

タの質が高いことが特徴である。

### C-2: バイオ医薬品とウイルス感染症

図 3 にバイオ医薬品を被疑薬とする、ウイルス感染症に関する有害事象のヒートマップを示す。ヒートマップ上でのデータは、有害事象と各医薬品の ROR 値を以下で示す  $z$  値で正規化し、各有害事象で、どの医薬品に注目すべきかを可視化した。

$$z = \frac{x - \mu}{SE} \quad \text{式 3}$$

ここで、SE は標準誤差、 $\mu$  は平均を示す。

RS ウイルス、ポリオーマウイルス、アデノウイルス等の有害事象を考えると、注意すべきバイオ医薬品が赤で表示されている。RS ウイルスでは、パリビズマブが赤で表示されている。パリビズマブは、抗 RS ウイルスヒト化モノクローナル製剤であり、ROR 値は高くなっているが、パリビズマブにより RS ウイルス感染が起こったというよりは、RS ウイルス感染にパリビズマブが使用されていると考えるのが妥当である。ポリオーマウイルス、アデノウイルス、サイトメガロウイルス、エプスタイン・バーン・ウイルス感染では、免疫抑制剤のバシリキシマブが赤で表示されており、他のバイオ医薬品と比較して ROR 値が高いことが分かる。同様に、ヘルペスウイルスではセルトリズマブペゴル、ウイルス感染その他ではバシリキシマブ、インフルエンザ感染ではアバタセプト、肝炎ウイルス感染ではリツキシマブ、パラボウイルス感染ではウステキヌマブが他の薬剤と比較して ROR 値が高いことが分かる。上部のデンドログラムは、最遠隣法による階層的クラスタリングの結果を示している。パリビズマブは抗 RS ウイルスヒト化モノクローナル

製剤であるパリビズマブは他のバイオ医薬品とウイルス感染関連の有害事象の ROR 値のパターンが大きく異なっている。サイトカインである TNF- $\alpha$ 、IL-6 の阻害薬であるゴリムマブ、アダリムマブ、インフリキシマブ、エタネルセプト、トシリズマブはその ROR 値のパターンが似ていることが分かる。抗 TNF- $\alpha$  抗体であるセルトリズマブ ペゴルは、他の抗 TNF- $\alpha$  抗体と比較して異なるパターンを示している。しかし、セルトリズマブ ペゴルの有害事象報告数はまだ少数（ケースレポート数 157）であるため、今後も注意して観察を続けることが必要であると思われる。ウステキヌマブもセルトリズマブ ペゴルと同様に有害事象報告数（ケースレポート数 158）がまだ少ないものの、IL-12、IL-23 を標的とするウステキヌマブは、パラボウイルス感染の ROR 値が高く、他のバイオ医薬品とまったく異なるパターンを示していた。

図 4 は、図 3 が有害事象方向に各バイオ医薬品の ROR 値を  $z$  値で正規化したのに対し、医薬品方向に各有害事象の ROR 値を  $z$  値で正規化し、各医薬品で、どの有害事象に着目すべきかを可視化した。ウイルス感染の重篤度も考慮すべきであるが、単純に ROR 値で考えると多くのバイオ医薬品でヘルペスウイルスの ROR 値が高い傾向にあることが分かる。

### C-3: バイオ医薬品とウイルス感染症のリスク因子の推定

医薬品と副作用の発現にはリスク因子が存在する可能性がある。リスク因子の推定を決定木を用いて行った。決定木の作成には CART (Classification And Regression Trees) を用い、分岐点の計算に

は以下のジニ係数 (GI; 分布の不純度の尺度) をとして用いた.

$$GI = 1 - \sum_{i=1}^c [p(i|t)]^2 \quad \text{式 4}$$

ここで,  $t$  はノード,  $i$  はクラス,  $p$  は分割されたケースレポートがクラスに属する比率である. この  $GI$  が 0 に近づくようにケースレポートを分類した. 図 4 は, 図 2 でサイトメガロウイルス感染リスクが他のバイオ医薬品と比較して高い可能性が示唆された, バシリキシマブ使用時のサイトメガロウイルス感染のリスク因子に関する決定木分析例である. 2014 年 12 月時点での JADER 中のバシリキシマブが被疑薬となっているレポート 596 件に対し, 最も多い有害事象はウイルス感染症である. その中でもサイトメガロウイルス感染に関する報告は最も多く, バシリキシマブの有害事象報告数の約 45%, 267 件がサイトメガロウイルス感染に関する報告である. バシリキシマブとサイトメガロウイルス感染の ROR 値は 81.6, ROR 値の 95%信頼区間の下限値は 68.3 であった. 決定木の根から葉に向かって図 4 を観察すると, ガンシクロビルが併用されているとサイトメガロウイルス感染のリスクが高いと判断されている. ガンシクロビルは, 言うまでもなくサイトメガロウイルス感染の治療薬であり, サイトメガロウイルス感染のリスク因子というよりは, サイトメガロウイルス感染後にガンシクロビルが投与されたと考えるのが妥当であると考えられる. 機序不明ではあるが, スルファメトキサゾール・トリメトプリム併用 42 例中 25 例 (約 60%) とサイトメガロウイルス感染の割合が高くなっている. スルファメトキサゾール・トリメトプ

リムの併用によってサイトメガロウイルス感染率が高まるか, 本剤を使用している患者が他の感染症にすでに罹患していて, サイトメガロウイルス感染のリスクが高かった可能性が考えられる. 免疫抑制剤であるミコフェノール酸 モフェチルを併用では, 34 症例中 23 症例でサイトメガロウイルス感染が認められ, 免疫抑制剤であるバシリキシマブとミコフェノール酸モフェチルを併用することで, サイトメガロウイルス感染のリスクが高まる可能性があることが分かる. 糖尿病を原疾患にもつ 26 症例中 18 症例でサイトメガロウイルス感染が認められ, 糖尿病がサイトメガロウイルス感染のリスク因子である可能性も示唆された.

