived exclusively from circularized DNAs of single DNA origin. By this method, we can make a comprehensive analysis of cancer genome in a highly precise manner, and we could obtain several candidate mate-pair information in ATL(adult T-cell leukemia/lymphoma) samples. This method is useful for analysis of structural changes of cancer genome and will provide us information for diagnosis and treatment of cancer.

研究分野：腫瘍学

キーワード：ゲノム解析

1. 研究開始当初の背景

腫瘍診断における遺伝子解析の必要性は高くなってきており、特に腫瘍ゲノムを網羅的に解析することは予後・治療法を含め腫瘍の診断に重要な役割を果たしていくと考えられる。腫瘍ゲノム解析においては、現在 CGH (comparative genomic hybridization) 法が導入されたことで、腫瘍ゲノム解析、特に構造解析が飛躍的に進歩してきている。しかし、CGH 法は、ゲノムのコピー数異常 (増幅・欠失) の構造解析に優れているのに対し、転座・逆位等の構造異常の検出は困難であるという欠点がある。そのため、転座・逆位を含む包括的なゲノム構造解析を高精度・高感度に解析できる方法が求められている。

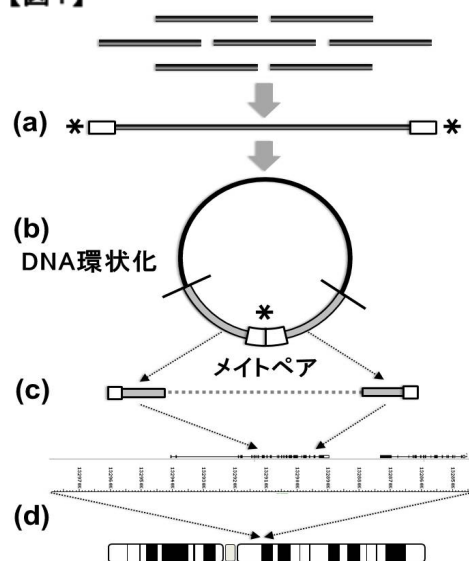
ゲノム構造解析法の候補としては、遺伝子塩基解析からゲノム構造解析までを含むトータルな遺伝子解析法としての全ゲノム解析 (WGA: whole genome analysis) がある。しかし、次世代シーケンサーは経済的になってきているものの、多数症例を対象とするには WGA は経済的負担がかかる解析法である。また、データ解析の負荷も大きく、short read の解析のため使用プログラムやデータ解析条件によりゲノム構造解析結果に曖昧さが生じるという報告もあり、WGA の一般検査への応用にはまだ時間がかかると考えられており現実的ではない。

本研究では、メイトペア解析法の応用により、転座・逆位を含むゲノム構造異常を検出可能とする新規システムを開発し、腫瘍解析へ導入する。

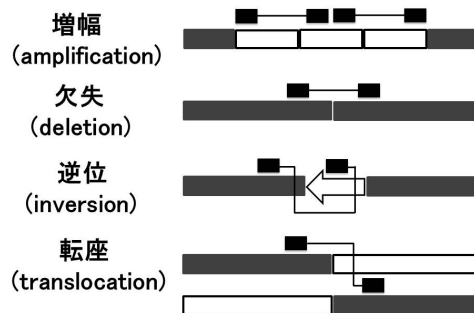
2. 研究の目的

(1)ゲノム構造解析を可能とする新規法開発
本研究では、転座・逆位を含むゲノム構造異常を検出する新規方法を開発し、新たな腫瘍診断法を確立することを目的としている。本システムは、メイトペア解析を基本としており、ゲノム DNA の環状化から得られる結合

【図1】



【図2】



両末端 (メイトペア) の配列解析 (図 1 a. b) を行う。このメイトペア配列情報をゲノム情報にマッピングすることにより、ゲノム構造を解析する (図 1 c. d)。具体的には、メイトペア情報のマッピング位置とパターンからゲノムの量的変化を含め様々な構造変化を検出することができるであろう (図 2)。

