

厚生労働行政推進調査事業費補助金（健康安全・危機管理対策総合推進事業）
分担研究報告書

原因不明疾患におけるゲノム解析法の構築

研究分担者 長島 真美 東京都健康安全研究センター
研究協力者 浅倉 弘幸 東京都健康安全研究センター

研究要旨

原因不明疾患等における次世代シーケンサー（NGS）の利用は、近年、重要視されており、年々増加傾向にある。一方で、解析法についてはこれといった標準法はなく、地方衛生研究所（地衛研）独自の解析方法の確立が急務である。今回、不明疾患解析検体およびアンプリコン NGS 解析を行う解析ソフト PaHuM ならびに MTAAP の評価を行った。結果として、どちらの解析ソフトも機能的に市販ソフトと同等以上であることが示され、十分に地衛研の NGS 解析に役立つことが示された。また、これらソフトの利用法を含めて、地衛研の職員を対象とした Web 研修会を実施した。

A. 研究目的

原因不明疾患における、次世代シーケンサー（NGS）の利用および解析は、従来より重要視されてきた。特に、茨城県衛生研究所が関与し、報告がなされたオズウイルス感染症の究明では、国立感染症研究所（感染研）の解析システムが使用された経緯がある。一方で、感染研の解析システムは、現在、地方衛生研究所（地衛研）の使用が不可能であり、地衛研独自の不明疾患検査体制の確立（解析方法の確立）が急務であった。

また、多数のプライマーを使用し、PCR を行った後に NGS 解析を行うアンプリコンベースの NGS 解析では、NGS 解析の感度・特異性向上で重要な要素であるが、最終的に使用したプライマーの除去が必須である。この手法は、新型コロナウイルス（SARS-CoV-2）解析で、全国で実施されてきた。この解析では、感染研の Web サイト：PathoGenS が利用されているが、SARS-CoV-2 以外の病原体解析には利用できない。

そこで、本研究では、原因不明疾患における網羅的解析やアンプリコン NGS 解析に適した解析ソフトの構築・導入と、地衛研への研修を目的とする。

B. 研究方法

1. 不明疾患解析検体の解析

東京都健康安全研究センターウイルス研究科に搬入され、ノロウイルスが検出された糞便検体 2 検体から抽出した RNA を供試し、NGS データをウイルス解析用パイプライン（PaHuM）（関塚産業有限会社）で解析を行い、検出したウイルスゲノムのリード数で解析ソフトの有用性を検討した。さらに、インフルエンザウイルスウイルスが検出された臨床検体より rRNA の除去の有無による解析データを検討した。

2. アンプリコン NGS 解析

東京都健康安全研究センターウイルス研究科に搬入された SARS-CoV-2 検体で、PCR 増幅後の産物を解析ソフト MTAAP（関塚産業有限会社）と従来法でのプライマー部分の除去、コンセンサス配列の取得を比較検討した。

3. Web 研修会

本研究班で検討したウイルス解析用パイプラインの使用例を地衛研セミナーとして、Web 研修を実施する。

（倫理面への配慮）

本研究班で実施する事項については、東

京都健康安全研究センター倫理委員会で承認されている（6 健研健第 906 号）。

C. 研究結果

1 不明疾患解析ソフトの使用

ウイルス解析用パイプライン（PaHuM）は（図 1）、臨床検体を用いた NGS 解析データ（メタゲノム）中に含まれる宿主（ヒト）ゲノムの除去、およびウイルスデータベースを用いたウイルスリードの検索を行うソフトである。ノロウイルスが Ct 値 30 前後で検出された糞便検体 2 検体の解析で、それぞれ Contig の長さは異なるが、2 検体からノロウイルス GI が検出された（図 2）。

