Forensic identification from mixed and dried blood spots using single-cell genomic technique

Yukiko Dozen, Yohei Miyake, Arisa Hayashi, Atsushi Nagai, Tomomi Michiue
Department of Legal Medicine, Graduate School of Medicine, Gifu University, Gifu, Japan

Introduction

Forensic identification of mixed samples is an important issue in forensic science. We analyzed DNA types of a single leukocyte isolated from mixed and dried blood spots using a single-cell genomic technique. A flow cytometer (FCM) was used for isolating single cells. We also performed whole genome amplification (WGA) for analyzing a small amount of DNA from a single cell.

Materials and methods

An outline of the methods used in this study is shown in Figure 1.

(1) **Preparation of mixed blood:** Peripheral blood samples were obtained from three Japanese donors named A, B, and C (one male and two females). These blood samples were mixed in equal amounts.

(2) **Preparation of dried blood spots:** Blood spots were prepared in Petri dishes and dried at room temperature.

(3) **Rehydration:** The blood spots were crushed using a bead crusher (Bio Medical Science) and rehydrated with distilled water. The samples were filtered using 70-µm and 40-µm cell strainers.

(4) **Isolation of single cells:** A lymphocyte was fluorescently labeled using an anti-CD3 antibody and isolated using a Cell Sorter SH800 (Sony). Figure 2 shows an example of a sorting gate to collect single cells.

(5) **WGA:** DNA from every cell isolated using FCM methods was amplified using the REPLI-g Single Cell Kit (Qiagen) according to the manufacturer's method.

(6) **Polymerase chain reaction (PCR):** DNA was purified using the QIAquick PCR Purification Kit (Qiagen). Fifteen autosomal short tandem repeats (STRs) and amelogenin were amplified using the AmpFlSTR Identifiler Plus PCR Amplification Kit (ABI).

(7) **Analysis:** After PCR, DNA was electrophoresed using the ABI PRISM 310 Genetic Analyzer (ABI) and genotyped using GeneMapper ID v.3.2 (ABI). The threshold was set at 50 RFU.

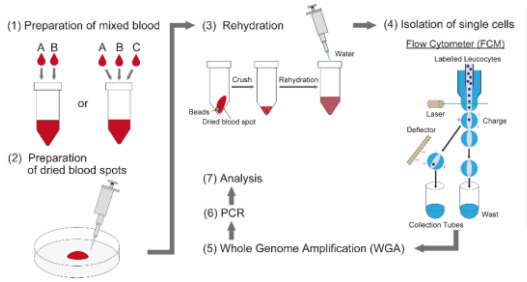


Figure 1: An outline of methods used in this study

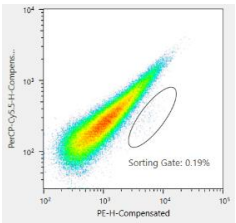


Figure 2: A fluorescence dot plot and a sorting gate on SH800 for dried and labeled samples

Results and discussion

Tables 1 and 2 show the analysis results of the leukocyte samples obtained from mixed and dried blood spots of two (A and B) and three donors, respectively. There were no lymphocyte samples in which all 16 loci were correctly typed. DNA types derived from one of the two donors were determined in 44 of 48 lymphocyte samples (Table 1). The average determination rate was approximately 30% and the maximum determination rate was approximately 75%. DNA types derived from one of the three donors were also determined in 37 of 40 lymphocyte samples (Table 2). The average determination rate was approximately 26% and the maximum determination rate was approximately 56%.

In the samples from two donors, the average rate of correctly detected alleles was approximately 46% and the maximum rate was approximately 77% (Table 1). However, in the samples from three donors, the average rate of correctly detected alleles was approximately 41% and the maximum rate was approximately 72% (Table 2).

Figure 3 shows peak height imbalances in heterozygous genotypes. The height of the allele peaks of the heterozygous genotype was almost the same in the sample without WGA (Figure 3A). However, there were imbalances in the heights of allele peaks in the samples with WGA (Figure 3B), and some loci showed allele dropouts in which only one of the heterotypic alleles was detected.

Because of these allele dropouts, the analysis results focusing on loci showed low determination rates. However, even if the number of correctly typed loci was small, we could detect donors when the number of detected alleles was large or when specific alleles were detected. Therefore, we also focused on the number of detected alleles. As a result, the analysis focusing on the number of alleles showed a higher detection rate than the analysis focusing on loci.

In this study, we did not have a lymphocyte sample in which all 16 loci were correctly typed. However, even if the amount of blood is small, the number of lymphocytes present in it is huge; therefore, we believe that it is possible to obtain lymphocyte samples with a nearly complete DNA profile using the single-cell genomic technique.

A: without WGA

B: with WGA

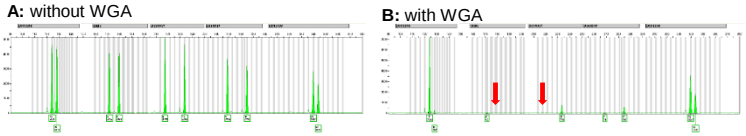


Figure 3: Peak height imbalance in heterozygous genotypes
A: Electropherogram of a blood spot sample without WGA
B: Electropherogram of a single leukocyte isolated from the blood spot of the same donor as A
Allele dropouts are indicated with arrows.

Table 1: Single-cell analysis results of mixed and dried blood spots isolated from two donors (A and B)

No.	Identified Donors	15 STR loci and amelogenin									Alleles		
		Correctly typed loci			Not correctly typed loci			Determination rate(%)			Number of alleles		Rate(%)
		Homo	Hetero	Total	Allele dropout	No peak	Total	Homo	Hetero	Total	Detected	Original	
1	B	2	2	4	9	3	12	100	14.29	25	15	30	50
2	B	2	2	4	7	5	12	100	14.29	25	13	30	43.33
3	B	2	3	5	5	6	11	100	21.43	31.25	13	30	43.33
4	A	1	3	4	5	7	12	33.33	23.08	25	12	29	41.38
5	A	1	2	3	7	6	13	33.33	15.38	18.75	12	29	41.38
6	B	2	6	8	7	1	8	100	42.86	50	21	30	70
7	B	2	3	5	4	7	11	100	21.43	31.25	12	30	40
8	B	2	10	12	1	3	4	100	71.43	**75**	23	30	**76.67**
9	B	2	3	5	4	7	11	100	21.43	31.25	12	30	40
10	A	3	4	7	5	4	9	100	30.77	43.75	16	29	55.17
11	A	2	0	2	8	6	14	66.67	0	12.5	10	29	34.48
12	B	1	6	7	6	3	9	50	42.86	43.75	19	30	63.33
13	B	2	8	10	4	2	6	100	57.14	62.5	22	30	73.33
14	B	2	5	7	6	3	9	100	35.71	43.75	18	30	60
15	A	2	4	6	7	3	10	66.67	30.77	37.5	17	29	58.62
16	B	2	7	9	3	4	7	100	50	56.25	19	30	63.33
17	A	3	1	4	9	3	12	100	7.69	25	14	29	48.28
18	B	1	0	1	8	7	15	50	0	6.25	9	30	30
19	A	3	6	9	5	2	7	100	46.15	56.25	20	29	68.97
20	A	3	2	5	6	5	11	100	15.38	31.25	13	29	44.83
21	-	Mixed											
22	B	1	2	3	8	5	13	50	14.29	18.75	13	30	43.33
23	B	1	1	2	4	10	14	50	7.14	12.5	7	30	23.33
24	B	1	1	2	6	8	14	50	7.14	12.5	9	30	30
25	A	0	3	3	3	10	13	0	23.08	18.75	9	29	31.03
26	-	Mixed											
27	B	2	3	5	3	8	11	100	21.43	31.25	11	30	36.67
28	B	2	2	4	7	5	12	100	14.29	25	13	30	43.33
29	A	3	2	5	3	8	11	100	15.38	31.25	10	29	34.48
30	B	2	3	5	4	7	11	100	21.43	31.25	12	30	40
31	-	Mixed											
32	B	1	2	3	6	7	13	50	14.29	18.75	11	30	36.67
33	B	1	3	4	9	3	12	50	21.43	25	16	30	53.33
34	-	Mixed											
35	B	1	2	3	7	6	13	50	14.29	18.75	12	30	40
36	B	2	3	5	6	5	11	100	21.43	31.25	14	30	46.67
37	B	1	3	4	8	4	12	50	21.43	25	15	30	50
38	A	2	1	3	5	8	13	66.67	7.69	18.75	9	29	31.03
39	B	2	2	4	7	5	12	100	14.29	25	13	30	43.33
40	B	1	1	2	6	8	14	50	7.14	12.5	9	30	30
41	B	1	3	4	5	7	12	50	21.43	25	12	30	40
42	B	2	6	8	6	2	8	100	42.86	50	20	30	66.67
43	B	2	1	3	9	4	13	100	7.14	18.75	13	30	43.33
44	B	1	2	3	9	4	13	50	14.29	18.75	14	30	46.67
45	A	2	4	6	7	3	10	66.67	30.77	37.5	17	29	58.62
46	B	2	2	4	6	6	12	100	14.29	25	12	30	40
47	A	2	2	4	7	5	12	66.67	15.38	25	13	29	44.83
48	B	0	2	2	6	8	14	0	14.29	12.5	10	30	33.33
Average		1.70	3.02	4.73	5.98	5.30	11.27	75	22.02	29.55	-		46.21

"Determination rate" indicates the percentage of correctly typed loci out of 16 loci containing amelogenin.
"Original" indicates the number of alleles containing amelogenin of donor.
Maximum percentages of determination rate and rate are indicated in red bold.