#### C-4: ロジスティック回帰分析による医薬品と有害事象の関係のモデル化

図 5 で, バシリキシマブ投与によるサイトメガロウイルス感染のリスク因子と推定された, 併用薬: ミコフェノール酸 モフェチル, 原疾患: 糖尿病についてロジスティック回帰分析でさらに詳細な考察を行った (ケースレポート数, 320106). 医薬品と有害事象の関係は以下のようにモデル化した.

$$\text{logit}(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 \quad \text{式 5}$$

ここで,  $y$  は注目する有害事象の時に 1, それ以外のときに 0 を取る応答変数.  $x_{1,3,\dots,n}$  は注目する医薬品, 事象の時に 1, それ以外の時に 0 を取る説明変数,  $\beta_{1,2,\dots,n}$  はモデルのパラメータとする.

ミコフェノール酸 モフェチル併用時は,

$$\begin{aligned} \text{logit}(y_{\text{CMV}}) \\ = -5.27697 + 1.63176x_{\text{Basiliximab}} \\ + 3.25054x_{\text{Mycophenolate mofetil}} \end{aligned}$$

となり,  $\beta_{\text{Basiliximab}}, \beta_{\text{Mycophenolate mofetil}}$  は

共に有意に 0 より大きかった (95% Wald 信頼区間が 0 を含まない確率  $2e-16$  以下). ミコフェノール酸 モフェチル, バシリキシマブは共に免疫抑制剤であり, 併用時はサイトメガロウイルス感染のリスクが高まることを示しており, バシリキシマブ (推定オッズ比  $\exp(1.63176) = 5.1$ ) よりもミコフェノール酸 モフェチル (推定オッズ比  $\exp(3.25054) = 25.8$ ) によるサイトメガロウイルス感染のリスクが大きいと推定された.

糖尿病を原疾患に持つ場合,

$$\begin{aligned} \text{logit}(y_{CMV}) = & -5.08885 \\ & + 4.56326x_{\text{Basiliximab}} \\ & - 0.29009x_{\text{Diabetes}} \end{aligned}$$

となり,  $\beta_{\text{Basiliximab}}$  は有意に 0 より大きかった (95% Wald 信頼区間が 0 を含まない確率  $2e-16$  以下). 一方で,  $\beta_{\text{Diabetes}}$  はわずかではあるが有意に 0 より小さかった (推定オッズ比  $\exp(-0.29009) = 0.75$ ). これは, 図 5 で糖尿病がバシリキシマブによるサイトメガロウイルス感染のリスク因子であるとする推定と矛盾する. 式 5 が, 糖尿病を原疾患とする患者における, バシリキシマブ使用によるサイトメガロウイルス感染を十分にモデル化できていない可能性が考えられた.

#### C-5: ロジスティック回帰分析医薬品相互作用の検討

医薬品と有害事象の関係を医薬品と医薬品, 医薬品と疾患等の相互作用を想定したモデルでさらに詳細に解析した.

$$\begin{aligned} \text{logit}(y) = & \beta_0 + \beta_1x_1 + \beta_2x_2 \\ & + \beta_{12}x_1x_2 \quad \text{式6} \end{aligned}$$

式 5 同様,  $y$  は注目する有害事象の時に 1, それ以外のときに 0 を取る応答変数.  $x_{1,3,\dots,n}$

は注目する医薬品, 事象の時に 1, それ以外の時に 0 を取る説明変数,  $\beta_{1,2,\dots,n}$  はモデルのパラメータとした.  $\beta_{12}x_1x_2$  項は相互作用項である.

ミコフェノール酸 モフェチル併用時は,

$$\begin{aligned} \text{logit}(y_{CMV}) = & -5.29939 + 4.33798x_{\text{Basiliximab}} \\ & + 3.39607x_{\text{Mycophenolate mofetil}} \\ & - 2.93630x_{\text{Baxililimab}}x_{\text{Mycophenolate mofetil}} \end{aligned}$$

となり,  $\beta_{\text{Baxililimab}}x_{\text{Mycophenolate mofetil}}$  は  $-2.93530 (< 0)$  となり, バシリキシマブとミコフェノール酸 モフェチルは相互作用は観察されなかった. 一方, 糖尿病を原疾患に持つ場合,

$$\begin{aligned} \text{logit}(y_{CMV}) = & -5.07909 + 4.50097x_{\text{Basiliximab}} \\ & - 0.47945x_{\text{Diabetes}} \\ & + 1.10410x_{\text{Baxililimab}}x_{\text{Diabetes}} \end{aligned}$$

となり,  $\beta_{\text{Baxililimab}}x_{\text{Diabetes}}$  は 1.10410 と有意に 0 より大きかった (95% Wald 信頼区間が 0 を含まない確率  $8.14e-06$ ) で, バシリキシマブと原疾患: 糖尿病の相互作用の推定オッズ比は  $\exp(1.10410) = 3.02$  となった. 相互作用を想定した場合, 糖尿病がバシリキシマブによるサイトメガロウイルス感染のリスク因子となり得る可能性が示された.