本法はシンプルな方法であるが、単なるメイトペア法では精度・感度ともに不十分で、研究目的を遂行することはできない。つまり、後述のように従来のメイトペア法では原理的な問題があり、本研究目的には不適である。そのため、本システムではこの問題を新しい方法論の導入で解決することにより、高い精度でゲノム構造変化を検出すること新規方法を開発する。

(2)本法の実証と腫瘍ゲノム解析への応用
本システムの構築後、システム条件設定を行うとともに、造血管悪性腫瘍の臨床検体への応用を実施する。具体的には、成人 T 細胞性白血病 (ATL) のゲノム解析に応用し、本新規システムの有用性を実証する。

3. 研究の方法

(1)メイトペア法を基本とする新規ゲノム構造解析法の開発

本システムでは、現在の遺伝子配列解析の技術を最大限に生かし、ゲノム構造異常の包括的な解析を高い精度でかつ経済的に可能にする。特に腫瘍ゲノムのコピー数のみならず CGH 法では困難とされる転座・逆位の解析を可能とすることで、腫瘍の診断および研究への応用を行う。

研究方法のポイントは、ゲノム DNA から作成されるメイトペア・ライブラリーにおいて単分子 DNA による自己環状化分子に由来する正確なメイトペア配列を、データ解析時にコンピューター上 (in silico) で選別することである。

本システムは、図 1、図 2 に示したようにメイトペア法を基本としており、ゲノム DNA 環状化から得られるメイトペア配列情報をヒトゲノム・データベース上にマッピングし、このゲノム上の位置とパターンを解析することでゲノム構造変化を明らかにする。

しかし、従来のメイトペア法には、重大な問題点があり、原理上、単分子 DNA 断片の

環状分子に由来する正確なメイトペアのみを選択することが不可能であり、多数の偽メイトペアが必然的に混入してしまうことになる。

例えば、3種類の DNA 断片を例として、これらを環状化する場合を想定してみると、DNA 断片の末端の形状は同じであるため互いに区別できない。そのため ligation を行った際には、目的の通りに (i) 単分子 DNA による環状化分子が形成される場合、(ii) 2種以上の複数 DNA が互いに ligation されて結果として環状化分子を形成する場合、(iii) 環状化されずに線状分子のコンカテマーにとどまる場合、の各々が起こりうる。線状分子については酵素的に分解除去することが可能であるが、環状化分子はスーパーコイルを形成するためサイズ分離も難しく、このような一定の頻度で形成される複数分子による環状化 DNA を排除することができないこととなる。その結果として、偽メイトペアが生じることとなり、この配列情報は複数の DNA 断片に由来するためゲノム構造異常と誤って解析されることになり、解析精度を大幅に低下させ、ゲノム構造変化検出の大きな障壁となる。

従来は、単分子 DNA による環状化の効率化を期待して、ポリエチレングリコール (PEG) 添加等による反応溶液の工夫や核酸濃度の最適化等が試行されており、その結果として複数分子による環状化 DNA (偽メイトペア) の頻度をある程度までは下げることができているが、完全に排除することは不可能であった。そのため、環状化の実験過程にこれ以上の改良を加えても、得られるメイトペア情報の精度には限界があることになる。そのため、従来のメイトペア解析では大量データの取得および統計処理等の対応がなされており、仮にゲノム構造異常に対応するメイトペア配列が解析データとして得られたとしても、それらは偽メイトペアによるノイズとして棄却されることになる。

本システムではこの問題を新しい方法論で解決し、メイトペア解析を次世代シーケンサーへ応用することにより、CGH 法に欠けている転座・逆位の解析を含むゲノム構造解析を可能とする包括的なゲノム解析法を開発する。

(2) 本法のシステム条件設定と腫瘍ゲノム解析への応用

本システム開発過程において、ゲノム構造解析の効率化を図るため、DNA 環状化を含めた実験過程の各ステップにつきシステムの条件最適化を行う。さらに、造血器悪性腫瘍を対象として臨床検体による実証研究を行う。