Table 2: Single-cell analysis results of mixed and dried blood spots isolated from three donors (A, B, and C)

No.	Identified donors	15 STR loci and amelogenin									Alleles		
		Correctly typed loci			Not correctly typed loci			Determination rate(%)			Number of alleles		Rate(%)
		Homo	Hetero	Total	Allele dropout	No peak	Total	Homo	Hetero	Total	Detected	Original	
1	A	3	1	4	5	7	12	100	7.69	25	10	29	34.48
2	B	2	3	5	7	4	11	100	21.43	31.25	15	30	50
3	B	1	2	3	4	9	13	50	14.29	18.75	9	30	30
4	B	1	2	3	6	7	13	50	14.29	18.75	11	30	36.67
5	A	2	0	2	3	11	14	66.67	0	12.5	5	29	17.24
6	C	3	4	7	3	6	9	75	33.33	43.75	14	28	50
7	A	2	7	9	5	2	7	66.67	53.85	**56.25**	21	29	**72.41**
8	B	1	2	3	6	7	13	50	14.29	18.75	11	30	36.67
9	A	3	6	9	4	3	7	100	46.15	**56.25**	19	29	65.52
10	-	Not detected											
11	B	2	2	4	5	7	12	100	14.29	25	11	30	36.67
12	A	0	5	5	6	5	11	0	38.46	31.25	16	29	55.17
13	B	1	1	2	5	9	14	50	7.14	12.5	8	30	26.67
14	B	1	5	6	6	4	10	50	35.71	37.5	17	30	56.67
15	A	2	2	4	6	6	12	66.67	15.38	25	12	29	41.38
16	A	1	3	4	6	6	12	33.33	23.08	25	13	29	44.83
17	B	2	3	5	3	8	11	100	21.43	31.25	11	30	36.67
18	A	0	2	2	2	12	14	0	15.38	12.5	6	29	20.69
19	B	2	1	3	6	7	13	100	7.14	18.75	10	30	33.33
20	B	2	0	2	3	11	14	100	0	12.5	5	30	16.67
21	B	2	2	4	8	4	12	100	14.29	25	14	30	46.67
22	A	1	5	6	4	6	10	33.33	38.46	37.5	15	29	51.72
23	-	Mixed											
24	C	1	1	2	5	9	14	25	8.33	12.5	8	28	28.57
25	B	2	2	4	5	7	12	100	14.29	25	11	30	36.67
26	A	3	0	3	7	6	13	100	0	18.75	10	29	34.48
27	B	2	2	4	5	7	12	100	14.29	25	11	30	36.67
28	C	1	2	3	7	6	13	25	16.67	18.75	12	28	42.86
29	C	3	2	5	6	5	11	75	16.67	31.25	13	28	46.43
30	B	1	3	4	8	4	12	50	21.43	25	15	30	50
31	B	2	1	3	8	5	13	100	7.14	18.75	12	30	40
32	-	Only common alleles were detected											
33	B	2	4	6	6	4	10	100	28.57	37.5	16	30	53.33
34	B	2	3	5	9	2	11	100	21.43	31.25	17	30	56.67
35	A	1	1	2	3	11	14	33.33	7.69	12.5	6	29	20.69
36	B	2	2	4	5	7	12	100	14.29	25	11	30	36.67
37	B	2	1	3	8	5	13	100	7.14	18.75	12	30	40
38	B	2	1	3	9	4	13	100	7.14	18.75	13	30	43.33
39	B	0	4	4	7	5	12	0	28.57	25	15	30	50
40	B	2	5	7	5	4	9	100	35.71	43.75	17	30	56.67
Average		1.68	2.49	4.16	5.57	6.27	11.84	70.72	18.53	26.01	-		41.43

"Determination rate" indicates the percentage of correctly typed loci out of 16 loci containing amelogenin.
"Original" indicates the number of alleles containing amelogenin of donor.
Maximum percentages of determination rate and rate are indicated in red bold.

Conclusion

Single-cell genome analysis is an effective method for identifying individuals from mixed and dried blood spots.

Email: legalmed@gifu-u.ac.jp (Y. Dozen)

All authors declare that they have no conflict of interest.

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

INSTITUT FÜR RECHTSMEDIZIN

Geschickt kombiniert zum Ziel

Katja Anslinger, Birgit Bayer, Marta Diepenbroek

Zusammenfassung

Die Gesetzesänderung des §81e der StPO ermöglicht den Strafverfolgungsbehörden in Deutschland seit Ende 2019 molekulargenetische Untersuchungen zur Vorhersage äußerlich sichtbarer Merkmale (konkret Augen-, Haar- und Hautfarbe) sowie des biologischen Alters eines unbekannten Spurenverursachers zu beauftragen. Bis August 2021 war es im Freistaat Bayern gemäß Polizeiaufgabengesetz darüber hinaus unter speziellen Voraussetzungen möglich, auch Untersuchungen zur biogeographischen Herkunft durchzuführen. Wie die Erfahrung der letzten Monate zeigte, existierten jedoch in einigen Fällen, in denen diese Analysen angefordert wurden, nur Mischspuren. Mischspuren, aus denen das autosomale oder Y-chromosomale DNA-Profil einer fremden Person abgeleitet werden konnte, die aufgrund ihrer Komplexität bzw. Mischungsverhältnisse jedoch für weiterführende DNA-Analysen nicht geeignet waren. Mehrere solcher Fälle konnten in unserem Labor durch die Kombination von Einzelpartikelanalyse zur gezielten Suche einer geeigneten Hautschuppe der entsprechenden Person und deren Phänotypisierung mithilfe eines sensitiven MPS-Panels erfolgreich untersucht werden. Vorgehensweise und Ergebnisse werden an Fallbeispielen beschrieben.

Einleitung

Immer wieder treten im Rahmen der Bearbeitung aktueller Fälle, aber auch bei der Aufarbeitung von Altfallen, Situationen auf, in denen Ermittlungen stocken und sich die ermittelnden Behörden neue Impulse durch die Ergebnisse einer forensischen DNA Phänotypisierung erhoffen. Was aber tun, wenn für eine erweiterte DNA-Analyse, die idealer Weise zum Erreichen einer größtmöglichen Vorhersagegenauigkeit an 1-Personen-Spuren durchgeführt wird, kein geeignetes Spurenmaterial zur Verfügung steht? Die folgenden Beispiele sollen einen Weg aus dieser Misere aufzeigen.

Fall 1 (Altfall aus 1997): Schon bei den ersten molekulargenetischen Analysen unmittelbar nach der Tat konnten autosomale bzw. Y-chromosomale Merkmale einer unbekannten Person an mehreren Spuren - ausnahmslos Mischspuren - nachgewiesen werden. Als Untersuchungsmaterial für die 2020 beauftragte Bestimmung von Augen-, Haar- und Hautfarbe wurde auf Faserfolien von der Bekleidung der Geschädigten bzw. dem Tatort zurückgegriffen.

Fall 2 (aktuelles versuchtes Tötungsdelikt): Bei der routinemäßigen Aufarbeitung der Spuren konnten an mehreren Stellen der systematisch und vollständig abgetasteten Oberbekleidung der Geschädigten neben ihren, bzw. den Merkmalen weiterer Familienmitglieder, die Merkmale einer unbekannten männlichen Person nachgewiesen und abgeleitet werden. Die Einzelpartikeluntersuchung wurde gezielt an den Folien der entsprechenden Bereiche durchgeführt. Die in Auftrag gegebene erweiterte DNA-Analyse umfasste sowohl Vorhersagen zu äußerlichen Merkmalen, als auch zur biogeographischen Herkunft.

Material und Methoden

Aus den Folienabtastungen wurden lichtmikroskopisch 153 (Fall 1) bzw. 340 (Fall 2) hautschuppenähnliche Partikel präpariert. Nach Extraktion mittels ReadyAmp™ (Promega) erfolgte die Quantifizierung der DNA durch RealTime PCR (Quantifiler® Trio DNA Quantification Kit, ThermoFisher). Proben mit humaner DNA wurden unter Verwendung des PowerPlex® ESX17 Fast Systems (Promega) typisiert. Die Vorhersagen zu Augen-, Haar- und Hautfarbe (Fall 1 und 2) sowie der biogeographischen Herkunft (nur Fall 2) erfolgten, jeweils in Doppelbestimmung, mithilfe des von uns entwickelten Customer Panels Ion AmpliSeq™ PhenoTrivium [2] unter Verwendung der Converge™ Software (beides ThermoFisher Scientific) bzw. des HIrisPlex-S-DNA-Phänotypisierungs-Webtool [3,4] und SNIPPER [5,6,7]. Zur Verbalisierung der Ergebnisse wird auf die Publikation zur Validierungsstudie des Panels [2] verwiesen.

Fall 1

braun	schwarz	dunkel
p = 0,998	p = 0,519	p = 0,878
	braun	
	p = 0,479	
	dunkel	
	p = 0,997	

Fall 2

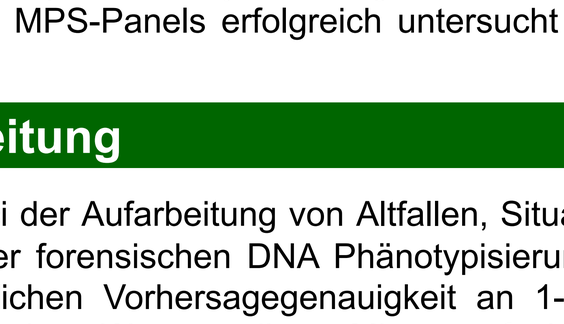
blau	blond	blass
p = 0,874	p = 0,436	p = 0,450
	braun	**inter**
	p = 0,417	p = 0,530
	hell	
	p = 0,901	

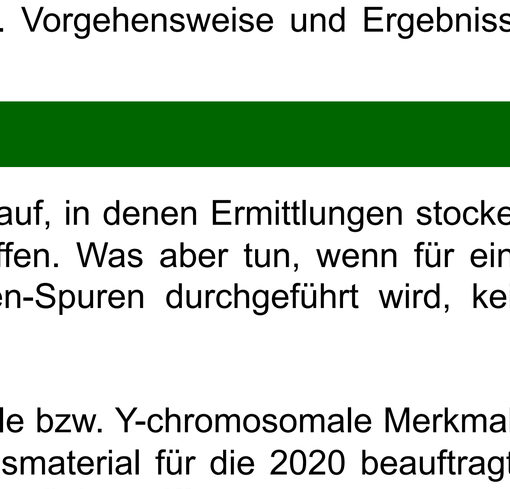
Abbildung 1: Erzielte Vorhersageergebnisse (jeweils höchste p-Values für Augen-, Haar- und Hautfarbe). Zur Veranschaulichung des vorhergesagten Phänotyps werden die Merkmale bildlich dargestellt. Der maximale AUC loss für eine Kategorie betrug 0,004 und hatte keinen Einfluss auf die Vorhersangen (Daten nicht gelistet) .

Ergebnisse

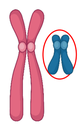
Von den 153 (Fall 1) bzw. 340 (Fall 2) Einzelpartikeln konnten nach STR-Analyse 7 (Fall 1) bzw. 3 (Fall 2) der jeweils gesuchten fremden Personen zugeordnet werden. Die Gesamtmenge an erhaltener DNA pro Probe variierte zwischen 150pg und 6.3ng (Fall 1) bzw. 50pg und 4.65ng (Fall 2). Für die weiterführende DNA-Untersuchung wurde auf die Proben mit dem jeweils höchsten DNA-Gehalt zurückgegriffen. Aus diesen konnten fast vollständige Konsensusprofile generiert werden (Fall 1: 39 der 41 HIrisPlex-S Marker; Fall 2: 40 der 41 HIrisPlex-S Marker; Fall 2: 161 der 163 autosomalen Abstammungsmarker bzw. 115 der 120 Y-chromosomalen Marker). Abbildung 1 gibt einen Überblick über die mithilfe des HIrisPlex-S Webtools erhaltenen Vorhersageergebnisse. Die Vorhersageergebnisse der biogeographischen Herkunft in Fall 2 sind der Abbildung 2 zu entnehmen.

Admixture | **PCA**





Väterliche Abstammung



Haplogruppe R

Unterhaplogruppe R-M269

Abbildung 2: Biogeografische Herkunftsvorhersage mittels Admixture-Analyse von Converge und PCA-Analyse von SNIPPER. Beide deuten auf eine europäische Abstammung hin. Die Analyse von Y-chromosomalen Markern weist auf Haplogruppe R, Linie M269 hin, die der vorhergesagten biogeographischen Abstammung entspricht.