#### D. 考察

JADER には, 身長, 体重, 生活習慣に関するデータがあまり含まれておらず, これらの影響を考察することが難しい. また, 副作用の発生頻度の算出が困難等の限界がある. しかし, 網羅的に効率よくデータ解析を行うことで, 有用な情報を抽出できる可能性がある. 本研究では, JADER

中のバイオ医薬品とウイルス感染症に関する ROR 値を可視化し、決定木分析、ロジスティック回帰分析により、ミコフェノール酸 モフェチル併用がバシリキシマブ使用時のサイトメガロウイルス感染のリスク因子であること、糖尿病の治療に使用している医薬品がリスク因子となっている可能性も否定できないが、糖尿病がバシリキシマブにおけるサイトメガロウイルス感染のリスク因子の一つである可能性を示すことができた。

#### E. 結論

JADER データベースを構築するソフトウェアを開発し、効率的に解析を進めることのできるシステムを構築した。本有害事象自発報告データの解析システムは、バ

イオ医薬品のウイルス感染性因子の安全性評価のみならず、すべての医薬品の安全性評価に活用できるものとする。

#### F. 研究発表

1. 論文発表
  - 1) 川崎ナナ：生物製品の局方収載の現状と課題. レギュラトリーサイエンス学会誌, vol.4, No.2, 149-154 (2014)
2. 学会発表  
なし

#### G. 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

表 1. 2 × 2 クロス集計表. JADER データベース中で医薬品 X が被疑薬の有害事象 Y の件数を a, 医薬品 X が被疑薬で有害事象が Y 以外である件数を b, 医薬品 X 以外が被疑薬の有害事象 Y の件数を c, 医薬品 X 以外が被疑薬で, 有害事象も Y でない件数を d とカウントする.

	Adverse event (Y)	Not adverse event (Y)	Total
Using drug X	a	b	(a + b)
Not using drug X	c	d	(c + d)
Total	(a + c)	(b + d)	(a + b + c + d)

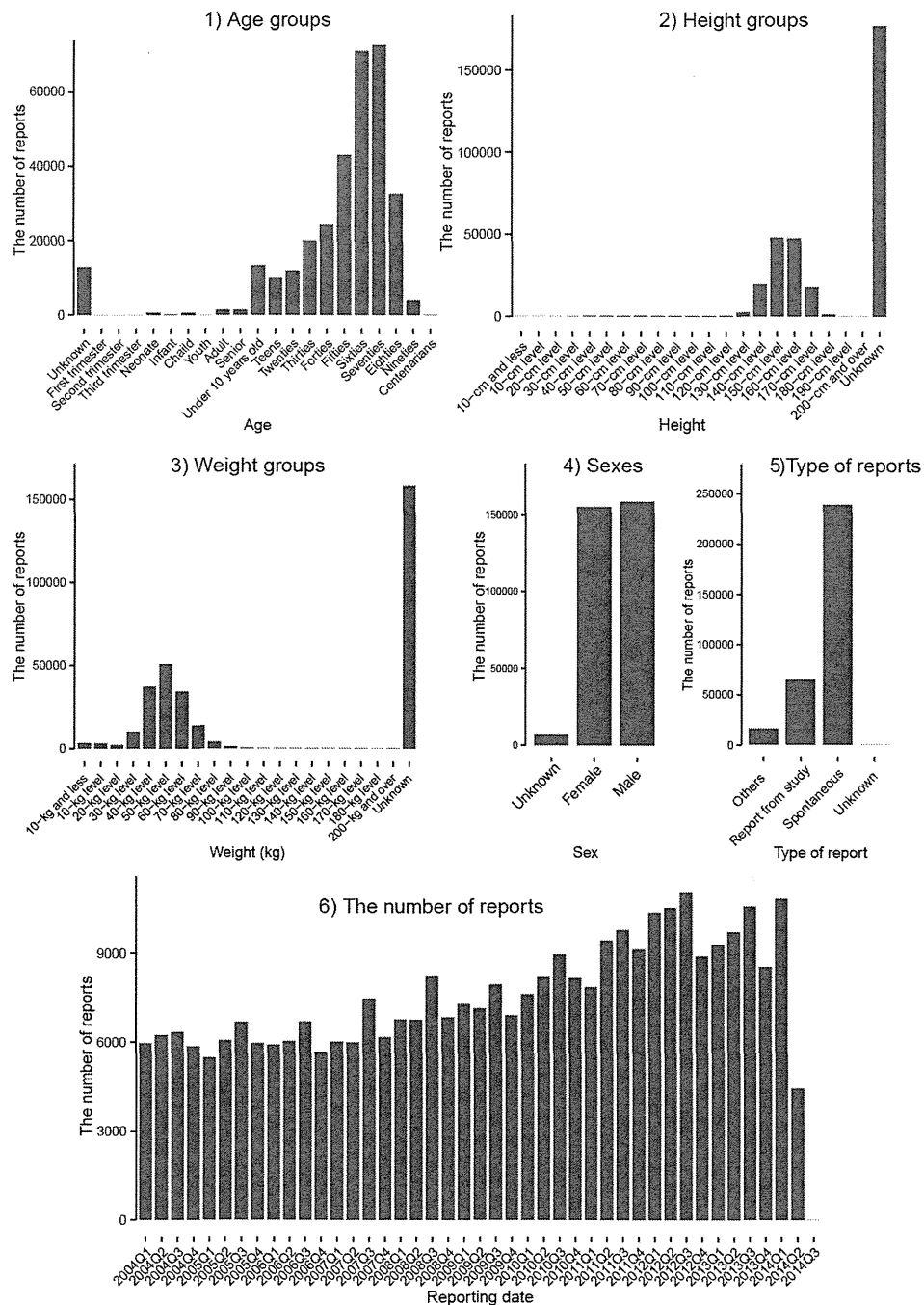
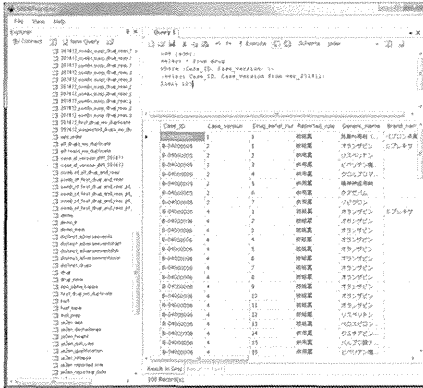