また、rRNA の除去効果を見る目的で、検討したところ、rRNA 除去した方が、各セグメントのリード長、カバレッジともに高い結果が得られた（図 3）。

2 アンプリコン NGS 解析

解析ソフト MTAAP は、アンプリコン内のプライマー配列を除去するのみではなく、コンセンサス配列の取得や、変異解析まで可能である（図 4）。

今回、SARS-CoV-2 の解析を本ソフトで実施したところ、従来、Ubuntu（プライマー除去）+CLC Genomics Workbench（変異解析）で 40 分であったものが、MTAAP では 5 分程度であった（図 5）。また、解析能力については、CLC Genomics Workbench の解析と概ね一致した。一部では、CLC で変異のないとさらた部分が、MTAAP では Mix 塩基と判定されるなど、有用性が示された（図 6）。

3. Web 研修会

テーマ：「不明疾患における NGS 解析法について考えてみる」で地衛研 Web セミナー第 4 回「Mini」を、テーマ：「不明疾患における NGS 解析ソフト PaHuM と PCR 増幅後の NGS 解析用ソフト MTAAP のご紹介」で、地衛研 Web セミナー第 5 回「Mini」を実施した。

●地衛研 Web セミナー第 4 回「Mini」

① 江原勇登（埼玉県衛生研究所）

「不明疾患における NGS 使用と解析について」

② 浅倉弘幸（東京都健康安全研究センター）

「糞便材料を使用した NGS 解析により判明したアストロウイルス」

③ 黒木絢士朗（東京都健康安全研究センター）

「海外のサイトを利用した NGS 解析（原因不明疾患）の比較と問題点」

④ 木村博一（群馬パース大学大学院保健科学研究科）

「メタゲノム解析・ウイルスゲノム解析用バイオインフォマティクスパイプラインプログラムの使用経験」

⑤ 関塚剛史（関塚産業有限会社）

「メタゲノム解析・ウイルスゲノム解析用バイオインフォマティクスパイプラインプログラム PaHuM のライセンス・仕様等について」

●地衛研 Web セミナー第 5 回「Mini」

① 関塚剛史（関塚産業有限会社）

「メタゲノム解析・ウイルスゲノム解析用パイプライン PaHuM の使用方法の説明について」

② 江原勇登（埼玉県衛生研究所）

「マルチプレックススタイレング PCR 増幅断片解析用 バイオインフォマティクスパイプラインプログラム MTAAP の使用経験（その 1）」

③ 浅倉弘幸（東京都健康安全研究センター）

「MTAAP の使用経験（その 2）」

④ 関塚剛史（関塚産業有限会社）

「MTAAP の仕様等について」

D. 考察

今回、不明疾患解析検体およびアンプリコン NGS 解析を行う解析ソフト PaHuM ならびに MTAAP の評価を行った。不明疾患解析での NGS 利用は、オズウイルス感染症の報告から、地衛研の業務としての期待が大きい。一方で、解析を担ってきた感染研の Web 解析ソフトが使用できない現状あった。また、海外でも同様の Web 解析サイトはあり、同様の結果は得られるものの、ヒトゲノムを含む可能性もあり、躊躇する地衛研も多い。今回、新たに導入可能なソフトの有用性を示すことができ、今後の地衛研における使用拡大が期待される。特に、MTAAP については、解析ソフトの開発をこちらから持ち込み、共同開発を行った。

結果として、どちらの解析ソフトも機能

的に市販ソフト以上であることが示された。

一般に、NGS 解析ソフトの値段は著しく高い。今回、検討した解析ソフトは、十分に地衛研の NGS 解析に使用可能であることが示された。

E. 結論

不明疾患解析検体およびアンプリコン NGS 解析を行う解析ソフト PaHuM ならびに MTAAP は、地衛研における NGS 解析用ソフトとして、安価かつ高機能であり、利用可能である。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

1) Kaku Y, Okumura K, Kawakubo S, Uriu K, Chen L, Kosugi Y, Uwamino Y, Begum MM, Leong S, Ikeda T, Sadamasu K, Asakura H, Nagashima M, Yoshimura K; Genotype to Phenotype Japan (G2P-Japan) Consortium; Ito J, Sato K.: Virological characteristics of the SARS-CoV-2 XEC variant., Lancet Infect Dis. 2024 Dec;24(12):e736. doi: 10.1016/S1473-3099(24)00731-X. Epub 2024 Nov 6.