Diskussion und Fazit

Obwohl der erste Fall nur die Vorhersage des Phänotyps umfasste, sind die erhaltenen p-Values und die damit korrespondierenden vorhergesagten phänotypischen Merkmale für Europa nicht üblich und könnten wertvolle Ermittlungshinweise darstellen. Der zweite Fall zeigte, dass der vorhergesagte Phänotyp (blaue Augen, hellbraunes Haar, blasse bis mittlere Haut) mit der vorhergesagten Abstammung (europäisch und Y-chromosomale Haplogruppe R-M269) übereinstimmt. Einmal mehr konnte die Technik der Einzelpartikelpräparation ihr Potenzial zur Lösung fast unlösbarer Fälle unter Beweis stellen.

Literatur

[1] Schneider H., Sommerer T., Rand S., Wiegand P. Hot flakes in cold cases. *Int. J. Legal Med.* 2011, 125, 543-548.

[2] Diepenbroek M., Bayer B., Schwender K., Schiller R., Lim J., Lagacé R., Anslinger K. Evaluation of the Ion AmpliSeq™ PhenoTrivium Panel: MPS-Based Assay for Ancestry and Phenotype Predictions Challenged by Casework Samples. *Genes*. 2020, 11, 1398

[3] Chaitanya, L.; Breslin, K.; Zuñiga, S.; Wirken, L.; Pośpiech, E.; Kukla-Bartoszek, M.; Sijen, T.; De Knijff, P.; Liu, F.; Branicki, W.; et al. The HIrisPlex-S system for eye, hair and skin colour prediction from DNA: Introduction and forensic developmental validation. *Forensic Sci. Int. Genet*. 2018, 35, 123–135.

[4] Breslin, K.; Wills, B.; Ralf, A.; Garcia, M.V.; Kukla-Bartoszek, M.; Pospiech, E.; Freire-Aradas, A.; Xavier, C.; Ingold, S.; De La Puente, M.; et al. HIrisPlex-S system for eye, hair, and skin color prediction from DNA: Massively parallel sequencing solutions for two common forensically used platforms. *Forensic Sci. Int. Genet*. 2019, 43, 102152.

[5] Phillips, C.; Salas, A.; Sánchez, J.J.; Fondevila, M.; Gómez-Tato, A.; Alvarez-Dios, J.; Calaza, M.; Casares de Cal, M.; Ballard, D.; Lareu, M.V.; et al. Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. Forensic Sci. Int. Genet. 2007, 1, 273–280.

[6] Pereira, R.; Phillips, C.; Pinto, N.; Santos, C.; Batista dos Santos, S.E.; Amorim, A.; Carracedo, A.; Gusmão, L. Straightforward Inference of Ancestry and Admixture Proportions through Ancestry-Informative Insertion Deletion Multiplexing. PLoS ONE 2012, 7, e29684.

[7] Fondevila, M.; Phillips, C.; Santos, C.; Freire-Aradas, A.; Vallone, P.M.; Butler, J.M.; Lareu, M.V.; Carracedo, A. Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies. Forensic Sci. Int. Genet. 2013, 7, 63–74.

Analysis of haplotype in mitochondrial DNA mixtures using LNA-mediated PCR clamping

Masaru Asari[1], Yuta Takahashi[1], Chisato Hoshina[1], Katsuhiro Okuda[1], Kanae Mori[1], Ryo Namba[2], Keiko Shimizu[1]
1 Department of Legal Medicine, Asahikawa Medical University, Asahikawa, Japan
2 Sapporo Higashi Tokushukai Hospital, Sapporo, Japan

[Objective]

Locked nucleic acid (LNA) has been widely used for various genetic analyses, and has many benefits, in terms of the specificity or sensitivity of amplification, because LNA-containing primers/probes form more stable duplexes with template DNA than probes lacking LNA. Here, we developed a new method for discriminating HV1 haplotypes from mitochondrial DNA (mtDNA) mixtures by applying PCR clamping using LNA. PCR clamping is based on the selective inhibition of amplification using LNA-containing probes, which can discriminate single-nucleotide differences.

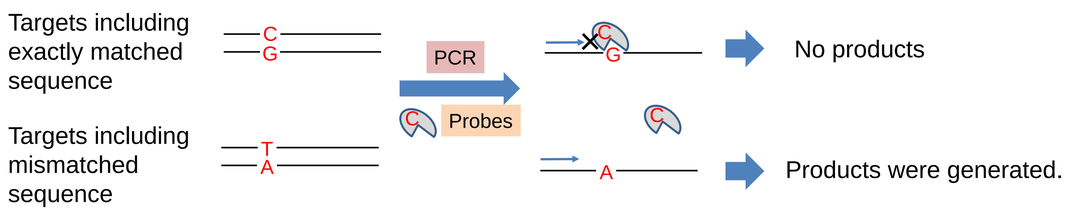


Fig.1 Inhibition of amplification using PCR clamping

[Method]

1. Sample

Buccal cells were collected from 100 healthy Japanese individuals and DNA was extracted from them using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). This study was approved by the Ethics Committee of Asahikawa Medical University.

2. Design of probes

We selected eight variations (16209C, 16217C, 16257A, 16261T, 16297C, 16298C, 16304C, 16362C) based on their frequencies in the Japanese population. Moreover, we added 16362T (the sequence without 16362C) because this sequence was particularly common. We designed seven probes with substitutions of LNA nucleotides (Table 1 and Fig .2), and each probe contained one or two target sites

Table 1 Design of LNA-containg probes

No.	Name	Sequence (5'-3')*
1	16209C	acaAgCaAgCaCaGcAaTc
2	16217C	aCaGcAaCcAaCcCtc
3	16257A/16261T	caAaGcAaCcTcTcAcCca
4	16297C/16298C	ccCacCcCCaAcaGtAc
5	16304C	cCtTaaCaGCaCaTaGTaC
6	16362C	cTtCtCgCcCcCaTgG
7	16362T	tCcAtGgGgAcGaGaAgG

*LNA nucleotides are in uppercase.

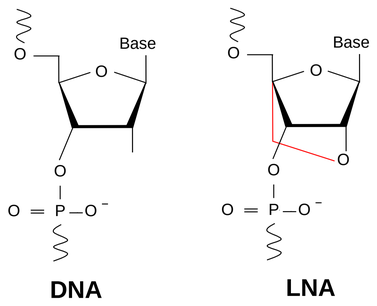


Fig. 2 Structure of DNA and LNA

3. Specificity of probes

Inhibition analysis using each probe except for probe 16362T was performed. Each extract was amplified in a 20-μL total reaction volume containing 300 copies of mtDNA, 0.25 μM each of fluorescently labeled forward and unlabeled reverse primers for the HV1 region, 0.5 μM of each probe, and 1 × QIAGEN Fast Cycling PCR Master Mix (Qiagen). The inhibitory effect of probe 16362T at indicated different concentrations was analyzed. The inhibitory effects of each probe were compared with the results from no probe. Detection of PCR products was performed on a 3500 Genetic Analyzer (Applied Biosystems). The data were analyzed using GeneMapper software v4.0 (Applied Biosystems).

4. Mixture analysis

We prepared mixed samples by mixing DNA from two individuals at a ratio of 1:9, 1:4, 1:1, 4:1, or 9:1, and then performed Sanger sequencing analysis after PCR clamping with each probe. We also prepared the mixtures of samples from three individuals with representative haplotypes at a ratio of 1:1:8 and performed sequencing analysis with/without indicated probes.

[Results and discussion]

1. Specificity of seven probes

The inhibition ratio was independently determined using each probe, except for Probe 16362T (Fig. 3). When the exactly matched templates were analyzed, the amplification was greatly inhibited, and the inhibition ratios were less than 0.1 for the six probes. In contrast, these ratios ranged from 0.43 to 0.80 when the mismatched templates were analyzed.

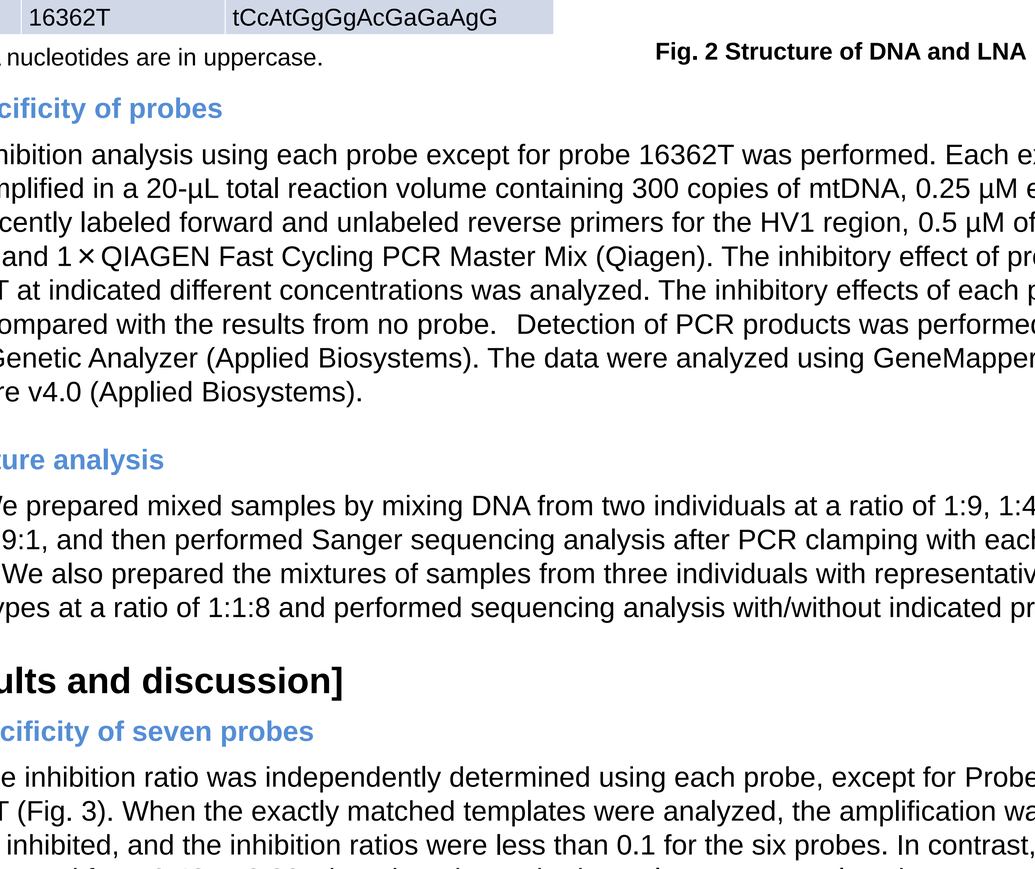


Fig. 3 Inhibitory effects of amplification using six designed probes. Mismatched or exactly matched template (300 copies of mtDNA) was amplified using probe 16209C (A), 16217C (B), 16257A/16261T (C), 16297C/16298C (D), 16304C (E), or 16362C (F). Amplification was performed using QIAGEN Fast Cycling Kit, and the extension was processed at 68ºC.

