

令和 6 年 6 月 24 日現在

機関番号：82505

研究種目：若手研究

研究期間：2019～2023

課題番号：19K19492

研究課題名（和文）STR型検査におけるPCR副産物の発生量予測モデルの開発

研究課題名（英文）Development of the model to predict PCR byproduct generation in STR typing.

研究代表者

深川 貴志（Fukagawa, Takashi）

科学警察研究所・法科学第一部・研究員

研究者番号：90801572

交付決定額（研究期間全体）：（直接経費） 2,000,000円

研究成果の概要（和文）：本研究では、「(1)人工合成DNAのスタター比率及びスタター発生率を測定する系の確立」、「(2)繰返し配列の塩基の並び順によるスタター比率の相違」、「(3)スタター発生率は、PCRサイクルで一定ではない可能性」の3つの成果を得た。このうち、(2)及び(3)については、当初想定していなかった結果であった。特に(2)は、従来の定説である「反復配列のAT含有量が高いほどスタター比率が高くなる」が4塩基の反復配列には当てはまらず、塩基の並び順に影響されることを強く示唆するものであり、非常にインパクトのある新たな知見が得られたといえる。

研究成果の学術的意義や社会的意義

現在、STR型検査におけるスタターの性質の詳細及び発生メカニズムについては、理論面において理解が進んでいないため、実務においては多数の試料の実測値からその性質を推定してスタターを扱っている。本研究の成果は、スタターの性質の詳細及び発生メカニズムに関する新しい知見であり、かつ従来の知見が当てはまらないことを示したものである。スタターの詳細について理解が進むことはSTR型検査、ひいてはDNA型鑑定の理論面のさらなる解明につながることから、本研究の成果は学術的意義・社会的意義の大きいものであると考えている。

研究成果の概要（英文）：In this study, three main results were obtained: 1. Establishment of a system to measure the stutter ratio and stutter generation ratio of synthetic oligonucleotides, 2. Differences in stutter ratios depending on the order of bases in repeat sequences, and 3. The possibility that stutter generation ratio is not constant across PCR cycles. Of these, the results of 2 and 3 were not initially expected. In particular, the result of 2 strongly suggests that the conventional theory, "the higher AT content of repeat sequence, the higher the stutter ratio", does not apply to 4 base repeat sequences, but is affected by the order of the bases. This is a very impactful new finding.

研究分野：法医学

キーワード：スタター フラグメント解析 合成DNA

### 1. 研究開始当初の背景

DNA 型鑑定に用いられる STR (Short tandem repeat) 型検査は、PCR により STR 領域 (アレル) を増幅し、DNA フラグメント解析で蛍光標識された増幅産物をピークとして検出した後、型判定解析をすることで行われる (図 1)。STR 型検査における PCR では、スタターと呼ばれる、本来のアレルより 1 回繰返し回数が少ない副産物が生じる (図 1)。現在、アレルかスタターかの判断の根拠は、スタター位置のピーク高とメインアレルのピーク高の比 (スタター比率) 及び検査人の経験である。しかしながら、スタターは生成量が微量であることから PCR の確率的バイアスが強く作用するため、スタター比率は PCR ごとに変動する。それゆえ、スタター比率は定量性に乏しく、スタター判断の客観的根拠としては弱い。さらに検査人の経験は主観的であり、客観性は乏しい。そのため、ときにスタターの判断を誤り、STR 型検査の目的である個人識別を果たせないことがある。これを防ぐため、より定量的かつ客観的な新たな指標が望まれている。

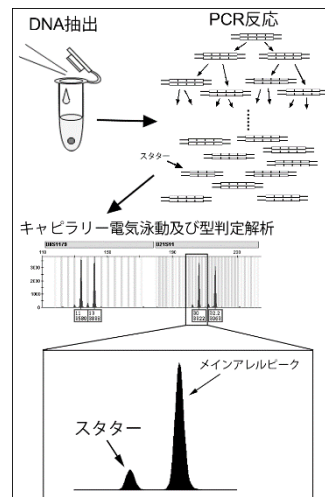


図 1. STR 型検査の概要とスタター

このような指標には、スタターの発生メカニズムに関する要素のうち、確率的バイアスに影響されず一定の値を持つ要素が最も適していると考えられる。しかし、スタター発生メカニズムに関する要素が完全に解明されていないことから、定量的かつ客観的な指標となり得る要素は見出されていない。スタターは PCR で発生することが確認されていることから、PCR でスタターが発生する確率、つまりスタター発生率は、繰返し回数・繰返し配列の種類・STR 領域ではない部分の寄与等の DNA 配列の様々な要素及び温度等の PCR 条件の要素が複合的に作用していることが判明している。つまり PCR 条件が一定であれば、スタター発生率は DNA 配列に関する要素のみに依存し、前サイクルまでのスタター生成量などの確率的な事象に影響されないため一定の値になると考えられる。このことから、スタター発生率はスタター比率よりも定量的な数値・指標となり得ると考えられるが、スタター発生率は測定方法がないため、報告されたことはない。そこで、本研究においてこれらを測定することで、スタター比率よりも定量的な数値・指標の確率を試みた。