図 1. JADER の統計情報。年齢の分布 (1), 身長(2), 体重 (3) の分布, 性別 (4), 報告の種類 (5), 報告数の推移 (6)を示した。

a) データベース作成, 管理ソフトウェア



b) Webベース, 有害事象解析システム



図 2. 作成した JADER リレーショナルデータベース作成, 管理ソフトウェア (a) と web ベース有害事象解析システム. a) PMDA で公開されている CSV ファイルから RDB を作成し, 着目すべき任意のレポートを抽出できる. b) 様々な解析を web ベースで行うことができる.

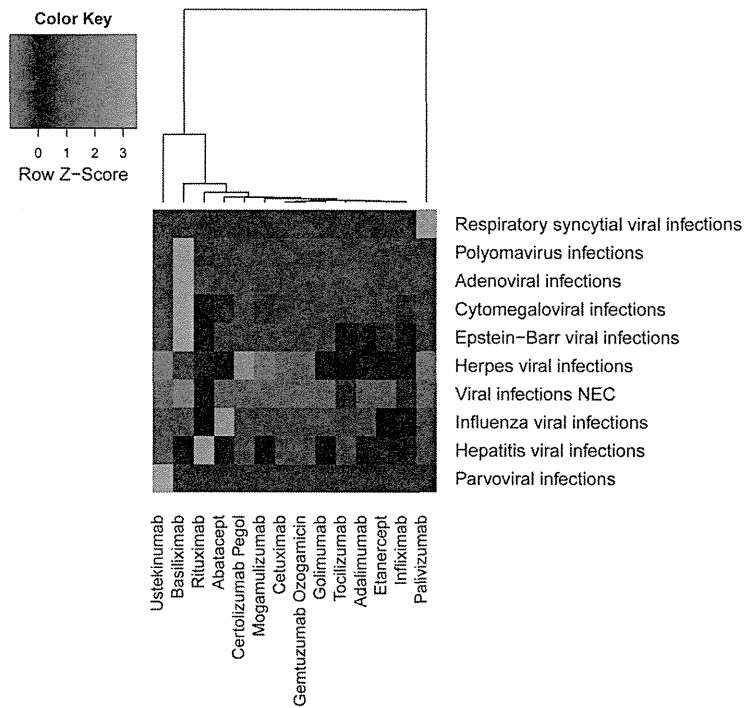


図 3. ヒートマップを用いた着目すべき薬剤の可視化。有害事象ごとに各医薬品の ROR を Z スコアで正規化している。赤に近いほど Z スコアが高く、優先的に考察すべき薬剤であることを示す。