In the case of probe 16362T, the amplification using exactly matched templates was completely inhibited (Fig. 4A). However, extensive inhibition was also identified in the analysis of mismatched templates. To reduce this unexpected inhibition, we analyzed the effects of probe 16362T at a higher temperature (72ºC) in the extension step (Fig. 4B), and the amplification with mismatched templates was not inhibited and the inhibition ratios ranged from 0.88 to 1.06. In this study, the allele A on the complementary sequence was alternatively used in the design of probe 16362T, and high extension temperature (72ºC) also improved the specificity of amplification by this probe.

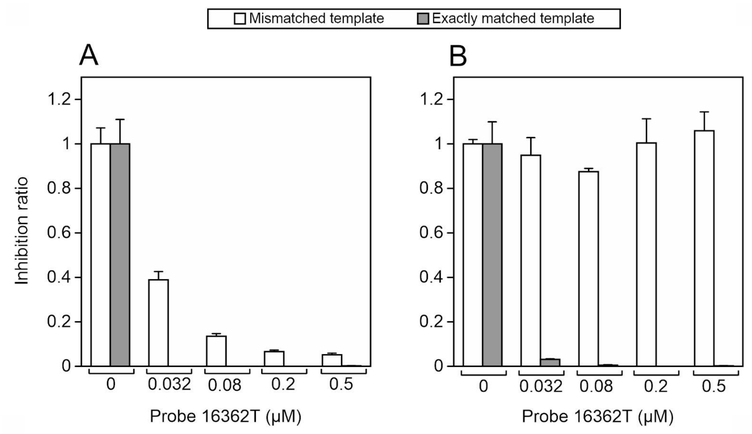


Fig.4 Inhibitory effects of amplification using probe 16362T. Mismatched or exactly matched template (300 copies of mtDNA) was amplified at 68ºC (A) or 72ºC (B) in the extension step. Amplification at the higher extension temperature was performed using AmpliTaq Gold 360 DNA Master Mix.

3. Mixture analysis

The results for mixed samples with material from two individuals (with DNA at a ratio of 1:1) are shown in Fig. 5. As expected, complete inhibition of exactly matched template was detected for each probe. At other mixing ratios (1:9, 1:4, 4:1, or 9:1), the exact matched sequences also showed inhibited amplification, and the mixed bases were not identified.

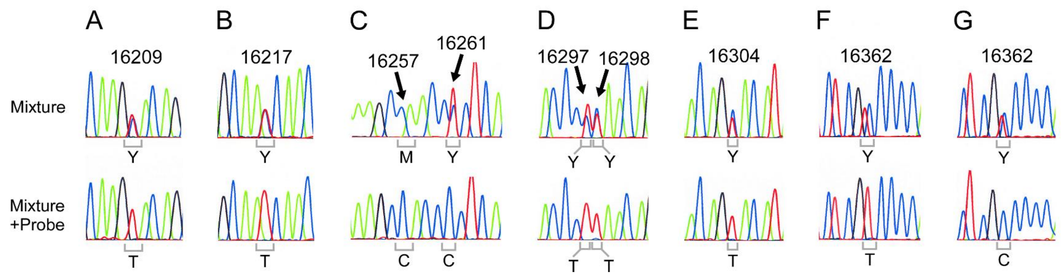


Fig. 5 Sequencing analysis of mixed samples containing material from two individuals. Mixed samples were independently amplified using probe 16209C (A), 16217C (B), 16257A/16261T (C), 16297C/16298C (D), 16304C (E), 16362C (F), or 16362T (G). Mixed bases (top row) were not observed by adding probes (bottom row).

Mixed samples from three individuals were detected with two selected probes, and representative results from three individuals mixed at a ratio of 1:1:8 are shown in Fig. 6. For example, the amplification using probe 16209C and probe 16362C would be expected to inhibit two haplotypes (16209C-16223T-16324C and 16245T-16362C, respectively). In fact, one haplotype (16129A-16183C-16189C-16223T-16297C-16298C) forming nucleotide mismatch by two selected probes was predominantly identified. At the mixing ratios of 1:1:18, 1:18:1, and 18:1:1, lower-intensity signals were identified and the haplotypes were not determined for some combinations of two selected probes (data not shown).

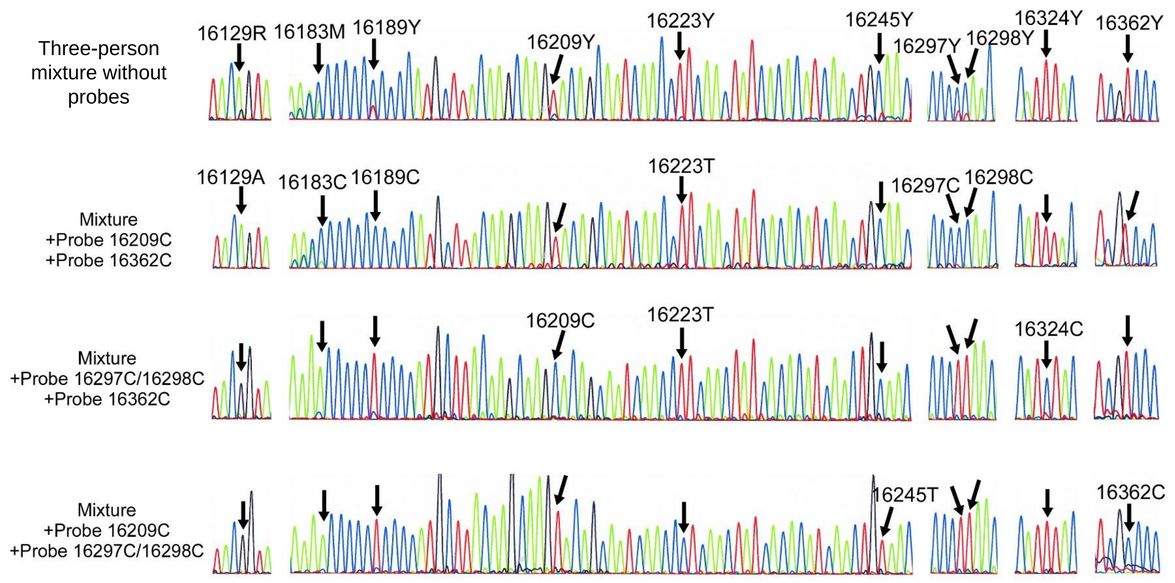


Fig. 6 Sequencing analysis of mixed samples containing material from three individuals with/without PCR clamping using two selected probes. This mixture was prepared by mixing three representative haplotypes (16245T-16362C, 16209C-16223T-16324C, and 16129A-16183C-16189C-16223T-16297C-16298C) at a ratio of 1:1:8. Arrowheads indicate the positions of nucleotide differences, and the variations including mixed bases are described above the arrowheads.

Analysis of the number of contributors to a mixed sample and the likelihood evaluation were performed based on autosomal STR profiles in mixed samples; a case with a Y-chromosomal STR mixture has also been introduced elsewhere. However, it would be practically difficult to interpret mixed profiles of mtDNA, although only mtDNA fragments might be successfully amplified. The ability to discriminate haplotypes could demonstrate that samples contain material from two individuals (the suspect and the victim) and provide additional information about the number of contributors to samples with material from three or more individuals. Therefore, this method to discriminate haplotypes in mixed samples could add the potential value of mtDNA analysis to criminal investigations.

[Conclusions]

We developed a new method for discriminating HV1 haplotypes from mitochondrial DNA mixtures by applying PCR clamping using LNA. Seven probes discriminate nucleotide differences and haplotypes were determined in two- or three-person mixtures. The ability to discriminate haplotypes in mixed samples by using LNA-mediated PCR clamping indicates the potential value of mtDNA analysis in criminal investigations.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Quantitative Analyse von miRNA in Myokardinfarkten und plötzlichen Herztodesfällen in Blut und Gewebe

Mildeberger, Luise[1]; Wilmes, Verena[1]; Scheiper-Welling, Stefanie[1]; Verhoff, Marcel A.[1]; Kauferstein, Silke[1]

[1]Institut für Rechtsmedizin, Universitätsklinikum Frankfurt, Goethe-Universität, Frankfurt am Main, Deutschland

Rechtsmedizin Frankfurt am Main

Einleitung

Die World Health Organisation (WHO) verzeichnete im Jahr 2016 17,9 Millionen Tote aufgrund kardiovaskulärer Erkrankungen. Damit stehen diese an erster Stelle der Todesursachen weltweit. Um die verschiedenen Arten dieser Krankheiten voneinander zu unterscheiden, bedarf es spezifischer und sensitiver Biomarker. Einen der vielversprechendsten Biomarker stellen zurzeit mikroRNAs (miRNAs) dar. Bei ihnen handelt es sich um kurze, nicht kodierende RNAs, die regulierend in viele biologische Prozesse eingreifen. Sie binden im Bereich des 3'UTR ihrer Ziel mRNA und können so entweder deren Translation inhibieren oder zu deren Abbau führen.

Im Fokus unserer Studie stand der Vergleich der Expressionsmuster verschiedener miRNAs in Myokardinfarkten (MI) und plötzlichen Herztodesfällen (*sudden cardiac death* = SCD), sowie zwischen Gewebe und zirkulierenden miRNAs. Ziel der Studie sollte sein, einen geeigneten Biomarker zur Unterscheidung zwischen MI und SCD zu finden.

Methoden

Alle für diese Studie verwendeten Proben wurden post-mortem entnommen und vom Institut für Rechtsmedizin in Frankfurt zur Verfügung gestellt. Es wurden Herzgewebeproben von an Myokardinfarkt verstorbenen Personen und von Personen, die an einem Plötzlichem Herztod verstorben sind, verwendet. Darüber hinaus wurde, soweit möglich, jeweils von den selben Individuen Vollblut verwendet. Im Falle der MI Herzen wurde zwischen I Gewebe (infarziertes Gewebe, n = 15) und NI Gewebe (nicht-infarziertes Gewebe, n = 14) unterschieden. Aus SCD Herzen wurden jeweils zwei biologische Replikate entnommen (n = 14). Die Vollblutproben von MI und SCD hatten jeweils eine statistische Größe von n = 15. Als Kontrollen (C) wurde Herzgewebe aus Vorder- und Hinterwand von Suiziden oder Unfalltoden verwendet (n = 4).

Nach der miRNA Isolation und Quantifizierung wurden die Expressionen von verschiedenen miRNAs in Herzgewebe und in Vollblut von MI und SCD relativ mittels q-RT-PCR unter Berücksichtigung der MIQE Richtlinien bestimmt.

Ergebnisse

In SCD Gewebeproben konnte eine signifikante Hochregulation an miR-1 im Vergleich zu den Kontrollen detektiert werden (p = 0.036). Die Expression von miR-1 war ebenfalls stark hochreguliert im SCD Gewebe im Vergleich zu infarziertem MI Gewebe (p = 0.054). Ein Unterschied zwischen C, I und NI war im Gewebe jeweils nicht zu beobachten (Abb. 1, A & B).