### 2. 研究の目的

スタター発生に関しては繰返し回数・配列の種類、及び STR 領域ではない部分の寄与を含む様々な要素が複合的に作用しているが、繰返し回数以外の各要素を個別に調べた研究はない。これはスタターに関する研究で通常用いられるヒトゲノム DNA では DNA 配列の変更ができず、スタター発生に関する各要素を独立させることができないためである。しかし、スタター発生メカニズムについて理解するためには、各要素がそれぞれどのような作用をしているかを把握しておく必要がある。そこで、本研究では DNA 配列の変更が可能な人工合成 DNA を用い、繰返し回数・繰返し配列の種類に焦点を絞って研究を行う。

そこで、本研究ではスタター発生に関する要素が最も少ない単純な STR 領域、つまり 1 種類の繰返し配列 (とプライマー結合領域) からなる人工合成 DNA を用いて、

- ①スタター発生率の測定手法の確立
- ②測定したスタター発生率を用いてスタター比率の分布予測モデルの確立
- ③繰返し回数及び繰返し配列の種類に由来するスタターの定量的な性質の解明

を目的とする。

### 3. 研究の方法

#### (1) 人工合成 DNA のスタター比率及びスタター発生率を測定する系の確立

4 塩基が複数回反復した配列 (以下「繰返し配列」とする。) とその両端にプライマー結合領域だけを持つ一本鎖 DNA 及びその逆相補鎖の一本鎖 DNA を合成し PAGE 精製したものを等量混ぜて二本鎖 DNA にしたものに、蛍光標識されたプライマーを入れて PCR を行い、得られた PCR 増幅産物について、DNA フラグメント解析をすることで、アレルのピーク及びスタターピークのピーク高を取得し、スタター比率を算出する系の確立を行った。

まず、通常の PCR を行うことで得られるスタター比率を算出する系について、PCR のサイクル数、適切な鋳型 DNA 量及びプライマー濃度の検討を行った。次に、サイクル数を 1 回にした PCR を行うことで得られるスタター発生率を算出する系について、1 サイクル PCR の条件、適切な鋳型 DNA 量及びプライマー濃度の検討を行った。なお、スタター比率及びスタター発生率は、4 塩基短い産物のピークの高さを本来のアレルのピークの高さで除算した値とした。

#### (2) スタター比率の分布のシミュレーション

1 サイクルあたりのスタター発生率及び鋳型 DNA 量から、本来の繰返し数の DNA 分子数及びス

タターの DNA 分子数をシミュレーションするプログラムを作成した。プログラムの原理は次のようにした。

- ① 繰返し配列を持つ DNA の初期 DNA 分子数を、鋳型 DNA 量から算出する。
- ② 1 サイクルで分子数を 2 倍にする。このとき、スタター発生率に基づいて繰返し数が 1 少ない分子を生じさせる。
- ③ スタター比率を算出する系の PCR サイクル数と同じ 29 サイクル分②を繰返し、元の繰返し配列を持つ DNA 分子数とスタターの DNA 分子数からスタター比率を算出する。

#### 4. 研究成果

##### (1) 人工合成 DNA のスタター比率及びスタター発生率を測定する系の確立

最初にスタター比率の測定系の確立を行った。種々の検討を行った結果、PCR 反応には AmpliTaq Gold™ 360 Master Mix (サーモフィッシャーサイエンティフィック) を用いることにした。鋳型 DNA 量は約 1 amol の二本鎖 DNA が、プライマー濃度はフォワードプライマー、リバースプライマーのいずれも 1 μM が、最も適切であった。また、PCR 反応は 95°C 10 分加熱後、95°C 15 秒、59°C 30 秒、72°C 30 秒を 29 サイクル行い、最後に 72°C において 60 分加熱というプログラムにした。PCR 増幅産物を 3500xL Genetic Analyzer (サーモフィッシャーサイエンティフィック) により泳動してピークを検出し、GeneMapper™ ID-X Software v1.6 (サーモフィッシャーサイエンティフィック) により解析することで、スタター比率 (4 塩基短い産物のピークの高さ/本来の塩基数のピークの高さ) を測定する系 (図 2、一番高いピークが本来の塩基数のピーク、矢印は 4 塩基短い産物のピーク) を確立した。