4. 研究成果

(1) ゲノム構造の高精度解析を可能とする新規方法の開発

① ゲノム断片へ付加する特異的配列を並置したアダプターの作成

本研究のゲノム構造解析では、ゲノム DNA を制限酵素により 10Kbp~20Kb に部分消化した DNA 断片を使用した。DNA 断片へ付加するアダプターは、DNA 断片への結合サイトに加え、対側末端に制限酵素サイト (X: NotI) を布置し、アダプター内には2箇所の相同的な特異的塩基配列 (N) を設置している。その特異配列間に2箇所の制限酵素サイト (Y: EcoRI) を配置した。ここでの特異的配列のランダム塩基 (N) は同一アダプター内では同じ配列 (相同配列) が並ぶよう設計している。この同一の特異的配列の並置には、ランダム塩基配列を有する合成 DNA にループ構造を付加し、strand replacement polymerase による塩基伸長を行うことで作成した。また、ゲノムを対象とするため、次に示す2段階目の制限酵素切断 (EcoRI) によりゲノム断片が同時に切断される可能性がある。そのため準備した DNA 断片は事前に EcoRI メチラーゼによって修飾し EcoRI での切断が起きないようにしておくようにしている。

② メイトペア選択のための2段階 DNA 環状化

当該アダプターを DNA 断片に付加した後、制限酵素 (X: NotI) での切断後に ligation を行い1段階目の自己環状化を行う。この時、環状化されていない DNA は酵素的に除去する。次に制限酵素 (Y: EcoRI) によりアダプター内に2箇所ある相同的特異配列の間で切断し、再度 ligation 反応を行うことで2段階目の自己環状化を行う。手順としてはシンプルなのこの2段階 DNA 環状化を行うのみであるが、この結果、各環状化の過程により、アダプター内の特異的配列の組み合わせパターンが変化することになっている。

つまり、1段階目 DNA 環状化が目的通りに単分子 DNA で形成され、さらに2段階目 DNA 環状化もやはり単分子 DNA であった場合、この2段階目の環状化 DNA から得られるメイトペア断片は、正しい配列情報を有することとなり正確にゲノム構造を反映する。この時、アダプター内の特異的配列は必ず同一配列の組み合わせとなり、環状化過程が単分子 DNA により正しく進行したことの指標となる。この特異的配列の組み合わせによるメイトペア配列を選別できることが本法の最大の特徴である。一方、1段階目あるいは2段階目のどちらかの環状化において複数分子 DNA により不適切な環状化 DNA が形成された場合、つまり目的から外れて偽メイトペアが形成されてしまった場合、アダプター内の2箇所の特異的配列は互いに異なる配列の組み合わせとなることになり合致しないことになり、この特異的配列の組合せを指標に、メイトペア作成が予定通り進んだか、そうでないかを判定することとなる。

従って、これらメイトペア配列を次世代シーケンサーによる配列解析を行うことで、正しいメイトペア配列：特異的配列が同一組

み合わせである正確なメイトペア、エラーメイトペア配列：特異的配列の組み合わせが合致しない環状化過程でエラーが生じた偽のメイトペア、不完全配列：環状化過程そのものが正常に進行せず EcoRI 配列間が残存する不完全なメイトペア、と環状化過程の進行具合を区別することができる。これらをコンピュータ上で選別することで、正確なメイトペア配列情報のみを抽出することが可能となっている (in silico selection)。つまり、従来のメイトペア解析では原理的に不可避であった複数 DNA 分子による偽メイトペア配列のノイズを完璧に排除することができることとなる。これらの正確なメイトペア配列をゲノム情報にマッピングし、ゲノム・ライブラリー作成時の DNA 断片サイズを指標として、互いのメイトペア配列の推定距離および配列方向から増幅・欠失・転座・逆位を判定した。

従来行われている大量データの取得と統計処理では、メイトペア情報のノイズ処理に限界があり、また実際のゲノム構造異常もノイズとして処理されてしまうため腫瘍のクローン解析へ応用することができない。本システムでは、偽メイトペアを排除し、正確なメイトペアのみを選択することが可能となっており、高い精度でのゲノム構造解析を遂行できる基盤的技術を確立することができる。

