

令和 3 年 5 月 13 日現在

機関番号：14301

研究種目：挑戦的研究（萌芽）

研究期間：2019～2020

課題番号：19K22758

研究課題名（和文）ジェネオロジー理論を発展させた全ゲノム情報による多人数同時身元確認法の開発

研究課題名（英文）Development of a personal identification method by multiple kinships analysis using whole-genome information.

研究代表者

玉木 敬二（Tamaki, Keiji）

京都大学・医学研究科・教授

研究者番号：90217175

交付決定額（研究期間全体）：（直接経費） 4,900,000 円

研究成果の概要（和文）：本研究では3人以上が関与する複雑な血縁関係についてその有無に関する尤度比を算出して、血縁関係の判断ができるソフトウェア（KinBN）を開発した。ソフトウェアKinBNはネット上に挙げており、無償でダウンロードして利用できる。また、濃度やDNA断片化を段階的に変化させた模擬試料を作成して、ゲノム情報がどの程度まで判読可能かをマイクロアレイでのSNPs解析法を用いて検討した結果、ある程度の変性があっても第3度血縁までは判断できることがわかった。また、誤判定がなければSNPs数を2万5000まで減少させても、第3度血縁まで確実に判定できることがわかった。

研究成果の学術的意義や社会的意義

現行のDNA型検査に伴う遠縁や複数人の参照者を用いる複雑な血縁鑑定においても、多数ローカス間の連鎖を考慮した正確な尤度比が算出できるようになった。このため、在宅試料がなく異同識別検査では身元を確認できない大規模災害などで、多数の死亡者が出た際の身元確認に非常に有効となる。また、DNAマイクロアレイによるSNPsタイピングはある程度DNAが断片化した変性試料においても、血縁の判断が可能となった。このため、精度の高い血縁者検索法を提案することが可能となり、戦没者遺骨の身元確認のような変性した微量のDNA検査にも福音をもたらすと期待される。

研究成果の概要（英文）：We have developed software named KinBN that can determine the kinship by calculating the likelihood ratio about the presence of complex kinship such as involving more than two people. KinBN is now listed online and anyone can download and use it free of charge. We also prepared simulated denatured DNA samples in which the DNA concentration and fragmentation were changed stepwise, and examined how much the SNPs information was obtained using a microarray. The results showed that it was possible to determine up to the third-degree blood relationship even if there was some DNA denaturation. It was also found that even if the number of 25,000 SNPs can confirm the relationship up to the third-degree blood relationship unless there are no erroneous base calls.

研究分野：法医学

キーワード：DNA多型 ジェネオロジー 血縁鑑定 身元確認 SNPs ICS

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

2011年に発生した東日本大震災では、身元不明遺体に関して身元鑑定を行う際に、第1度血縁(親子、兄弟姉妹)よりも離れた血縁関係にあたる方からの血縁関係の有無の判断は技術的に困難であるためDNA試料の提供はされなかった。この状況を改善するため、われわれは網羅的SNPs解析を利用した二者間の血縁鑑定法(Morimoto C: PLoS One 2016)を開発したが、同法では常染色体に存在する174,254個のSNPsを解析し、染色体共有指標(Index of Chromosome Sharing: ICS)という二者間の染色体共有度合いを示す値を求め、その値によって血縁関係を判断する。ICSを利用すると、第五度血縁者(またいとこ)でも非血縁者との鑑別が高精度で可能であると期待される。しかし、これまでの研究ではSNPsタイピングに比較的多量の高分子DNAを必要とするため、大規模災害などの身元確認に用いられるDNAは微量で低分子化しており検査ができないおそれがある。また、比較対象が多人数存在する場合は評価方法が複雑であり、SNPsを用いる方法は1対1の2者間での血縁鑑定に留まっている。

一方、近年遺伝子検査ビジネスによるジェネオロジー(genealogy:系図学)が注目されている。また、古代人の全ゲノムシーケンス(Carpenter ML. Am J Hum Genet. 2013)の報告など、微量変性試料のDNA鑑定に関しても、ゲノム解析技術の進歩により克服されつつある。そこで、ジェネオロジーを参考に、ネットワーク学習系のAIを介在した、全ゲノム情報による血縁判定プログラムを作成して、法医遺伝学領域に導入すれば、不明者に対してその家族と疑われる複数人を同時に評価できる精度の高い血縁鑑定システムができるという発想に至った。