Im Blut zeigte sich eine anderes Bild. Dort konnte, im Vergleich zum Gewebe, eine Hochregulation der miRNAs in MI Blut festgestellt werden. Im Falle von miR-133a war der Unterschied zwischen der Expression in MI Blut und SCD Blut signifikant (p = 0.015). Die Expression von miR-133a in MI war ebenfalls erhöht im Vergleich zu C (p = 0.059) (Abb. 1, C &D).

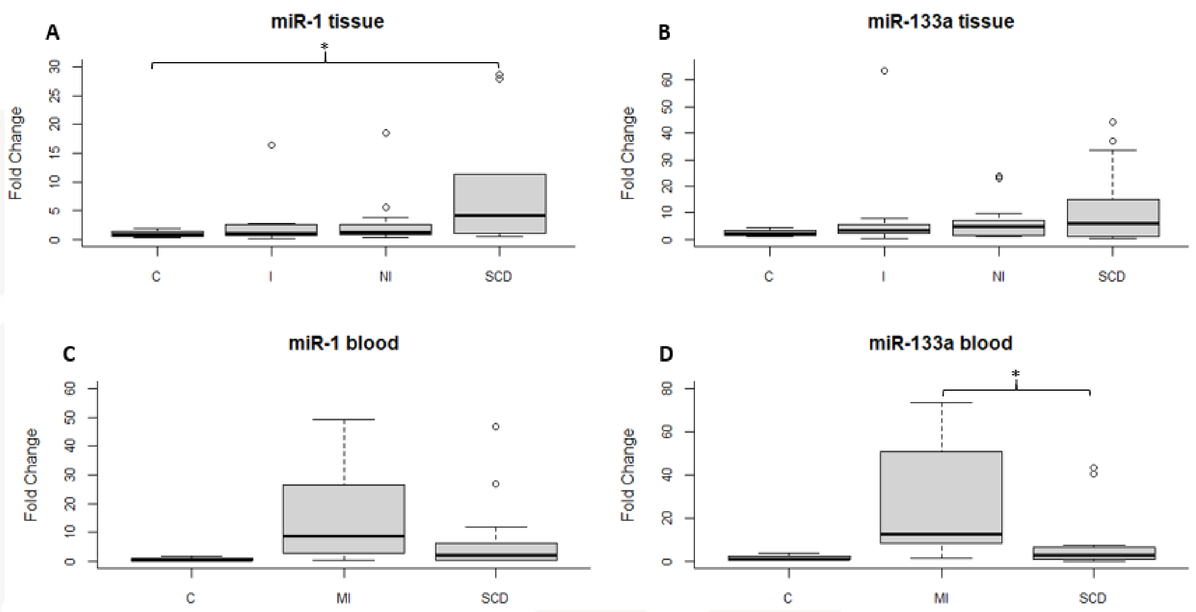
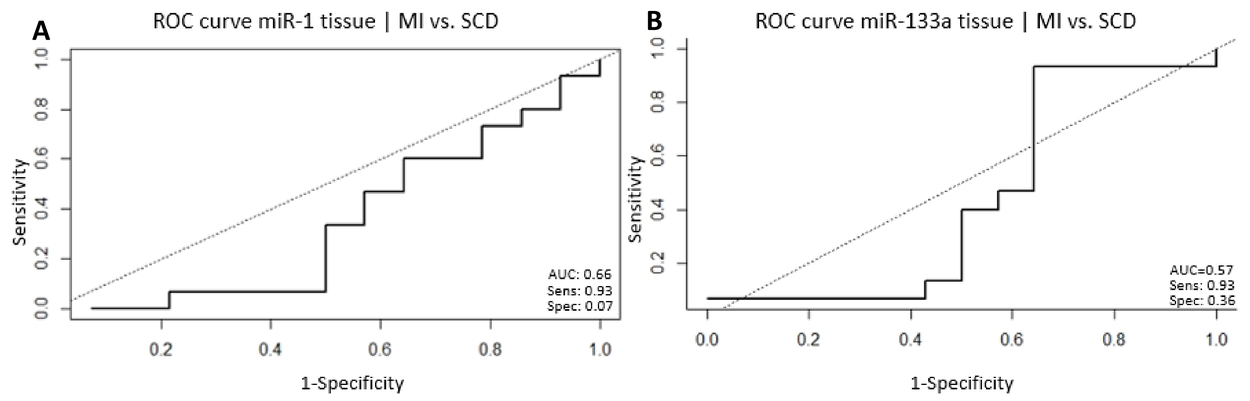


Abb. 1: *Boxplots der Fold Change Werte von **A** miR-1 im Gewebe, **B** miR-133a im Gewebe, **C** miR-1 in Vollblut und **D** miR-133a in Vollblut*



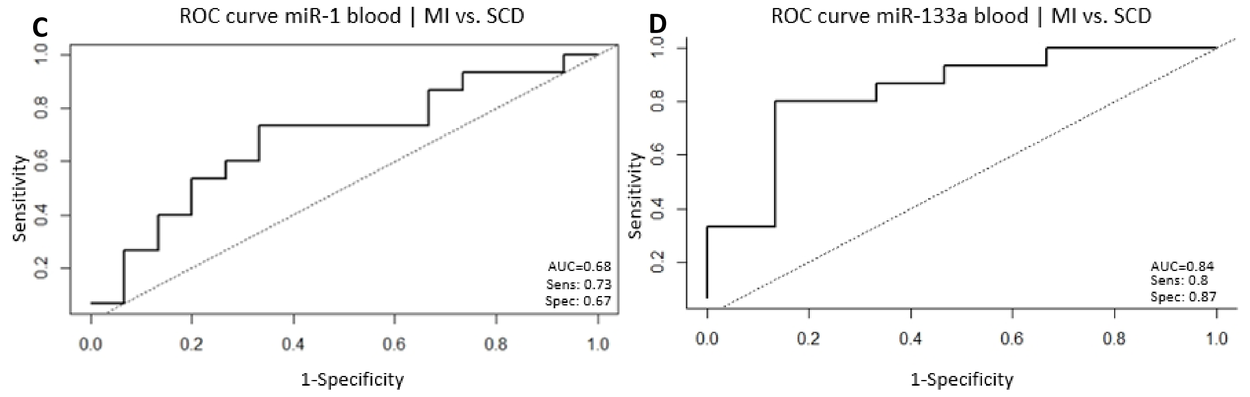


Abb. 2: *ROC Kurven vom Vergleich **A** miR-1 bei MI und SCD im Gewebe, **B** miR-133a bei MI und SCD im Gewebe, **C** miR-1 bei MI und SCD in Vollblut und **D** miR-133a bei MI und SCD in Vollblut*

Als spezifischster und sensitivster Marker zur Abgrenzung von MI zu SCD stellte sich miR-133a in Vollblut heraus (Abb.2). Mit nur 2 falsch positiven Zuordnungen konnte mit Hilfe dieses Markers die meisten MI Fälle als solche erkannt werden.

Allgemein konnte durch die Studie festgestellt werden, dass sich Marker im Vollblut besser zur Unterscheidung von MI und SCD eignen als Marker in Gewebe.

Zusammenfassung

In dieser Studie konnte nachgewiesen werden, dass sich die miR-1 und die miR-133a Expression in MI und SCD sowohl im Gewebe als auch im Vollblut unterscheiden.

Zusätzlich zum Bericht von Pinchi et al. (2019) [1], die u.a. miR-1 und miR-133a als Marker in Herzgewebe zur Unterscheidung zwischen MI und SCD nachgewiesen haben, konnte diese Studie belegen, dass sich diese Marker auch im Vollblut zum Nachweis und zur Differenzierung eignen. Als spezifischster und sensitivster Marker erwies sich dabei miR-133a.

Es ist bereits bekannt, dass sich die miRNA Expression nicht nur im zeitlichen Verlauf des MI Geschehens, sondern auch je nach Art des MI ändern kann [2][3]. Daher wäre eine genauere Einteilung der MI Fälle, z.B. anhand MI Marker wie cTnT im Serum oder Plasma, sinnvoll.

Quellen

[1] Pinchi, Enrica; Frati, Paola; Aromatario, Mariarosaria; Cipolloni, Luigi; Fabbri, Matteo; La Russa, Raffaele et al. (2019): miR-1, miR-499 and miR-208 are sensitive markers to diagnose sudden death due to early acute myocardial infarction. In: *Journal of cellular and molecular medicine*

[2] Kuwabara, Yasuhide; Ono, Koh; Horie, Takahiro; Nishi, Hitoo; Nagao, Kazuya; Kinoshita, Minako et al. (2011): Increased microRNA-1 and microRNA-133a levels in serum of patients with cardiovascular disease indicate myocardial damage. In: *Circulation. Cardiovascular genetics*

[3] Feng Wang; Guangwen Long; Chunxia Zhao; Huaping Li; Sandip Chaugai; Yan Wang et al. (2013): Plasma microRNA-133a is a new marker for both acute myocardial infarction and underlying coronary artery stenosis. In: *Journal of Translational Medicine.*

Kindly supported by FORENSISCHES FORUM

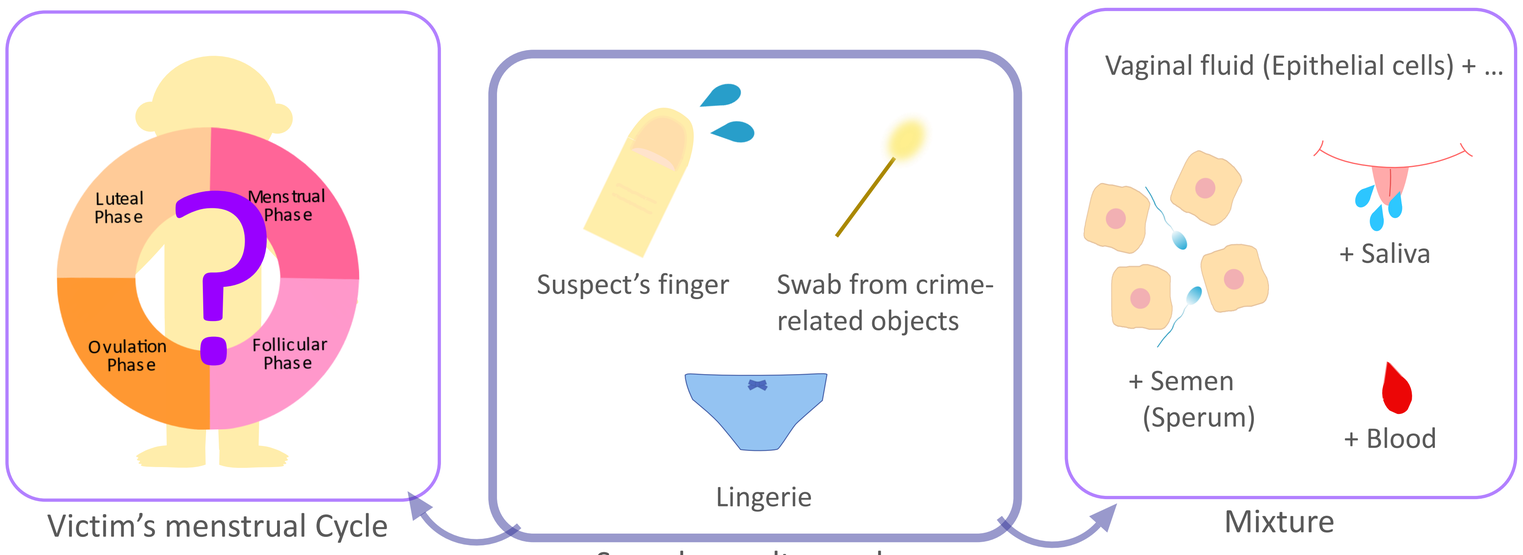
Robustness of *Lactobacillus crispatus* DNA-based vaginal fluid identification