(2) 本法のシステム条件設定と腫瘍ゲノム解析への応用

① 本法のシステム条件設定

本システムの構築後、より効率よくゲノム構造解析を行うため、DNA 環状化を含めた実験過程の各ステップにつきシステムの条件最適化を行った。

表 1 にメイトペア解析による配列結果の一部を示すが、アダプター配列およびアダプター内に相同の特異配列を有し、環状化過程が正常に進行したと評価できるクローンからは、ゲノム断片長に相当したメイトペア情報を得ることができている。下段の 2 配列は、ひとつは異なる染色体にマップされ構造異常が疑われるが、アダプター内の特異的配列が合致せずエラークローンであることが分かる。もうひとつは、メイトペア情報がリピート配列にあたりゲノム上の多数の部位にマッピングされ解析不能であるクローンである。

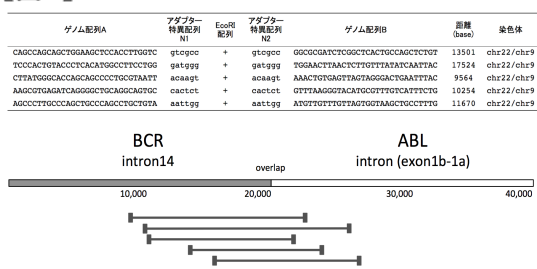
【表 1】

ゲノム配列A	アダプター特異配列 N1	EcoRI 配列	アダプター特異配列 N2	ゲノム配列B	距離 (base)	染色体
GAATCTGGGGCCAGCTTCCTCCACCGC	aaacgg	+	AAACTGAAAGATTCTTTTTCATTTGTC	11368	chr3	
GGGGGGGGGGGAGTACAGCAGCAGCAG	aaactg	+	TTTTTGTATGTCAGAACTGCACTCAAAAAT	12559	chr7	
GGAAAAGGATCCCCCACTACACCTCCG	acpagg	+	CATGCCCTTAAAGATGTTCTGGGAAGT	14415	chr4	
GAACCTGAGACTGTTATGAAACCCAGT	agacgt	+	CATCCACCCACTCACTCACTCTCTACCCA	18115	chr22	
AAAGTGTAAATACGAAGGCTACCTCTTT	attctt	+	AAACATAAATCAAAATCAACATTTATAA	14732	chr1	
TGAGGCTGAAATGTACGGGATTCACAGT	caactg	+	ATTTCTAGTTCACATGATTTTCTCTAG	15382	chr2	
AAATACAGGAAAGATGAAATGATTTACTG	ccactt	+	TTCCATTTAAATCTATTTTCTTAGC	9670	chrX	
AGTTAGATATATGAGAGCTTACAAACAT	cccttg	+	TCTCTCACTCACTTGGACAGCCCTCCA	14882	chr16	
ATRAAAATGTAAGAACCATTAACCATAC	gataaa	+	ACCTCTCTCCAGGGCATCAAGCTGACAC	13471	chr1	
TGACTGTATTTTCTCTAAATACAAC	gcaagg	+	TGGAAGAGGATGAAAGCTTCTCACTGCC	10704	chr4	
GCATCTCCACTGTAGGAGCTCTTTCTCT	gcoectt	+	CCATCGAGCCTCTCTCTTTTGTAGAGAC	17411	chr6	
TACATCTTTCTTTCTTTCTTTCTTTCT	gcoact	+	TTTTTTTCTTCACTTTGAAAGAGCTGAG	16360	chr2	
ATTTCAAAATATGCTTTGCTTTGGAGAGC	tgaaag	+	KCACAGAGCAGACACGAAAAGTGGCTCC	14265	chr15	
ATCCAGATAGGCTGCCAGCAGGGCTGCT	tggttc	+	CAGAATACTATGAAATTCACCAAT	13265	chr5	
GGAGTCACCTCTCTTAAGGCTCTCTCT	ttctct	+	ATTTTCTTGTACTCTGGCAGAGTCTGC	19371	chr10	
CTCCTCTCCCTCTCTCTCACTCCACT	agotta	+	GGAGCCCGAGCTCTCTCACTCCAGACC	-	chr2/chr3	
AAATATATGCTTTTCAATATATATAGC	ccatga	+	ATAAAGGAAGAGTCTTACTCTCAGAT	-	chr2/7	