2. 研究の目的

(1)大規模災害時や孤独死における身元不明遺体の身元確認には、DNA型による血縁鑑定が利用される。2019年4月より実務導入されたDNA鑑定法は、マイクロサテライト(STR)21ローカスの繰り返し数をキャピラリー電気泳動にて型判定する検査(GlobalFiler検査)であるが、少子高齢化に伴い、第1度血縁者が存在しない事例の増加が予想され、孫やいとこなど、より離れた血縁関係の鑑定精度向上が求められている。さらに、大規模災害など死亡者が多人数存在する場合には、ある不明者と複数家族のスクリーニングだけでなく、不明者と該当家族内の遠縁を含めた複数の搜索依頼者(例:叔父と従兄弟)との血縁判定が必要となるが、その評価方法は確立していない。そこで、本研究では、まず、この新しく導入されたDNA検査法の結果を用いて、3人以上が関与する複雑な血縁関係についてもその有無に関する尤度比を算出して、血縁関係の判断ができるソフトウェアの開発を目的とする。

(2)ジェネオロジーにみられるように次世代シーケンシング(NGS)(法医学分野ではMPS: massive parallel sequencingという)による全ゲノム情報から、血縁鑑定を行う技術が発展している。しかし、実際の分析は2者間に限定されており、その分析精度も低いため実際の法医実務には使えない。また、われわれの開発したDNAマイクロアレイを使った一塩基多型解析による血縁鑑定法は有力な方法であるが、DNA量が少なく変性している法医試料における検討には至っていない。そこで、本研究ではジェネオロジー理論をさらに発展させ、MPSまたはマイクロアレイによるゲノム情報から精度の高いSNPsハプロタイプを比較することにより、身元不明者のスクリーニングと該当家族内の複数の搜索依頼者の血縁関係が判定できるシステムの構築を行う。遺体と家族の塩基配列レベルの情報を多人数間で比較して血縁者を検索できれば、これまでの法医領域には全くない新しいシステムとなり、大規模災害時などにおける多数の個人識別が迅速に行えることが期待される。

3. 研究の方法

(1)現行のGlobalFiler検査キットのような多数のSTRローカスのタイピングにおけるローカス間の連鎖についても考慮した尤度比が算出できる血縁鑑定のソフトウェアを開発する。また、単に尤度比の算出のみでなく、問題となる人や複数の参照者が関わる血縁関係のSTR型を多く合成して、尤度比を算出するシミュレーションを行い、ある検査システムにおける件の血縁関係の尤度比の分布を示して、その検査システムを評価できる機能を付加するよう工夫する。

(2)法医学試料からの全ゲノムシーケンスに向けた検討

DNA 量が 3ng から 200ng まで 4 段階に変化させた模擬試料や、DNA を 150 塩基程度まで超音波切断した模擬変性試料から、MPS によって全ゲノムシーケンスがどの程度まで判読可能かをリード数、マップ率などを検討する。この結果を基に、実際の身元確認で使用される検体(骨、歯、爪)数例から DNA を抽出して、全ゲノム解析を行いどの程度のデータが得られるか確認する。もし、法医学試料から信頼性のある全ゲノムの一塩基多型(SNPs)や塩基配列のデータが得られない場合でも、解析する領域を限定したり、DNA マイクロアレイの代替による SNPs 判定によりハプロタイプデータが得られるので、法医学試料からの第 3 度から第 5 度までの遠縁の血縁鑑定に向けた基礎資料は得られる。

法医学における身元確認では厳密な精度が要求されるので、血縁者であるかどうかの判定については、われわれの研究で提唱した ICS を用いる方法をさらに発展させて、不明者に対してその家族と疑われる複数人のデータを同時に評価できる新しい指標を考案する。以上のシミュレーション及び統計解析にはプログラミング言語 R または Python によるプログラムを使用する。