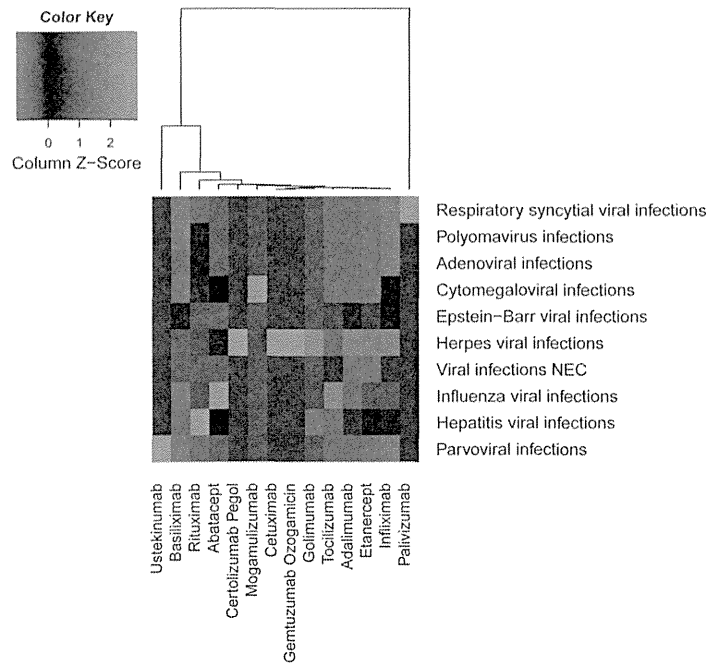


図4. ヒートマップを用いた着目すべき有害事象の可視化. 有害事象ごとに各医薬品の ROR を Z スコアで正規化している. 赤に近いほど Z スコアが高く, 優先的に考察すべき有害事象であることを示す.

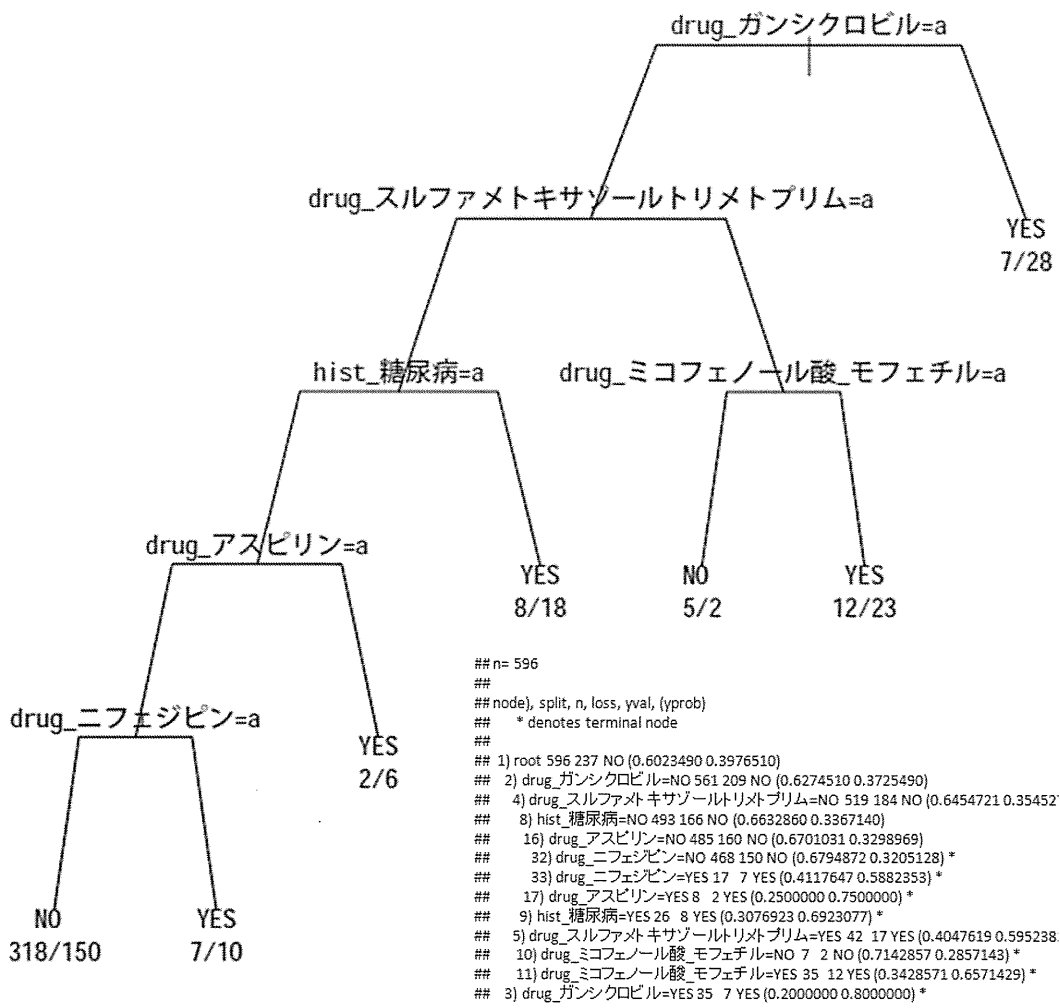


図 5. 決定木を用いたバシリキシマブ使用時の、サイトメガロウイルス感染のリスク因子の推定.

厚生労働科学研究費補助金（医薬品等規制調和・評価研究事業）  
分担研究報告書

細胞組織加工医薬品のウイルス安全評価に関する研究  
研究分担者 遊佐敬介 国立医薬品食品衛生研究所 生物薬品部

研究要旨： 細胞組織加工製品のウイルス安全性の確保は、製品の安全性の重要な課題のひとつである。しかし細胞組織加工製品の製造工程ではバイオ医薬品の場合と異なり、細胞や組織が製品そのものであるため、ウイルス不活化・除去のための処理工程を組み込むことが困難な場合が多い。そこで少量の検体を使って短時間、高感度に網羅的にウイルスを検出するシステムの開発を行った。本研究では次世代シーケンサー(NGS)による RNA-seq データを用いて細胞のウイルス感染の有無を判定するためのパイプラインを作った。この手法にさらに精密で迅速な改良を行うことで、細胞組織加工製品の製造に用いられる細胞ストック (iPS, ES 細胞等), フィーダー細胞, 中間製品, 最終製品のウイルス試験として実用化が可能であろうと考えられる。