2) Kumagai R, Kawakami M, Kasuya F, Yamada K, Takahashi A, Murayama K, Morita K, Kaneyasu T, Tamai N, Hasegawa M, Harada S, Amano A, Fujiwara T, Miyake H, Nagashima M, Sadamasu K.: Sewage surveillance of SARS-CoV-2 in Tokyo (May 2020-June 2022): method using high-throughput, fully-automated platform., Virusdisease. 2024 Sep;35(3):520-524. doi: 10.1007/s13337-024-00880-z. Epub 2024 Jul 2.

3) 黒木 絢士郎：ノロウイルス遺伝子型別における RNA-Seq NGS 解析サイトの活用、東京都微生物検査情報、45, 6-8, 2024

2. 学会発表

なし

H. 知的財産権の出願・登録状況

1. 特許取得

なし

2. 実用新案登録

なし

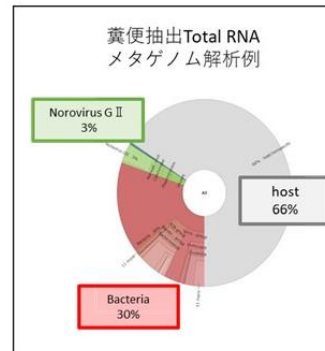
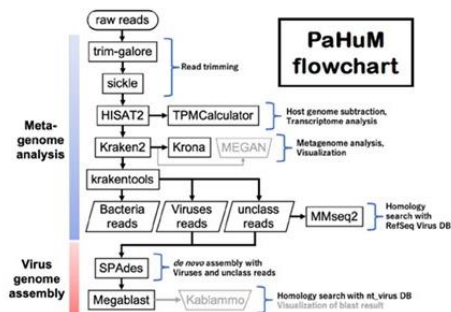
3. その他

なし



メタゲノム解析・ウイルスゲノム解析用パイプライン

- ・メタゲノム中に含まれるホスト（ヒト）ゲノム除去
- ・ウイルスデータベースを用いたウイルスリードの検索



Kraken2 detected virus name	MMseqs2 detected virus name with Kraken2 unclass reads	Num. virus sequence
Norwalk virus (295); Primate norovirus (7)	Norwalk virus (400)	4

Contig length	Coverage read depth	Subject sequence id	Hit seq title
6727	11.0	MK596174.1	Norovirus GI isolate G29-G32 nonstructural polyprotein (ORF1) gene, partial cds; VP1 (ORF2) gene, complete cds; and VP2 (ORF3) gene, partial cds

図1. ウイルス解析用パイプライン (PaHuM) について

ライブラリ調製試薬： SEQuoia Express Stranded RNA Library Prep Kit
 rRNA除去（ヒト）： SEQuoia RiboDepletion Kit
 次世代シーケンサー： iseq 150bp×2 PE
 解析： PaHuM

Sample : 1
 Virus : Norovirus
 GI
 RNA : 糞便乳剤
 Ct : 30.0

Kraken2 detected virus name	MMseqs2 detected virus name with Kraken2 unclass reads	Num. virus sequence
Norwalk virus (295); Primate norovirus (7)	Norwalk virus (400)	4