4. 研究成果

(1)ローカス間連鎖を考慮した血縁鑑定ソフトウェアの開発

これまで血縁鑑定法のソフトウェアは有償無償を含めて数種類開発されているが、現行の GlobalFiler 検査のように多数の STR ローカスを用いるシステムには不正確な結果を提供するおそれがある。というのも、多数のローカスを用いるとローカス間の遺伝的距離が近くなり、連鎖を生じるため各ローカスから算出された尤度比をそのまま掛け合わせるができないためである。GlobalFiler 検査は常染色体 21 ローカスをタイピングしているが、ローカス間に連鎖のあるものが 3 組ある。この連鎖を考慮しながら尤度比を算出する過程は煩雑だが、ベイジアンネットワーク(BN)を用いることで解決した。但し、BN では出力が事後確率として表されるのに対し、事前確率が必ずしも明確ではない身元確認や血縁鑑定では尤度比(LR)が求められるので尤度比になるようプログラム上で変換した。また、血縁鑑定における尤度比算出の 2 つの対立仮説は、「H1: A と B は兄弟である。」に対し、「H2: A と B はいどこである」というように、H2 が必ずしも件の両者は非血縁であるというものではないため、そのような尤度比が算出されるようプログラムを工夫した。プログラムは統計解析ソフト R を用いて作成し、また、使用者が使いやすいように GUI 化をはかった。完成したソフトウェア(KinBN)についてその算出結果の検証を行った。米国国立標準技術研究所が提供する血縁データを用い、他の既存のソフトウェア(FamLink)と比較して、ローカス間の連鎖のある場合においても尤度比を正確に算出できることを確認した。さらに、KinBN では様々な複数人が関与する血縁関係におけるシミュレーションを行い、尤度比の分布をグラフ化できる機能も付加した。このため、使用した検査キットの血縁鑑定の総合的な判別力を評価することができる。ソフトウェア KinBN はネット上(<https://github.com/ChieMorimoto/KinBN/releases>)に挙げており、誰でも無償でダウンロードして利用できる。

(2)MPS の検出とタイピングにおける課題と対応

当初は MPS によってヒトゲノム中の SNPs を判定する予定であったが、MPS の実験を進めていく過程で MPS の深刻な技術的障壁が見つかった。まず、法医学分野において標準とされる DNA 検査はマイクロサテライト(STR)の型判定であるが、MPS による STR 型判定が予想外に難航した。STR の領域を PCR 増幅してライブラリーを作成する際、PCR 増幅によるスターなどの非特異 PCR 産物の増加による対象シーケンス情報に対する妨害が非常に深刻であった。また、膨大なシーケンス情報源から STR 領域を切り取って型判定するソフトウェアを複数種類使用したが、それらの STR に関するその処理能力は非常に低かった。パラメータの変更など工夫を凝らしても、出力結果はこれまでのキャピラリー電気泳動法による STR 型との整合性がないだけでなく、多くのローカスにおいて、同一と考えられる塩基配列が切り取り方が異なるため、違うアレルのように検出報告されてしまった。このため、判定結果の信頼性は非常に低く、MPS による利点と考えられた同一長さのアレルの塩基配列による区別を利用したヘテロ接合度の高い STR 型システムにおけるアレル頻度調査などの作業が全く進まない状況であった。国際学会においても、MPS における STR アレルの命名法について議論が進んでいるが、未だにデータベースの集積が滞っているのは命名法だけでなく、このような技術的課題が克服されていないのも原因であろうと思われる。MPS による SNPs 型判定では、STR に比べてこのような課題は多少軽減されると考えられるが、そもそも微量試料を基本とする法医学試料において、MPS 自体の基本的性能が十分に追いついている

かどうか大きな懸念となっしまい、MPSの費用対効果を勘案すると、まずマイクロアレイでのSNPs検討を経るべきであろうと判断し、マイクロアレイを用いたSNPs解析法を優先することにした。