研究協力者

苑 宇哲 国立医薬品食品衛生研究所 生物薬品部

A. 研究目的

細胞組織加工製品のウイルス安全性をどのように確保するは、製品の安全性の重要な課題のひとつである。バイオ医薬品では、医薬品製造用細胞に関して厳密なウイルス試験が行われる。レトロウイルス試験や *in vitro* ウイルス試験, *in vivo* ウイルス試験, 抗体産生試験等である。また未精製バルクでも培養中に迷入した可能性を考慮しウイルス試験が行われる。これに続く精製工程でもウイルス不活化・除去工程が組み込まれている。また精製工程のウイルスクリアランス能を評価することによって、製品へのウイルス混入の可能性をさらに低減させている。これに対して細胞組織加工製品では、ドナーの適格性、あるいは製造に携わる人の安全のために原材料となる細胞組織のウイルス試験を行うことになっている。しかしそれ以降の工程、即ち中間製品や最終製品に関する効果のあるウイルス不活化・除去やウイルス試験に関する具体的な試験法については今後の課題となっている。

細胞組織加工製品を汚染する可能性のあるウイルスは、そのウイルス侵入経路の違いから、三つに分類して考えることができる。一つめは、原材料となるヒト細胞・組織からのウイルスの迷入である。これはドナーの適格性の観点から、HBV, HCV, HIV, HTLV 等をはじめとしていずれもヒト感染ウイルスが対象である。二つめは、製造工程全般を通じたウイルスの製品への「飛び込み」である。飛び込みウイルスに関しては、無菌操作や設備等の問題と密接に関係してい

るので本研究では特別に対象としない。三つめは、生物由来原料を介して汚染されるケースである。バイオ医薬品の製造工程でのウイルス汚染の事例では、汚染は動物由来成分を含む原材料を介してバイオリクターを汚染するケースがいくつか報告されている。従って細胞組織加工製品でも、三つめの経路をどう防ぐかについての方策を立てることが重要となる。フィーダー細胞, 製造工程以降の細胞ストックや中間製品, 最終製品に関するウイルス試験法をより感度が高く、信頼性に富んだものにするためには、新規の検出系が必要になる。ウイルス検出法として感度が優れ、短時間で多数のウイルスを検出できる方法に核酸増幅法がある。しかしこの手法は検出すべきウイルス個々に特異的なプライマー設定が必要なため、検出できるウイルスの種類が限定されてしまうという欠点がある。また少量のサンプルで迅速、簡便、高感度にウイルスを検出法の選択肢を増やしておく必要から新規のウイルス検出システムの検討をしておくことは、将来の多様な細胞組織加工製品のウイルス安全性をより確かなものとする。今後多様な製品が実用化されるようになるとウイルス安全性に関して既存の手法で対応できないタイプの製品が開発された可能性があるからである。細胞（細胞ストック, フィーダー細胞, 中間製品, 最終製品等）のウイルス感染の有無を調べる場合、製造スケールが小さく、簡便に、且つ高感度にウイルスを検出する必要が出てくる。バイオ医薬品を想定した ICH Q5A 等にある方法は大量製造の製品を想定

した試験法であり、必ずしも細胞組織加工製品の製造工程のウイルス安全性評価法として最適とはいえない場合が想定されるからである。そのために、本年度は次世代シーケンサー（NGS）を使ったウイルス検出法について検討し、RNA-seq データを用いて細胞ストック、フィーダー細胞、中間製品、最終製品等のウイルス感染の有無を高感度、網羅的に検出するシステムが構築できるかどうかを検討した。

## B. 研究方法

### (1) 細胞とウイルス

HEK293 細胞、SNL76/7 細胞は ATCC から購入した。FCV (F4 株) は野田博士 (衛研) からいただいた。また MVM は山口博士 (衛研) からいただいた。FCV は、m. o. i. 0.1 で CRFK 細胞に感染後、2 日めの上清を 0.22  $\mu\text{m}$  のフィルターでろ過したものを  $-80^{\circ}\text{C}$  で保存した。MVM は、m. o. i. 0.1 で A9 細胞に感染後、2 日めの上清を 0.22  $\mu\text{m}$  のフィルターでろ過したものを  $-80^{\circ}\text{C}$  で保存した。ウイルス保存液の感染感受性の測定は、FCV は CRFK 細胞、MVM は A9 細胞を使って行った。

### (3) 次世代シーケンサーによる解析及び RNA-seq データの解析

非感染 HEK293 細胞 ( $1 \times 10^6$ )、FCV を m. o. i. 1 で感染させ、12 時間、24 時間後に RNeasy mini kit (Quiagen) を用いて RNA を分離精製し、6  $\mu\text{g}$  RNA を出発材料として、HiSeq 2000 で  $2 \times 100$  bp、50M リードサイズの解析を行った (受託サービス Eurofin Genomics)。解析データを使って、先ずヒト遺伝子転写物を同定し、続いて NCBI のウイルスデータベース viral. 1. 1. を用いてウイルス様シーケンシスの検索を行った。