Contig length	Coverage read depth	Subject sequence id	Hit seq title
6727	11.0	MK596174.1	Norovirus GI isolate G29-G32 nonstructural polyprotein (ORF1) gene, partial cds; VP1 (ORF2) gene, complete cds; and VP2 (ORF3) gene, partial cds
550	3.8	MK596175.1	Norovirus GI isolate G19-G20 nonstructural polyprotein (ORF1) gene, partial cds; VP1 (ORF2) gene, complete cds; and VP2 (ORF3) gene, partial cds
434	1.8	MK596176.1	Norovirus GI isolate BA81265 nonstructural polyprotein (ORF1) gene, partial cds; and VP1 (ORF2) and VP2 (ORF3) genes, complete cds
257	0.8	PQ894187.1	Norovirus GI isolate IC_285 nonstructural polyprotein (ORF1) and VP1 (ORF2) genes, complete cds; and VP2 (ORF3) gene, partial cds

Sample : 2
 Virus : Norovirus
 GI
 RNA : 糞便乳剤
 Ct : 29.8

Kraken2 detected virus name	MMseqs2 detected virus name with Kraken2 unclass reads	Num. virus sequence
Norwalk virus (60); Primate norovirus (44); Shallot latent virus (1); Tobacco mild green mosaic virus (1)	Norwalk virus (110); Primate norovirus (51); Shallot latent virus (2)	10

Contig length	Coverage read depth	Subject sequence id	Hit seq title
826	7.0	PQ632217.1	Norovirus GI isolate NS19217 nonstructural polyprotein (ORF1), VP1 (ORF2), and VP2 (ORF3) genes, complete cds
1906	6.3	MN922742.1	Norovirus GI isolate Hu/GI.3/TW3108/2019/TWN nonstructural polyprotein, RdRp region, (ORF1) and VP1 (ORF2) genes, partial cds
372	5.2	PQ632217.1	Norovirus GI isolate NS19217 nonstructural polyprotein (ORF1), VP1 (ORF2), and VP2 (ORF3) genes, complete cds
710	4.2	PQ632217.1	Norovirus GI isolate NS19217 nonstructural polyprotein (ORF1), VP1 (ORF2), and VP2 (ORF3) genes, complete cds
1498	3.9	PQ632217.1	Norovirus GI isolate NS19217 nonstructural polyprotein (ORF1), VP1 (ORF2), and VP2 (ORF3) genes, complete cds
842	3.3	PQ632217.1	Norovirus GI isolate NS19217 nonstructural polyprotein (ORF1), VP1 (ORF2), and VP2 (ORF3) genes, complete cds
915	3.0	PQ632217.1	Norovirus GI isolate NS19217 nonstructural polyprotein (ORF1), VP1 (ORF2), and VP2 (ORF3) genes, complete cds
291	0.8	MK596174.1	Norovirus GI isolate G29-G32 nonstructural polyprotein (ORF1) gene, partial cds; VP1 (ORF2) gene, complete cds; and VP2 (ORF3) gene, partial cds

