

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

研究協力者 江原 勇登 埼玉県衛生研究所 専門研究員
牧野 由幸 埼玉県衛生研究所 主任
(現 埼玉県熊谷保健所)

研究要旨

網羅的ゲノム解析ツールである MePIC、PaHuM 及び CZ ID を比較し、PaHuM 及び CZ ID が MePIC の後継解析ツールとなり得るか調べた。MePIC で病原ウイルスゲノムを検出した 3 検体の NGS データを用い、PaHuM 及び CZ ID で追加の解析を行い、検出したウイルスゲノムのリード数をそれぞれの解析ツール間で比較した。3 検体の内 2 検体は病原ウイルスを 3 つの解析ツール全てで同リード数検出し、残りの 1 検体はリード数には若干の違いがあるものの、目的とする病原ウイルスを検出した。いずれの解析ツールもメタゲノム解析のツールとして相互の信頼性が高いと考えられた。

A. 研究目的

次世代シーケンサー(以下 NGS という。)を用いた原因不明疾患における網羅的解析は、健康危機管理を担う公的検査機関として非常に重要である。これまで国立感染症研究所ゲノム解析研究センターが開発した網羅的ゲノム解析ツールである MePIC により解析を実施していたが、その解析システムは 2023 年に廃止になったことから、後継となる解析ソフトウェアが必要となっていた。そこで、MePIC の開発者が独立後に新規に開発した PaHuM と、国立感染症研究所や地方衛生研究所で解析実績のある海外の解析プラットフォームの 1 つである CZ ID で解析結果を比較し、後継の解析ソフトウェアとなり得るか調べた。

B. 研究方法

2023 年に、埼玉県の感染症重大事案対策事業として当所に搬入された 15 歳以下の急性脳炎・脳症、無菌性髄膜炎及び不明ウイルス感染症等 6 症例 23 検体のうち、NGS データを MePIC により解析し、病原ウイルスゲノムを検出した 3 検体(検体 1~3)の NGS データを用いた。3 検体の NGS データを、PaHuM 及び CZ ID で追加の解析を

行い、検出したウイルスゲノムのリード数をそれぞれの解析ツール間で比較した。

なお、NGS 検査におけるライブラリーキットは、SMARTer Stranded RNA-Seq Kit HT(タカラバイオ)を使用し、AMPureXP (BeckmanCoulter) で精製後、MiSeq Reagent Kit v3 150bp (Illumina) でシングルエンドシーケンスを行った。

(倫理面への配慮)

本研究は、埼玉県感染症重大事案対策事業の一環として実施した。データの学術的利用については埼玉県衛生研究所倫理審査委員会の審査を受け、許可を得た(承認日: 2018 年 8 月 29 日)。

C. 研究結果

1 検体 1

MePIC では全ウイルス 47 リードのうち 4 リード(9%) からムンプスウイルスを検出、PaHuM では全ウイルス 132 リードのうち 4 リード(3%)ムンプスウイルスを検出、CZ ID では全ウイルスリード数はカウントされないが、ムンプスウイルスが 4 リード検出した。MePIC、PaHuM、CZ ID の全ての解析ツールでムンプスウイルスを 4 リー

ド検出した（表 1）。

表1 各ゲノム解析ツールによるリード数の比較(検体1)

解析ツール	全リード数	全ウイルス	ムンプスウイルス	その他のウイルス
MePIC	711,316	47	4	43
PaHuM	711,316	132	4	128
CZ ID	711,316		4	

2 検体 2

MePIC では全ウイルス 8,937 リードのうち 9 リード (0.1%) からパラインフルエンザ 3 型を検出、PaHuM では全ウイルス 6,055 リードのうち 9 リード(0.1%)パラインフルエンザ 3 型を検出、CZ ID では全ウイルスリード数はカウントされないが、パラインフルエンザ 3 型を 9 リード検出した。MePIC、PaHuM、CZ ID の全ての解析ツールでパラインフルエンザ 3 型を 9 リード検出した（表 2）。

表2 各ゲノム解析ツールによるリード数の比較(検体2)

解析ツール	全リード数	全ウイルス	パラインフルエンザ3型 ウイルス	その他の ウイルス
MePIC	2,926,738	8,937	9	8,928
PaHuM	2,926,738	6,055	9	6,046
CZ ID	2,926,738		9	

3 検体 3

MePIC では全ウイルス 1,136 リードのうちヒトヘルペスウイルス 6 が 57 リード(5%)、ヒトヘルペスウイルス 6B が 1,048 リード (92%) を検出、PaHuM では全ウイルス 1,142 リードのうちヒトヘルペスウイルス 6A を 12 リード(1%)、ヒトヘルペスウイルス 6B を 1,098 リード(96%)検出、CZ ID では全ウイルスリード数はカウントされないが、ヒトヘルペスウイルス 6 を含む突発性発疹ウイルス(Roseolovirus)が 1,063 リード検出した。リード数は異なるものの、MePIC、PaHuM、CZ ID の全ての解析ツールでヒトヘルペスウイルス 6A 及びヒトヘルペスウイルス 6B を含むヒトヘルペスウイルス 6 を検出した。

表3 各ゲノム解析ツールによるリード数の比較(検体3)

解析ツール	全リード数	全ウイルス	ヒトヘルペス ウイルス6	ヒトヘルペス ウイルス6A	ヒトヘルペス ウイルス6B	突発性発疹 ウイルス	その他の ウイルス
MePIC	5,426,312	1,136	57		1,048		31
PaHuM	5,426,312	1,142		12	1,098		32
CZ ID	5,426,312					1,063	

D. 考察

MePIC と PaHuM で比較すると、全ウイルスリード数に違いが見られたことから、解析リードの真核生物、細菌、ウイルスへの分類能には差が見られると考えられた。

病原ウイルスのリード数については、検

体 1 ではムンプスウイルス、検体 2 ではパラインフルエンザウイルス 3 型がそれぞれ同じリード数検出したが、検体 3 ではヒトヘルペスウイルス 6 を検出したものの、リード数やウイルス名の表記名は異なっていた。このことからリード数が増えると解析ツールによる誤差が見られることが示唆されたものの、大きな差はなく、全ての解析ツールで目的とする病原ウイルスを同様な信頼性で検出できると思われた。

E. 結論

検証したデータは少ないものの、MePIC、PaHuM、CZ ID で目的とするウイルスを検出することができたことから、いずれの解析ツールもメタゲノム解析のツールとして相互の信頼性が高いと考えられた。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

なし

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