(3)DNA変性試料がタイピング結果に及ぼす影響

SNPsタイピング成功率の低下

われわれの開発したDNAマイクロアレイによる網羅的SNPs解析による血縁鑑定法では、ICSを算出するが、法医鑑識領域でしばしば遭遇する劣化DNA試料におけるICSの値の影響をさらに明らかにするため、模擬劣化試料を作成して分析した。血液抽出DNAを超音波処理装置(Covaris Acoustic Solubilizer)により断片化し、まず3段階(DNA断片の中央値が450bp、300bp、150bp)の劣化DNA試料と作成し、各試料を希釈して検査DNA量として4段階の試料(3ng、12.5ng、50ng、200ng)を実験に用いた。各試料はエレクトロフェログラムを観察してその断片化を確認した。これらの試料を用いて現行の法医鑑識で用いられているGlobalFiler検査を行った。その結果、450bpの断片化試料では検査DNA量に関係なく全てのローカスで正確なタイピングができたが、300bpの断片化試料では比較的長い増幅領域を必要とする1ローカス(TPOX又はSE33)において、アリルドロップアウトが生じてアリルが検出されなかった。さらに、150bpの断片化試料では、4ローカス(CSF1PO、TPOX、D18S51、SE33)では全ての試料でアリルが検出されず、アリルが検出されたローカスでも、ヘテロ接合体の2つのうち片方がドロップアウトしているなど、全ての試料でアリルドロップアウトがみられた。したがって、現状のSTR型検査においてはDNA量よりも断片化によってドロップアウトが生じて誤判定に至る可能性が示唆された。

DNAが断片化した劣化試料において、これまでの予備実験ではSNPsタイピング成功率(call rate: CR)は全て99%を超えていたが、これは十分量のDNA量の場合であり、さらに少量化することにより、CRが低下する可能性がある。このため、前記DNAの断片化や少量化の調整を行った後、Infinium HD assayのHumanCore 24 Beadchip (Illumina)を用いてSNPsを解析した。その結果、DNA断片化をしない試料であってもDNA量を減らすと段階的にCRは低下し、3ngになると98.8%となった。この傾向は断片化した試料ではより顕著になり、450bpの断片化試料では十分量(200ng)であれば98.5%であったが、3ngでは88.0%となり、300bp断片化試料では74.9%までに下がった。特に150bp断片化試料では、十分量であってもCRは87.1%に留まり、少量化するにつれて低下し、3ngでは43.8%にまで低下した。

SNPsタイピング誤判定率の上昇

これまで、CRの低下のみを劣化試料におけるICS値の低下の因子と考えていたが、今回の検討によりCRとDNA断片化のない試料による型判定結果との不一致率(誤判定率 error rate: ER)はより緊密に関連することがわかった。模擬劣化試料ではCRが92%まではERは0.6%未満に留まっていたが、CRが低下するにつれてERは増加した。この結果について線形回帰分析を行ったところ、CRが90%でERは1.5%となり、CRが50%ではERは10.0%に増加することがわかった。

このデータを基に108組の実際の血縁ペアの一方のSNPsデータに対してCRとERの値を変化させてICSの値の変化を観察した。具体的には(1-CR)(%)分の無作為のSNPsデータを“未検出”としてデータを削除し、また、CR値に対応したER(%)分のSNPsの1塩基を変化させ、そのSNPsデータの一部を基に実際の血縁ペアからICS値の分布図を改めて作成した。そして、用意したテスト検体についてICS値を算出して、この分布図から確率密度を算出し、事後確率(W)を算出し、Wが90%以上の場合に、件の血縁関係を支持するとした。シミュレーションには統計解析ソフトRによる独自のプログラムを使用した。その結果、傍系(兄弟、いとこなど)では第3度血縁(いとこ)までであれば、CRが95%以上あれば、判断できることがわかった。

判定SNPs数の変化による血縁鑑定精度の変化

また、現行のSNPsマイクロアレイキットではなく、より少数のSNPs数を判定するキットを用いた場合にICS値による血縁鑑定の有効性がどのように変化するのか検討した。血縁関係が既知の2者の一方のSNPsデータを無作為に削除してSNPs数を減少させて算出したICS値について、前項と同様にある血縁におけるICS値の分布図を作成して、テスト検体のICS値と比較することにより事後確率を算出して判定した。但し、この場合ERは0とした。その結果、SNPs数を2万5000まで減少させても、第3度血縁(いとこ)までは確実に判定できることがわかった。