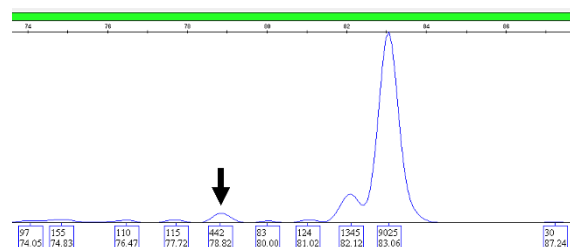


図 2. スタター比率の測定系で得られる電気泳動図

次にスタター発生率の測定系の確立を行った。最終的に、スタター比率測定系の条件のうち、鋳型 DNA 量を約 1 μmol に変更し、PCR 反応のサイクル数を 29 サイクルから 1 サイクルに変更した。その結果、スタター比率の測定系で得られる電気泳動図 (図 2) と同様の電気泳動図を得られる測定系を確立できた。

##### (2) 繰返し配列の塩基の並び順によるスタター比率の相違

確立した系を用いて、スタター比率及びスタター発生率の測定を試みた。まず、鋳型 DNA 量が少ないため DNA の消費が少ないスタター比率の測定から行った。先行研究より、繰返し配列の AT 含有率が高いほどスタターが生じやすい、つまりスタター比率が高くなることが報告されている。この報告の再現性を確かめるため、繰返し配列の塩基構成が A、A、T、G の 4 つで並び順が異なる全 12 種類の合成 DNA (繰返し回数は 12 回) についてスタター比率を測定したところ、先行研究とは異なり、並び順が異なるだけでスタター比率に差がある傾向がみられた (図 3、AATG ~ TAAG の 12 種類)。また、反復領域に出現する 4 塩基繰返し配列が共通する配列群 (例えば反復配列 AATG, ATGA, TGAA, GAAT はいずれも反復領域に「AATG」が出現する。この配列を「共通繰返しモチーフ」と定義した。図 3 の点線で囲まれた配列は共通繰返しモチーフである。) はスタター比率が類似の値になっており、共通繰返しモチーフが異なるとスタター比率が大きく異なる場合があることが判明した。なお、コントロールとして繰返し配列を持たない配列についてもスタター比率の測定を行ったが、スタター比率は 0.5% 未満となりほとんどスタターが生じなかった。わずかに生じた

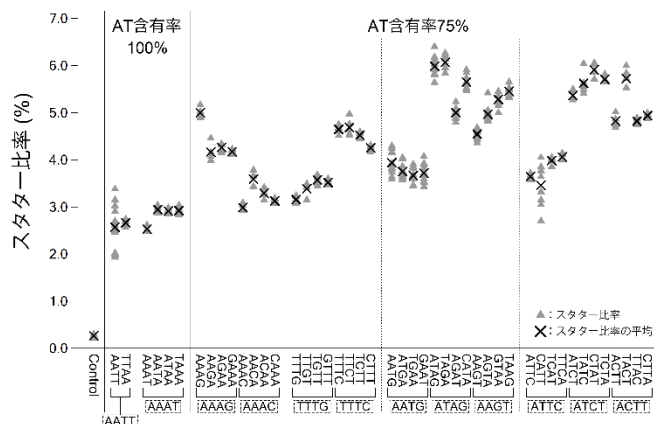


図 3. スタター比率測定結果 (コントロール、AT 含有率 100% 及び 75% の繰返し配列)

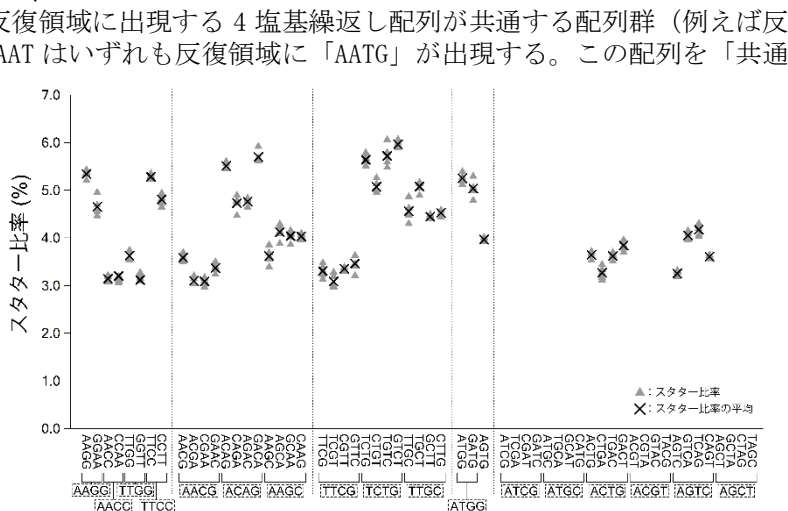


図 4. スタター比率測定結果 (AT 含有率 50% の繰返し配列)