Masako Ohmura[1], Jun Ohta[1,2]

[1] Forensic Biology Unit, Scientific Crime Laboratory, Kanagawa Prefectural Police, Kanagawa, Japan

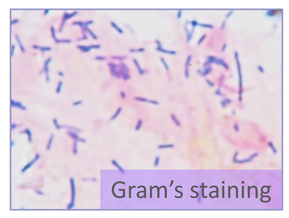
[2] Department of Forensic Dentistry, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

Objective

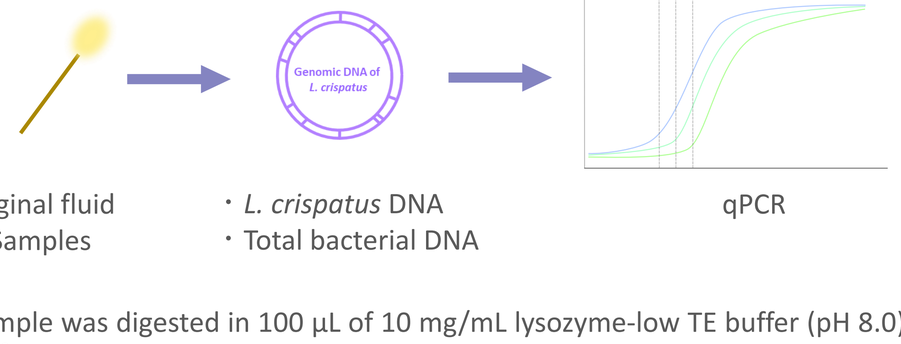


Lactobacillus crispatus DNA-based vaginal fluid identification (VFID) is a useful approach for investigating sexual assaults. However, it has two main limitations: 1) The constituent bacteria of vaginal flora are affected due to the menstrual cycle, and **vaginal fluid samples of various menstrual phases are collected for studying a case** and 2) **Semen-mixed vaginal fluid samples are handled using differential extraction protocols for human identification (HID) in a case**, thereby increasing sample consumption. This study aimed to evaluate the robustness of *L. crispatus* DNA-based VFID using various menstrual phase samples and semen-mixed vaginal fluid samples.

Methods

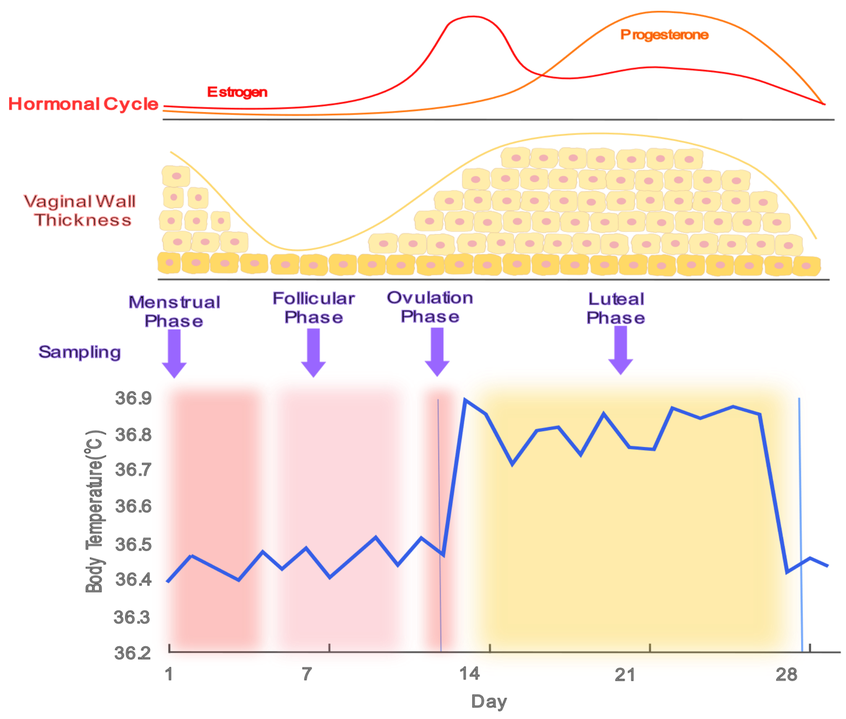


L. crispatus
- Gram-positive bacillus
- One of the predominant species in the vagina
- PCR target: 16S rRNA gene (V1; 137 bp)



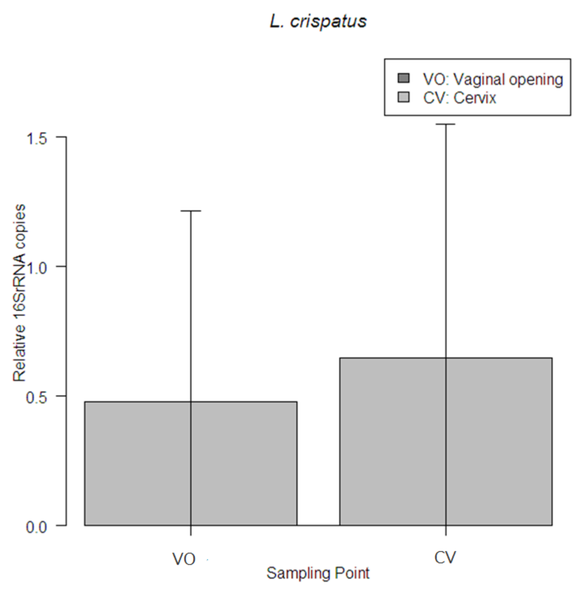
A sample was digested in 100 µL of 10 mg/mL lysozyme-low TE buffer (pH 8.0) at 37° C for 1 h. Total DNA of the lysate (500 µL) was extracted using the PrepFiler™ Forensic DNA Extraction Kit (Thermo Fisher Scientific) and an AutoMate *Express*™ Forensic DNA Extraction System (Thermo Fisher Scientific) and eluted in 50 µL of low TE buffer. *L. crispatus* DNA and total bacterial DNA were quantified by quantitative PCR (qPCR).

Effects of the menstrual cycle on *L. crispatus* DNA-based VFID

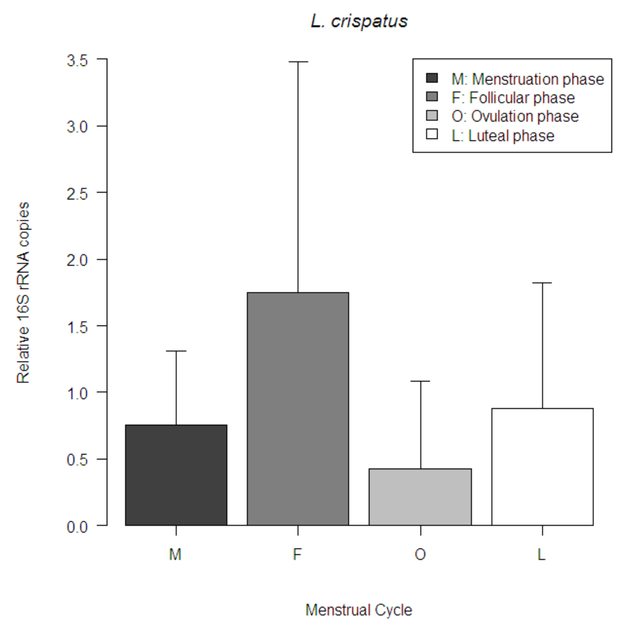


To evaluate the influence of menstrual cycle for *L. crispatus* DNA-based VFID, vaginal fluid samples were collected at menstruation, follicular, ovulation, and luteal phases. **The growth environment of microorganisms (e.g., temperature, intimal condition) at each phase greatly changes due to the influence of female hormones.**

(A) Sampling sites

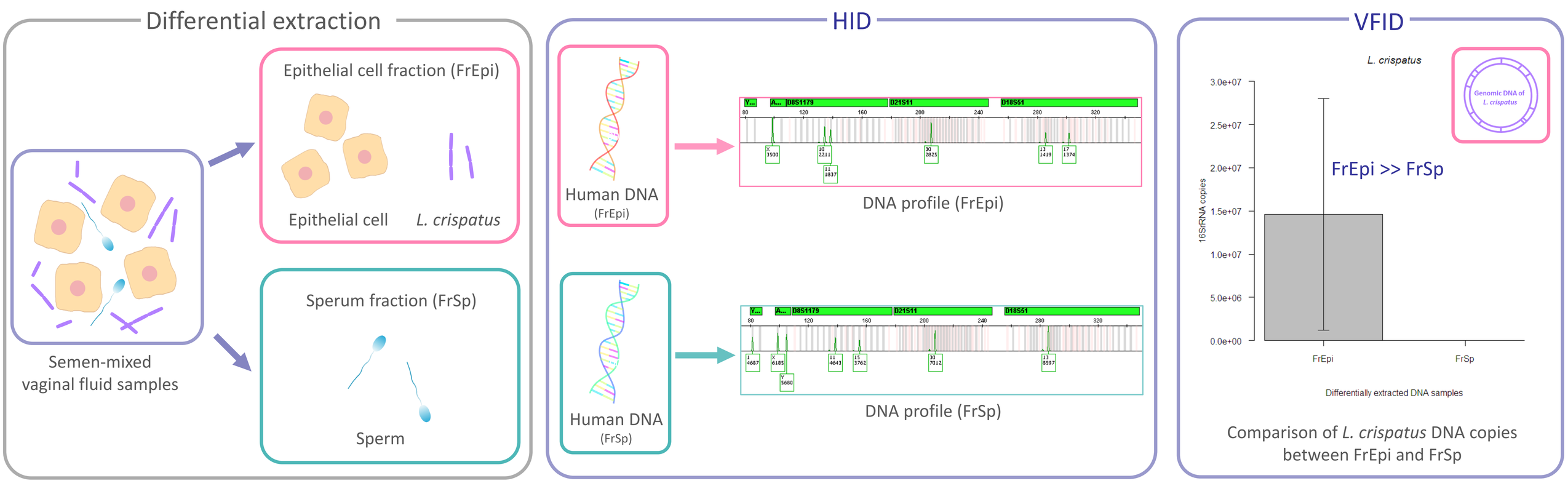


(B) Menstrual cycle



Vaginal fluid samples were collected from a healthy female donor from two sites and four sampling periods of the menstrual cycle. To normalize the analyzed samples, relative 16S rRNA copies were calculated by dividing the number of *L. crispatus* DNA copies to the total number of bacterial DNA copies. Relative 16S rRNA copies of *L. crispatus* DNA were not significantly different between the vaginal fluid samples collected from two different sites (A; Student' *t*-test, $P = 0.756$) as well as among samples collected in all phases of the menstrual cycle (B; Tukey HSD test, $P > 0.05$).