SNPsマイクロアレイを用いた方法は、検査するDNA量が少量であってもある程度のCRが得られればこれまで困難であった第3度血縁など遠縁の鑑定においても判定可能であり有効な方法であることが確認された。現在、ERが低い場合には既存のハプロタイプデータから、試料では判定できなかった

SNPsの genotype を推定する genotype imputation を始めており、血縁鑑定精度の向上にどの程度貢献するか検討する予定である。

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 2件）

1. 著者名 真鍋翔, 濱保英樹, 小谷泰一, 宮尾昌, ジェマイル・レイラ, 玉木敬二.	4. 巻 27(1)
2. 論文標題 GlobalFiler 21 ローカスにおけるDNA型的一致率と混合試料の数学的評価について.	5. 発行年 2019年
3. 雑誌名 DNA多型	6. 最初と最後の頁 137-142.
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 オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 平井瑛里子, 橋谷田真樹, 森本千恵, 藤本駿太郎, 濱野悠也, 尾関宗孝, 玉木敬二.	4. 巻 27(1)
2. 論文標題 Precision ID GlobalFiler™ NGS STR Panelを利用した次世代シーケンスによる混合試料解析の検討.	5. 発行年 2019年
3. 雑誌名 DNA多型	6. 最初と最後の頁 132-136.
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 Morimoto C, Tsujii H, Manabe S, Fujimoto S, Hirai E, Hamano Y, Tamaki K.	4. 巻 47
2. 論文標題 Development of a software for kinship analysis considering linkage and mutation based on a Bayesian network.	5. 発行年 2020年
3. 雑誌名 Forensic Sci Int Genet.	6. 最初と最後の頁 102279
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.fsigen.2020.102279	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

〔学会発表〕 計6件（うち招待講演 0件／うち国際学会 0件）

1. 発表者名 Chinyi Cheng, 真鍋翔, 玉木敬二, 山田亮.
2. 発表標題 User-friendly visualization of global DNA identification matching in multi-victim disaster.
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

1. 発表者名 大内 司, Guan Xueting, 平井瑛里子, 眞鍋 翔, 大林将弘, 橋谷田真樹, 赤根 敦, 安達 登, 玉木敬二, 舟山真人.
2. 発表標題 Precision ID GlobalFiler NGS STR Panel v2による日本人データベースの構築(第1報).
3. 学会等名 第104次日本法医学会学術全国集会.
4. 発表年 2020年

1. 発表者名 花村天斗, 眞鍋 翔, 藤本駿太郎, 平井瑛里子, 尾関宗孝, 玉木敬二
2. 発表標題 集団遺伝学モデルに基づいたY-STRによる血縁鑑定に関する検討
3. 学会等名 第104次日本法医学会学術全国集会.
4. 発表年 2020年

1. 発表者名 眞鍋 翔, 平井瑛里子, 藤本駿太郎, 花村天斗, 玉木敬二.
2. 発表標題 GlobalFiler 21ローカスのシーケンス情報を利用した混合試料の関与人数の推定.
3. 学会等名 日本DNA多型学会第28回学術集会.
4. 発表年 2019年

1. 発表者名 平井瑛里子, 眞鍋 翔, 藤本駿太郎, 橋谷田真樹, 玉木敬二.
2. 発表標題 Massively Parallel SequencingにおけるSTR型解析ツールの判定精度について.
3. 学会等名 日本DNA多型学会第28回学術集会.
4. 発表年 2019年

1. 発表者名 真鍋 翔, 花村 天斗, 藤本駿太郎, 平井瑛里子, 玉木敬二.
2. 発表標題 母集団内に存在する同一Y-STR数の推定法に関する検討.
3. 学会等名 日本法科学技術学会第25回学術集会.
4. 発表年 2019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
---------	---------