ものについては、スタターではなく DNA の合成の際に取り除けなかったエラー配列が PCR 増幅したものと考えられる。

さらに、塩基構成を拡張し、AT 含有率 100%、75%、50% 及び 25% の種々の配列（繰返し回数は 12 回）についてスタター比率の測定を行った（図 3、4、5）。その結果、塩基構成が A、A、T、G の場合と同様に、共通繰返しモチーフが同じ配列群はおおむねスタター比率が同程度であったのに対し、共通繰返しモチーフが異なるとスタター比率が大きく異なる場合があった。なお、AT 含有率 50% の配列のうち、共通繰返しモチーフが「ATCG」、「ATGC」、「ACGT」、「AGCT」の G と C が隣り合う配列群は、正しい塩基長の PCR 産物が得られなかったため、スタター比率をプロットしていない（図 4）。

また、図 3、4、5 より、AT 含有率にかかわらずスタター比率は 2.0~6.5% の間に収まっており、AT 含有率とスタター比率に関連性はみられなかった。

以上から、本研究により、従来の仮説であり過去に報告されていた「反復配列の AT 含有量が高いほどスタター比率が高くなる」<sup>①</sup>、<sup>②</sup>は、4 塩基の反復配列には当てはまらず、スタター比率は塩基の並び順に大きく影響を受けることが判明した。これは現行の定説を更新しうる大きな新規の知見であり、学術的に価値があるものといえる。なお、現行の定説の根拠となった文献のうち、<sup>②</sup>は 4 塩基の反復配列を扱っており、一見本研究の成果と矛盾するよう見えるが、<sup>②</sup>の文献では 2 種類の反復配列（「AGAT」と「AGCG」）しか扱っておらず、本研究でもその 2 種類だけに注目すれば<sup>②</sup>の文献と同じ結果が得られていることから、過去の報告と矛盾する結果ではないといえる。

本成果は当初予期していなかったものであり、本研究課題の計画の前提でもあった定説「反復配列の AT 含有量が高いほどスタター比率が高くなる」が異なることが示唆されたため計画を大きく変更する必要が生じたが、非常にインパクトのある新たな知見が得られたといえる。

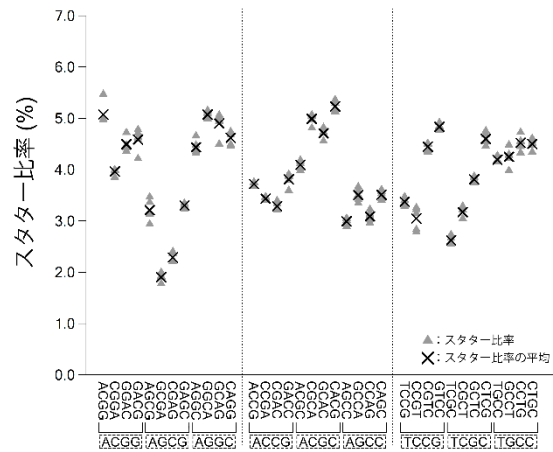


図5.スタター比率測定結果（AT含有率25%の繰返し配列）

### (3) スタター発生率は、PCR サイクルで一定ではない可能性

確立したスタター発生率の測定系を用いて、いくつかのスタター発生率の測定を試みたところ、約 0.5~6.0% となった。これらの値を用いてスタター比率の分布のシミュレーションを試みたところ、スタター比率が異常に高く（約 20~50%）シミュレーションされることが判明した。これを受けて今までのデータを精査したところ、スタター発生率の測定系の確立の過程で最後の 72°C の過熱時間を検討したデータにおいて、加熱時間が短いほど、本来の塩基長のピークの高さが低くなり、本来の塩基長より短いノイズのピークの高さが高くなることを見いだした。これは、加熱時間が短いほど、本来の塩基長まで伸長しきれていない DNA 断片が多いことを示唆している。通常の PCR サイクルでは加熱時間がほとんどないので、サイクル数を重ねるごとに、本来の塩基長より短い PCR 産物が多く生じていると考えられる。これは、本研究で作成したシミュレーションプログラムの前提である「PCR のサイクルにおいて伸長反応は 100% 完了する」から逸脱しており、正しくスタター比率をシミュレーションできていない可能性が示唆された。