図2. PaHuMを用いたノロウイルス解析例

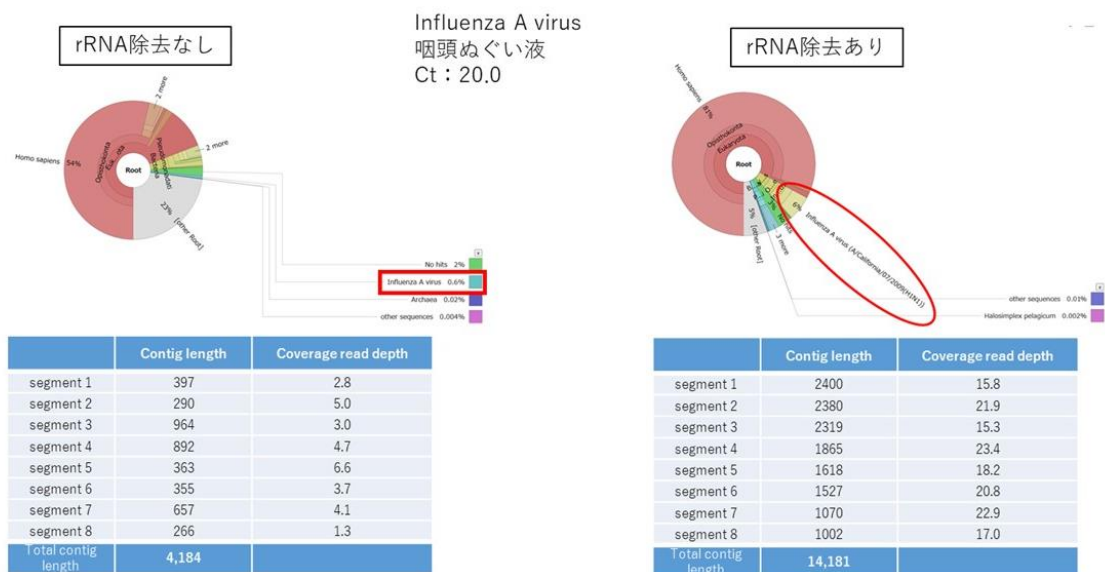
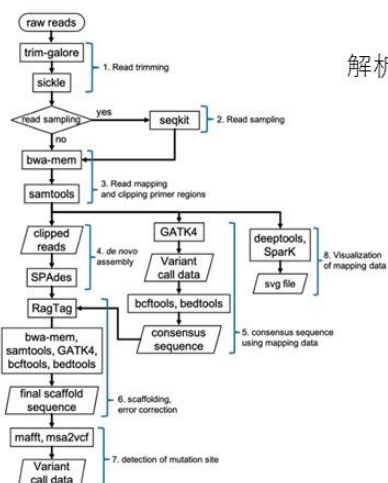


図3. rRNA除去の有無によるインフルエンザウイルス解析例



マルチプレックスPCR 増幅断片解析用パイプライン



解析対象ウイルスゲノムをマルチプレックスPCRで増幅

増幅産物のプライマー部分を除去し、
コンセンサス配列を取得

変異解析

ウイルス特異的な変異の検出

- 大きな欠損
- 混合感染
- ミックスアレル

図4. ウイルス解析用パイプライン (MTAAP) について

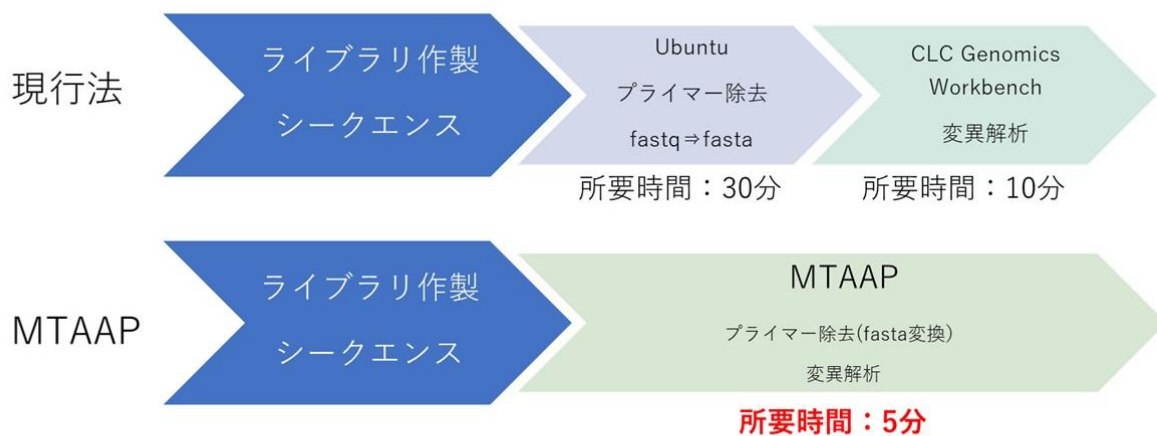


図5.MTAAPを使った新型コロナウイルス解析①



図6. MTAAPを使った新型コロナウイルス解析②