A combination approach of human DNA-based HID and *L. crispatus* DNA-based VFID



Semen-mixed vaginal fluid samples were treated using the differential extraction process (n = 3). Both the DNA extracts of epithelial cell and sperum fractions were amplified using the GlobalFiler™ Kit and examined using *L. crispatus* DNA-based VFID. *L. crispatus* DNA was quantified from epithelial cell fraction, and male donor-derived DNA profiles were obtained from sperm fractions.

Conclusions

- ***L. crispatus* DNA is a robust VFID marker.**
  L. crispatus DNA could be detected stably regardless of the sampling sites or the menstrual phases.
- ***L. crispatus* DNA-based VFID approach can be combined with HID approach for the analysis of semen-mixed vaginal fluid samples.**
  The combined approach is practical, as it can minimize sample utilization. We recommend that the use of FrEpi for *L. crispatus* DNA-based VFID.

Characterization of nonfunctional alleles of FUT2 in a database

Mikiko Soejima and Yoshiro Koda
Forensic Medicine, Kurume University School of Medicine, Japan

- Because the blood group antigens on the surface of mucosa or in secretions can be used as a scaffold for infection by some kinds of pathogens, secretor status may associate with susceptibility to infectious diseases. The expression of ABO antigens in secretion is regulated by the FUT2, which encodes a secretor type α(1,2)fucosyltransferase. Secretors express ABO antigens in saliva and non-secretors who are homozygotes of the null alleles (se) do not. The prevalence of non-secretors is about 20–25% in most geographic regions.
- Many single nucleotide polymorphisms (SNPs) and recombinant null alleles have been identified and some of them showed a population-specific pattern. For association studies on secretor status, we need to know FUT2 variation of the target populations unless we determine the secretor status by phenotype.
- Erythrogene v0.8 extracted the data of blood group alleles from 1000 Genome Project. To survey the uncharacterized nonfunctional alleles of FUT2, we selected 18 alleles listed in Erythrogene and obtained genomic DNA of 22 individuals who have these alleles. In this study, we experimentally confirmed the nonsynonymous SNPs and their haplotypes (alleles). We also analyzed the encoded enzyme activities by transient expression assays. In addition, we predicted the effect of each nonsynonymous SNP on the protein by several in silico methods.

MATERIALS AND METHODS

The study protocol was approved by the ethical committees of Kurume University School of Medicine (No. 342).

22 genomic DNAs were purchased from the Coriell Institute for Medical Research. The coding regions of FUT2 were amplified and sequenced. To determine the haplotypes, we cloned PCR products into a plasmid and sequenced the clones. To evaluate the significance of uncharacterized alleles, a transient expression experiment was performed. The FUT2 containing each SNP concerned, Se^{357} (wild-type allele), se^{428} (nonsecretor) or $se^{357,385}$ (weak-secretor) were inserted into pcDNA3.1(+) vectors. 3 μg of each construct together with 100 ng of the pGL3 Promoter was transfected COS-7 cells by means of TransIT-X2. After 2 days of culture, the cells were incubated with 1E3 antibody for H type 1-4, followed by incubation with FITC-conjugated goat anti-mouse IgM, and expression of H antigen on the cell surface was monitored by flow cytometry. The experiments were performed five times. The firefly luciferase activity was assayed using ONE-Glo Luciferase Assay System and the transfection efficiency was confirmed by the intensity of luciferase light. The effect of each nonsynonymous on the function of the enzyme were predicted using four free software, MutationTaster, Mutation Assessor, PolyPhen-2, and Sorting Intolerant From Tolerant (SIFT). Because functional analysis showed allele containing 357C>T; 856T>C; 863C>T to be nonfunctional, site-specific mutagenesis of the plasmid was performed to determine which substitution is responsible for inactivation of the enzyme. Finally, we obtained two constructs, pcDNA3.1(+) containing FUT2 of 357C>T; 856T>C and 357C>T; 863C>T.

RESULTS

1. *Sequence and haplotype determination of FUT2*

We detected all of the registered SNPs in the database and encountered 17 uncharacterized nonsynonymous SNPs: 58C>T, 80T>C, 205G>A, 262A>C, 370G>A, 539G>A, 542C>T, 544G>A, 566T>C, 610G>T, 616G>C, 715C>T, 856T>C, 863C>T, 964A>G, and 980C>A in 22 individuals. In addition to these, the alleles of 357C>T, 403C>T, and 950C>T were also investigated, although 403C>T and 950C>T themselves have been identified in Latin Americans ($Se^{357,403}$) and a Mongolian ($se^{357,950}$), respectively. However, the haplotypes of ten of 18 alleles were different from those in the database. Among the ten alleles, three SNPs (370G>A, 544G>A, and 863C>T) were found on the nonfunctional allele se^{428}, and 403C>Tand 950C>T were found on another chromosome. Accordingly, we performed functional analyses of the 13 alleles. 863C>T was on two different alleles, 357C>T; 856T>C; 863C>T and 428G>A (plus 4 additional SNPs); 863C>T.

2. *Functional analyses of candidates of non-secretor alleles*

To determine whether each allele encodes a functional Se enzyme or not, the cell surface H antigen was examined by flow cytometry. The expression levels of H antigens are shown in Table 1. The percentage of H antigen–positive cells transfected with pcDNA3.1 ligated with the FUT2 of 80T>C; 357C>T, 964A>G; 357C>T, 357C>T; 539G>A, 40A>G; 616G>C, 357C>T; 715C>T, 58C>T; 357C>T, 357C>T; 980C>A, 205G>A; 357C>T, and 357C>T; 542C>T were almost identical to those ligated with the positive control of Se^{357} (wild-type allele, 26.22 ± 5.06%).

On the other hand, the that of 566T>C (9.28 ± 0.99%) was less than half of Se^{357} but higher than that of $se^{357,385}$ (4.86 ± 0.76%), while that of 262A>C; 357C>T (2.03 ± 0.43%) was lower than that of $se^{357,385}$. Those of 610G>T or 357C>T; 856T>C; 863C>T was almost undetectable as was that of se^{428} (0.68 ± 0.19%). We categorized FUT2 of 566T>C as a secretor allele (Se^{566}) and that of 262A>C; 357C>T as a weak secretor allele ($se^{262,357}$, Se^{w}). Although the 566T>C substitution seemed to partially impair enzyme activity

3. *Determination of causal SNP of FUT2-357C>T; 856T>C; 863C>T for inactivation of the encoded enzyme*

FUT2 of 357C>T; 856T>C; 863C>T was shown to be a nonfunctional allele. However, it was unclear which SNP of 856T>C and 863C>T was involved in inactivation of the Se enzyme. We generated FUT2 of 357C>T; 856T>C and FUT2 of 357C>T; 863C>T by *in vitro* mutagenesis. 856T>C completely impaired and 863C>T somewhat impaired the activity.

4. *Estimation of significance of uncharacterized nonsynonymous SNPs*

We predicted the possible impacts of 14 amino acid substitutions on the functions of Se enzyme using four software programs.

The results were not always consistent with those of expression experiments (Table 1). We observed discrepancies for six to nine out of 14 SNPs, while all of the SNPs reside in the catalytic domain. The software generally tended to overestimate the impacts of the nonsynonymous SNPs we tested here.

Table 1 Evaluation of candidates for nonfunctional FUT2 alleles by expression of cell surface H antigens or in silico analysis of the nonsynonymous SNP

Allele	Expression level (± SD)	A.A change	Mutation Tester	Mutation Assessor (score)	Polyphen2 (score)	SIFT (score)	Allele
80T>C; 357C>T	25.79 (5.43)	M27T	P	N (-0.835)	B (0.000)	T (0.16)	$Se^{80,357}$
357C>T; 964A>G	19.57 (2.34)	I322V	D	M (2.990)	P (0.961)	A (0.00)	$Se^{357,964}$
357C>T; 539G>A	23.99 (3.69)	R180Q	P	N (0.620)	B (0.026)	T (0.29)	$Se^{357,539}$
610G>T	0.63 (0.36)	G204W	D	M (3.005)	P (1.000)	A (0.00)	se^{610}
40A>G; 616G>C	21.32 (2.75)	V206L	P	L (1.680)	B (0.172)	A (0.00)	$Se^{40,616}$
357C>T; 715C>T	20.45 (7.38)	R239W	D	M (2.915)	P (1.000)	A (0.05)	$Se^{357,715}$
357C>T; 856T>C; 863C>T	0.68 (0.27)						$se^{357,856,863}$
Alleles generated by mutagenesis							
357C>T; 856T>C	0.63 (0.28)	Y286H	D	M (3.005)	P (1.000)	A (0.00)	
357C>T; 863C>T	7.98 (0.55)	T288M	P	L (1.445)	P (0.993)	A (0.00)	
58C>T; 357C>T	24.31 (3.82)	R20W	P	M (2.915)	P (1.000)	A (0.02)	$Se^{58,357}$
566T>C	9.28 (0.99)	V189A	D	M (3.005)	P (1.000)	A (0.00)	Se^{566}
357C>T; 980C>A	20.32 (2.06)	S327Y	P	M (1.995)	P (0.989)	A (0.00)	$Se^{357,980}$
205G>A; 357C>T	19.27 (9.62)	A69T	D	M (3.005)	P (1.000)	A (0.00)	$Se^{205,357}$
357C>T; 542C>T	21.69 (1.50)	P181L	D	M (2.810)	P (0.973)	A (0.00)	$Se^{357,542}$
262A>C; 357C>T	2.03 (0.43)	M88L	P	M (3.005)	P (0.988)	A (0.00)	$se^{262,357}$ (Se^{w})

Note: P and D represent polymorphism and disease-causing, respectively (MutationTester). N, M, and L represent neutral, medium and low, respectively (MutationAssessor). B and P represent benign and probably damaging, respectively (Polyphen2). T and A represent tolerated and affect protein function, respectively (SIFT).

CONCLUSION

We identified two non-secretor alleles (se^{610} and $se^{357,856,863}$) and one weak secretor allele ($se^{262,357}$) in samples in the 1000 Genome Project. Experimental phasing and expression studies are desirable for analysis of FUT2.

CONFLICT OF INTEREST

The authors have declared no conflict of interest.

REFERENCE

1. Soejima M and Koda Y. Survey and characterization of nonfunctional alleles of FUT2 in a database. Sci Rep. Feb 4;11(1):3186, 2021 doi: 10.1038/s41598-021-82895-w.

KURUME UNIVERSITY

DNA methylation-based age prediction in the Japanese population by pyrosequencing analysis

Yuta Takahashi[1], Masaru Asari[1], Katsuhiro Okuda[1], Kanae Mori[1], Ryo Namba[2], Keiko Shimizu[1]
1 Department of Legal Medicine, Asahikawa Medical University, Asahikawa, Japan
2 Sapporo Higashi Tokushukai Hospital, Sapporo, Japan

[Introduction]

Age prediction based on DNA methylation levels has been reported as an informative method for forensic science because of the association of methylation levels with chronological age. However, large amounts of DNA (10–20 ng) are recommended for methylation analysis because of its low sensitivity. In this study, we investigated the correlation between age and methylation levels in a Japanese population. The effects of multiplex PCR on the sensitivity and accuracy of the predictions were also analyzed.