さらに、途中まで伸長した DNA 断片が PCR に及ぼす影響について検討した。フォワードプライマーを、プライマー領域に繰返し配列を n 回（n は 0~9）追加したオリゴヌクレオチドに変更し、スタター発生率を測定したところ、スタター比率は n が大きくなるにつれて上昇する傾向がみられた（図 6）。PCR の各サイクルにおいて本来の塩基長より短い PCR 産物が多く生じると仮定した場合、生じた PCR 産物の長さによってスタター発生率が異なることとなり、各サイクルで生じる PCR 産物の長さも異なることから、各サイクルでスタター発生率が異なると考えられる。

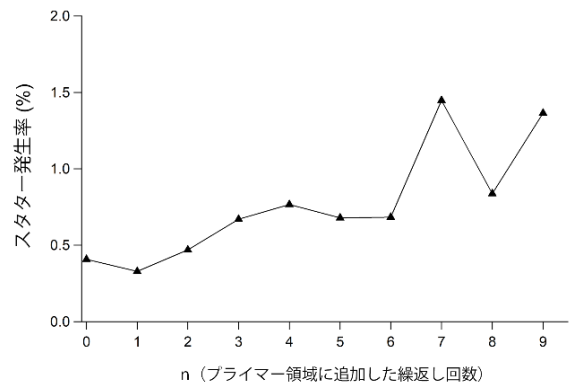


図6.プライマーの長さを変更したときのスタター発生率

この結果により、本研究課題で作成したシミュレーションプログラムは、前提としたモデルが現実と異なっている可能性が強く示唆された。本研究課題については、シミュレーションをこれ以上行うことは無意味と判断し、シミュレーションはここで打ち切った。しかし、今回のスタター発生率とシミュレーションの検討において得られた知見は、これまでに報告されたことのない新しいものであり、スタター発生メカニズムを理解するうえで非常に重要なものであると考えている。

#### (4)まとめ

本研究課題でスタター比率及びスタター発生率について検討をした結果、いずれについても予期せぬ結果となった。しかし、これらの結果は、今までほとんど明らかでなかったスタター発生メカニズムに関する新規の知見であり、現行のスタター発生メカニズムの理解を更新し得るものであると考えている。例えば、今までスタターの生じやすさはAT含有率という一次構造で決まるとされていたが、本研究ではAT含有率が同じでも並び順が異なるとスタター比率が異なる場合があったことから、一次構造だけでなく二次・三次構造も考慮に入れる必要があると考えられる。

今後の展望としては、本研究課題で得られた新たな知見についてより詳細な検討を行うことで、スタター発生メカニズムの理解を深め、新たなスタター発生モデルの構築が可能になると考えている。

#### <引用文献>

- ① Schlötterer C, Tautz D: Slippage synthesis of simple sequence DNA. *Nucleic Acids Res*, 20(2): 211-215, 1992.
- ② Brookes C, Bright JA, Harbison S, et. al.: Characterising stutter in forensic STR multiplexes. *Forensic Sci Int Genet*, 6(1): 58-63, 2012.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 深川 貴志、綿引 晴彦、三田 裕介、北山 哲史、藤井 宏治、水野 なつ子	4. 巻 30
2. 論文標題 反復配列の塩基構成とスタター比率に関する検討（第2報）	5. 発行年 2022年
3. 雑誌名 DNA多型	6. 最初と最後の頁 85-89
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 深川 貴志、綿引 晴彦、三田 裕介、北山 哲史、藤井 宏治、水野 なつ子	4. 巻 29
2. 論文標題 反復配列の塩基構成とスタター比率に関する検討	5. 発行年 2021年
3. 雑誌名 DNA多型	6. 最初と最後の頁 82-86
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 深川 貴志、綿引 晴彦、三田 裕介、北山 哲史、藤井 宏治、水野 なつ子
2. 発表標題 反復配列の塩基構成とスタター比率に関する検討（第3報）
3. 学会等名 日本DNA多型学会第32回学術集会
4. 発表年 2023年

1. 発表者名 深川 貴志、綿引 晴彦、三田 裕介、北山 哲史、藤井 宏治、水野 なつ子
2. 発表標題 反復配列の塩基構成とスタター比率に関する検討（第2報）
3. 学会等名 日本DNA多型学会第30回学術集会
4. 発表年 2021年

1. 発表者名 深川 貴志、綿引 晴彦、三田 裕介、北山 哲史、藤井 宏治、水野 なつ子
2. 発表標題 反復配列の塩基構成とスタター比率に関する検討
3. 学会等名 日本DNA多型学会第29回学術集会
4. 発表年 2020年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------