### (倫理面への配慮)

ヒト由来試料は使わないので倫理面の配慮は必要ない。

## C. 研究結果

### (1) 次世代シーケンサーを使ったウイルス検出法

次世代シーケンサーを使ったウイルス検出法の概略を示したのが図 1 である。6  $\mu\text{g}$  RNA を出発材料としてサンプルを調製し、これを次世代シーケンサーによって解析を行い

RNA-seq データを得た。このデータの中からウイルスシーケンシスが検出できるかどうかによって、細胞のウイルス感染の有無を判定する。通常培養細胞に含まれる RNA 量は少ない細胞で 1  $\mu\text{g}/10^5$  細胞、多い細胞では 3.5  $\mu\text{g}/10^5$  細胞といわれている。従って細胞数は数十万個あれば解析可能である。RNA を使った解析を行う理由は、DNA ウイルスも RNA ウイルスともに感染複製する際には、必ず RNA を産生する必要があるためである (図 2)。この検出系を作る際に重要な点は、どんな感度でどんな範囲のウイルス種が検出できるのかという点である。

### (2) モデル細胞 HEK293 とモデルウイルス FCV<sub>F4</sub> を使った検討

ここではモデル細胞としてヒト細胞株である HEK293 細胞、モデルウイルスとしてカリシウイルスの一種である FCV を用いた。非感染 HEK293 細胞とウイルス感染 HEK293 細胞の RNA-seq データの解析を行い、両者を比較することで、感染ウイルスのシーケンシスを検出できるかどうかを調べた。

非感染 HEK293 細胞とウイルス感染後 12 時間と 24 時間後の HEK293 細胞から RNA を分離し、poly(A)+RNA を精製した後、次世代シーケンサー用試薬を使ってサンプル調製を行い、Illumina 2500 を使って解析を行った。HEK293 細胞の RNA-seq データは 200 nt に満たない短いリードを 23M 本含み、88%以上のデータは比較的よいクオリティを持っていることがわかる (表 1)。

### (3) RNA-seq データの解析

非感染 HEK293 細胞の解析の結果、得られたこのシーケンシスを次の 3 つカテゴリーに分類した。① host seqs, ② unmapped seqs, ③ virus-like seqs である。得られたもののうち 98.3%は host seqs であった (図 3)。また unmapped seqs は 1.1%であり、virus-like seqs とされたものは 0.6%であった。非感染細胞には、内在性レトロウイルスや本来は host seqs だが、ウイルスによく似たシーケンシスが含まれている可能性がある。そこで virus-like seq に分類された 0.6%の内訳を調べてみると、99.7%アデノウイルス、0.01% ERV であった。残りの 0.29% は host seq 由来ではあるがウイルスのシーケンシスに似ているため virus-like seqs として同定された。アデノウイルスシーケンシスが高頻度に検出された訳は、HEK293 細胞を樹立する際に、トランスフォームの目的で adenovirus 5 を用いて

いたためである。HEK293 細胞の 19 番目の染色体には 4 kb ほどのアデノウイルスゲノムが組み込まれていることがわかっており、そのためアデノウイルスシーケンスが RNA-seq データに含まれていたものと考えられた。以上から非感染細胞には、ウイルス検出を難しくする高いバックグラウンドが存在することがわかった。

カリシウイルスを感染細胞の RNA-seq データで virus-like seqs と判別されたものは、非感染のものと同様に全 RNA-seq データの 0.6% を占めた。その内訳を見ると、感染後 12 時間、24 時間でそれぞれカリシウイルスが 0.21%、0.09% 検出することができた。しかしアデノウイルスに比べて検出頻度が極端に低く、感染後 12 時間では、アデノウイルスに比べて 1/500 程であった (図 4, 図 5)。そこで検出できたカリシウイルスのゲノムに対してマッピングを行ってみると、矢印で示されるようにマッピングできたのは、ゲノムのごく一部であることがわかった (図 6)。カリシウイルスは、RNA ウイルスの一種であり、数多くの株が存在し、多様性が高い。ここで用いたウイルスデータベース viral. 1. 1. にはそのうちの 1 株 Urbana しか含まれていなかった。そのため Urbana 株とホモロジーの高い一部のシーケンスしか検出できなかったことが明らかになった。そのためカリシウイルスの多くの株を含む拡張データベースを作成し、解析に用いたところ、カリシウイルス F4 株のゲノム全長をカバーする多くのリードを検出することができた。この拡張ウイルスデータベースを用いて再度 virus-like seqs フラクションを解析すると、感染細胞から作成した RNA-seq データのみカリシウイルスが、35.8%FCV, 17.9% 検出された (図 7)。改善の余地はあるが、NGS を用いてウイルス検出システムを構築することができるものと考えられた (図 8)。

#### (4) 次世代シーケンサーを使ったウイルス検出法で注意すべき点

以上から、次世代シーケンサーを使ったウイルス検出法に関して考慮しなければならない点が明らかになった。問題点として以下が挙げられる：① NSG によるシーケンシングのサイズと品質、② 適切なデータ解析プログラムの選択、③ ウイルスゲノムデータベースの拡張、④ 宿主細胞 RNA-seq に含まれるウイルス様配列 (tandem repeat, poly (A) などウイルス由来と判定される配列) をいかに除くか、⑤ 内在性レトロウイルス由来 seqs を外来性レ