[Materials and methods]

1. Correlation between methylation levels and chronological age in the Japanese population

1.1. Samples collection

Buccal cells were collected from 75 healthy Japanese individuals and genomic DNA was extracted. DNA quantification was performed using a Quantifiler Human DNA Quantification Kit (Life Technologies). The extracted genomic DNA was subjected to bisulfite conversion using an EpiTect Fast Bisulfite Conversion Kit (Qiagen).

This study was approved by the ethics committee of Asahikawa Medical University.

1.2. PCR amplification and analysis of DNA methylation using pyrosequencing technology

A total of 10 ng of the bisulfite-converted DNA was used for PCR amplification of 35 cycles for eight selected genomic regions that have been described previously for East Asian populations (Table1). DNA methylation levels at CpG sites in these regions were measured by pyrosequencing analysis.

1.3. Age prediction calculator in the Japanese population

The age predicted by multiple liner regression analysis was compared with the chronological age.

2. Effects of the amount of DNA and multiplex PCR on sensitivity and accuracy

2.1. Effects of amount of DNA

2.1.1. DNA sample

We used “EpiTect PCR control DNA, methylated/unmethylated and bisulfited converted (Qiagen)”. Fully methylated control DNA and fully unmethylated DNA were mixed in appropriate ratios to make 50% methylated control DNA.

2.1.2. PCR amplification and analysis of DNA methylation using pyrosequencing technology

We used 10 or 50 ng of DNA for PCR amplification for regions 5 and 6. Subsequent amplification and methylation detection were performed as described above.

2.2. Effects of multiplex PCR

2.2.1. Multiplex PCR

A total of 10 ng of DNA was used for multiplex PCR amplification of 10 cycles for the eight selected genomic regions.

2.2.2. PCR amplification and analysis of DNA methylation using pyrosequencing technology

Parts of the reaction solution were used for subsequent amplification of 35 cycles for each of the eight selected genomic regions. Methylation detection was performed as described above.

Table 1 Eight genomic regions selected for DNA methylation analysis.

Region	Gene	Marker		Size (bp)
1	TRIM59	cg07553761	10,11,12	142
2	TRIM59	cg07553761	18	194
3	RASSF5	cg08128734	4	138
4	C1orf132	cg10501210	11	256
5	No gene	cg10804656	6,7	200
6	PDE4C	cg17861230	26,27,28,29	181
7	CCDC102B	cg19283806	4	112
8	ELOVL2	cg21572722	11,12,13,14	296

These sites were described previously for East Asian populations (Feng et al., 2018)

[Results]

1. Correlation between age and methylation levels in the genomic DNA samples

The methylation levels of six of the eight selected genomic regions for the 75 samples are shown in Fig. 1. Regions 1, 2, 5, 6, and 8 exhibited positive correlation, region 4 showed negative correlation, and regions 3 and 7 exhibited lower correlation than the others between age and methylation levels.

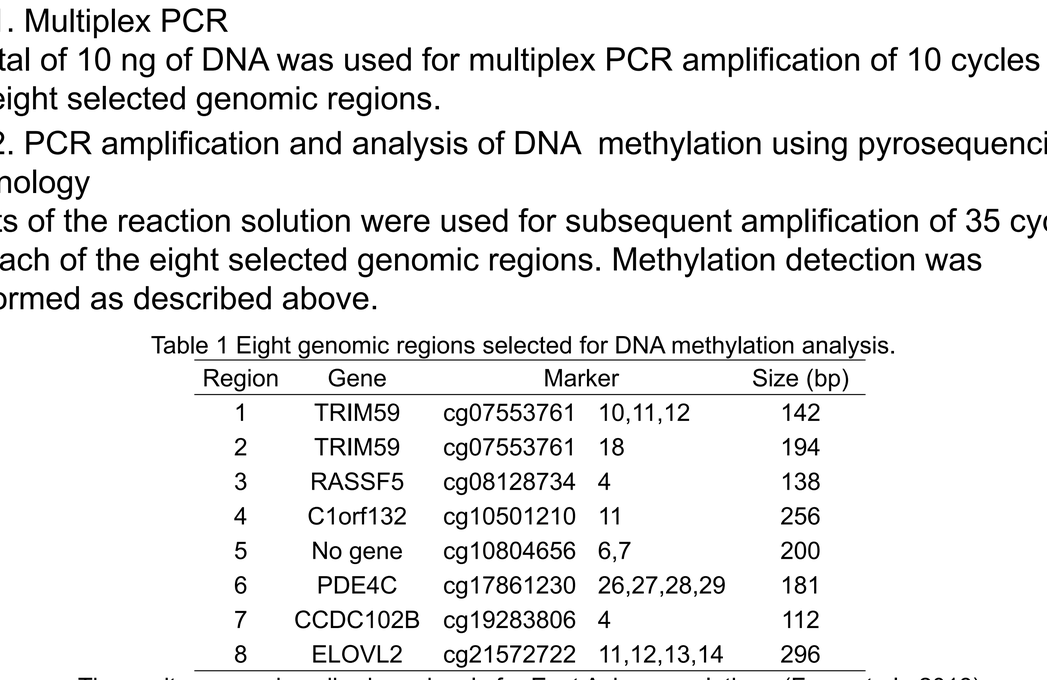


Fig. 1 Correlation between age and methylation levels in 75 health Japanese individuals
Six out of eight genomic regions exhibited correlation between age and methylation.
R^2, Pearson correlation coefficient

A model composed of seven of the eight selected regions exhibited strong correlation between predicted and chronological age, with mean absolute deviation from chronological age of 3.54 years (Fig. 2).

Predicted age =−0.154 × (No.1CpG10)
+0.447 × (No.1CpG11) +0.0508 × (No.1CpG12)
−0.401 × (No.3) −0.119 × (No.4) +0.453 × (No.5CpG6)
+0.342 × (No.5CpG7) −0.0467 × (No.6CpG26)
−0.0269 × (No.6CpG27) +0.231 × (No.6CpG28)
+0.276 × (No.6CpG29) +0.802 × (No.7)
−0.0518 × (No.8CpG11) +0.0110 × (No.8CpG12)
−0.0370 × (No.8CpG13) +0.347 × (No.8CpG14),
where No.1, No.3, No.4, No.5, No.6, No.7, and No.8 are regions 1, 3, 4, 5, 6, 7, and 8 respectively.

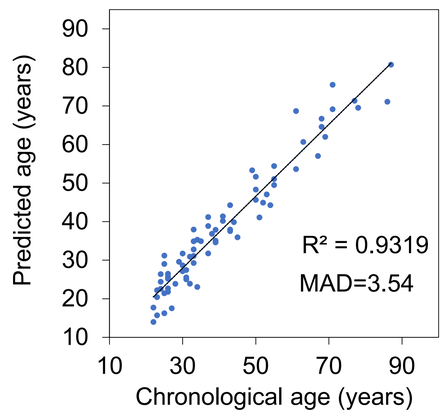


Fig. 2 Correlation between age predicted by multiple liner regression analysis and chronological age.
R^2, Pearson correlation coefficient;
MAD, mean absolute deviation

2. Effects of amount of DNA and multiplex PCR on sensitivity and accuracy

2.1. Effects of amount of DNA

The results for regions 5 are shown in Fig. 3. When the amount of DNA was large, the value was close to the expected methylation levels (50%). When the amount of DNA was small, the variation, as measured 3 times, was high. The results were similar for region 6.

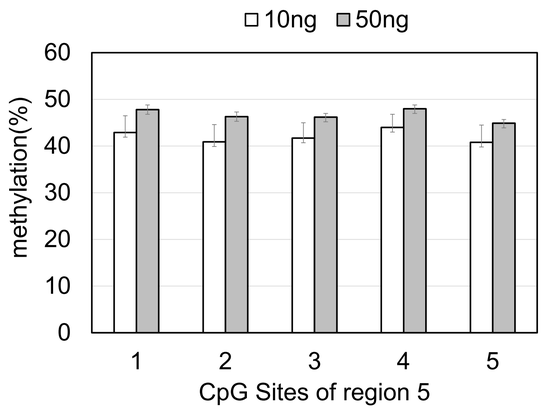


Fig. 3 Effects of amount of DNA on methylation levels
The graph shows the mean ± S.D. (n = 3), which is lower than the expected value of 50%.

2.2. Effects of multiplex PCR

The effects of multiplex PCR on methylation(%) and sensitivity are shown in Table2. The methylation levels increased or decreased slightly in many areas; only region six changed significantly. The result shows that the effect on the accuracy of the methylation levels was small, whereas the sensitivity was improved in many areas.

Table 2 Effects of multiplex PCR on methylation and sensitivity in the eight selected regions

Region	Methylation(%) and change in multiplex PCR			Sensitivity
	Multiplex PCR		Change	
	–	+		
1	33.78	30.40	↓	Increased 2 times
2	failed	failed	–	–
3	31.90	38.22	↑	Increased 2 times
4	38.60	48.65	↑↑	Reduced to half
5	36.94	34.95	↓	Increased 2 times
6	48.90	71.18	↑↑↑	Increased 4 times
7	41.68	43.99	↑	Increased 2 times
8	31.36	34.71	↑	Increased 2 times

[Discussion]

1. Correlation between age and methylation levels in the Japanese population

Table3 shows the comparison between this study and age prediction methods by methylation analysis reported in recent years. Our results indicate that the seven regions can be used to predict chronological age in the Japanese population.

Table 3 Comparison between the results of this study and other age prediction studies reported recently

	Hamano et al. (2016)	Hong et al. (2017)	Feng et al. (2018)	Dias et al. (2020)	Fleckhaus et al. (2020)	**This study**
Sample	Japanese	Korean	Chinese	Portuguese	German	**Japanese**
Method	HRM	SBE	PS	SS	PS	**PS**
Number of positions analyzed	2	6	8	5	8	**7**
MAD (years)	7.44	3.13	2.89	6.08	3.07	**3.54**

HRM, high resolution melting; SBE, single base extension; PS, pyrosequencing; SS, sanger sequencing; MAD, mean absolute deviation.

2. Effects of amount of DNA and multiplex PCR on sensitivity and accuracy

Multiplex PCR affected the sensitivity and accuracy, and it improved the sensitivity in many regions. However, the accuracy differed among the eight regions, probably because the sizes of their amplification products ranged from 112 to 296 bp. Therefore, to correctly analyze methylation levels by multiplex PCR, it may be necessary to optimize the primers and sequences to be analyzed.

[Conclusions]

Our results indicate that the seven selected regions can be used to predict chronological age in the Japanese population. Multiplex PCR affected the sensitivity and accuracy, and it improved the sensitivity in many regions.

[Conflict of interest statement]

The authors declare that they have no conflicts of interest.