一方、解析の過程でアダプター内の特異配列が同一であったクローンに関わらず、複数分子由来の環状化分子による偽配列のノイズが一定頻度存在することが分かった。この原因については、データ解析により 2 段階目の制限酵素切断時に制限酵素サイト (EcoRI) の切断不十分によると判明した。アダプターは遺伝子合成によって作成したが、約 1% の割合で合成エラーが生じ得る。そのため合成エラーが EcoRI サイトに生じたアダプターは制限酵素では切断されず、ある一定の頻度でノイズを形成してしまうことが明らかとなった。そのため、新規アダプターとして、EcoRI 切断 DNA 断片をあらためて ligation することにより、特異配列および 4 カ所の正確な EcoRI サイトを有するアダプターに改良した。この新規アダプターは、目標通り正しい EcoRI サイトで構成されていることが確認され、導入によりノイズを軽減することができた。

対象とした K562 細胞には Ph1 染色体が存在し、9 番および 2 番染色体間で転座が生じている。従って、本システムが正常に動作すれば、この転座したゲノム部分をカバーするメイトペア配列を得ることができるとなる。解析の結果、図 3 のように転座部に該当するメイトペア群が検出されており、本システムによるゲノム構造異常の検出が可能であることが示された。

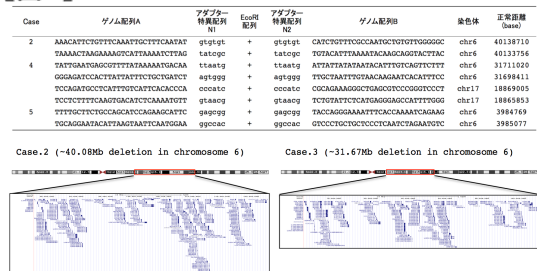
【図 3】



② 成人 T 細胞性白血病 (ATL) のゲノム構造解析への応用

本システムの実証試験のため、5 症例の ATL 検体のゲノム構造解析に応用した。明らかなゲノム構造異常として 3 症例で 6 番染色体長腕の欠失領域を示す複数のメイトペアが得られ、1 例で 1 7 番染色体短腕の欠失を示すメイトペア群が得られている (図には代表配列および 2 症例の欠失領域を示している)。

【図 4】



実際には、転座を含むゲノム構造異常を示唆するメイトペアが多数に認められているが、情報クローンが少数であることから、さらにシーケンスを深く読むことでそれらの中から正しくゲノム構造異常を検出できる可能性があり解析を継続している。新たな問題点としては、複数のゲノム箇所にマッピングするメイトペア配列がリピート配列以外にも多く観察されており、メイトペア配列の解析塩基長をさらに伸ばすことによる精度の向上を進めている。

本研究において、メイトペア法を基本とする新規ゲノム構造解析法を開発し、実際にゲノム構造異常を検出できることを実証した。さらにシステムの精度向上を図り一般化するとともに、臨床検体への応用によりゲノム構造解析を進めていく。

5. 主な発表論文等

〔雑誌論文〕（計 0 件）

〔学会発表〕（計 0 件）

〔産業財産権〕

○取得状況（計 1 件）

名称：METHOD FOR PRODUCING CIRCULAR DNA FORMED FROM SINGLE-MOLECULE DNA

発明者：MIZUNO Shinichi, NAGAFUJI Koji, OKAMURA Takashi

権利者：KURUME UNIVERSITY

種類：特許

番号：US 8,962,245 B2

出願年月日：2013年06月27日

取得年月日：2015年02月24日

国内外の別：外国

6. 研究組織

(1) 研究代表者

水野 晋一 (MIZUNO, Shinichi)

九州大学・先端医療イノベーションセンター

・准教授

研究者番号：40569430