トロウイルスと判別、⑥ 解析にかかる時間とコスト、⑦ 検出感度、検出限界、他の検出法との比較。

① NSG によるシーケンシングのサイズと品質について検討を行った (図 9)。次世代シーケンサーのデータサイズが、検出感度に関わる重要な要因であることを示している。当然リード数が大きくなると、それぞれの転写産物のカバレッジも増加するが、リード数を上げると、それに従って解析コストも増加するからである。この点についてはさらに検討を進める必要があるが、リード数を 5M に下げても、ウイルスゲノムをカバーすることができることがわかった。

② 適切なデータ解析プログラムの選択については表 2 にまとめた。RNA-seq データからウイルス様シーケンスを同定するのにかかった時間と得られた結果についてまとめたものである。tophat では解析に 66 時間かかるが、star ではわずか 3 分で解析を終了することができる。このように使用する解析プログラムによって解析にかかる時間に大きな差が出る。今後さらに効率の良い解析プログラムが利用できるようになる可能性があり、システムの改良と共に解析時間を短縮できると考えられた。

③ ウイルスゲノムデータベースの拡張については、既に述べたように多様性のあるウイルスに関しては、解析に使うウイルスデータベースを拡張し、検出できるウイルスの範囲を広げる必要がある。最新のウイルスデータを加えて改訂していく必要がある。

④ 宿主細胞 RNA-seq に含まれるウイルス様配列をいかに除くか、と⑤ 内在性レトロウイルス由来 seqs を外来性レトロウイルスと判別、⑦ 検出感度、検出限界、他の検出法との比較、については、ウイルス由来シーケンスを判別する際の重要な課題であり、今後検討する必要がある。

#### (5) 次世代シーケンサーを使ったウイルス検出法のコストについて

⑥ 解析にかかる時間とコストも実用化を睨んで重要な問題である。次世代シーケンス解析には、高額な機器が必要であり、実用化をにらんで価格面に関しての検討も必要になってくる。試薬のコストについてまとめたのが、表 3 である。近年の次世代シーケンサーの発達と改良は目を見張るものがある。解析に要する時間も Illumina の製品を例にとると、HiSeq 2500 では 40 時間かかるが、MiSeq 500 では 4 時間程に短縮できる。また価格面でも解析器材

さえそろえば、実質1サンプルにかかる試薬の費用は、約3万円程である。人件費等を併せても、1サンプル15万円程度に抑えることができるのではないかと考えられる。また器材の進歩と共に次世代シーケンサーによる解析の受託サービスは、今後は大きく価格ダウンすると思われる。また解析器材さえあれば、解析を行うのに必要な時間は、最短2日程度なると考えられる。そのためには、データ解析に要する時間を数時間程度にする必要があり、virus-like seq と判定するのに間違いがなくなるように判定に必要な知見の集積が今後必要である。

#### (5) 細胞組織加工製品製造に使われる細胞のウイルス試験

今回樹立したウイルス検出パイプラインを用いて、実際に細胞組織加工製品製造に使われる細胞のウイルス感染有無を判定できるかどうかについて調べた。ここではまずフィーダー細胞として使われる SNL76/7 細胞を取り上げた。この細胞にウイルス感染させて、それが検出できるかどうかを調べた。次世代シーケンサーによる解析の結果、今回は5,000万本以上の short reads が得られた(表4)。SNL細胞は通常放射線照射やマイトマイシン処理をして使うので、 $\gamma$ 線照射細胞についても実験を行った。得られたリードのうち99%近くがホストのシーケンスでした。またウイルス様シーケンスと判定されたシーケンスは全体の0.0018~0.0041%を占めることがわかった(図10)。ここで得られたごくわずかなウイルス様シーケンスをさらに詳細に調べると、非感染細胞では88%が内在性レトロウイルスであると判定された(図11)。またマウスパルボウイルス感染細胞では、パルボウイルスがウイルス様シーケンスと判定された0.0041%のうち76.2%がパルボウイルスシーケンスであると判定された。ここで検出されたリードをパルボウイルス遺伝子の全長にマッピングしてみると、全長にわたってマッピングできた。このことから、感染細胞からパルボウイルスが間違いなく検出できたことになる。

#### D. 考察

##### (1) 細胞組織加工製品のウイルス安全性確保

生物由来原料を使って製造される医薬品としてバイオ医薬品がある。バイオ医薬品製造におけるウイルス試験法に関してはガイドラインであるICH Q5A「ヒト又は動物細胞株を用いて製造されるバイオテクノロジー応用医薬

品のウイルス安全性評価」に詳しい。バイオ医薬品の製精製工程では、低pH処理等のウイルスの不活化のための工程やウイルス除去を行うためのウイルスフィルター処理等によってウイルス安全性を確保している。これに加えて製造用細胞や未精製バルクなどでのウイルス試験法、また精製工程のウイルスクリアランス値を評価することでさらに安全性を高めている。これに対して細胞組織加工製品では、最終製品が組み換えタンパク質等ではなく、「細胞」そのものであるため、低pHやウイルスフィルターといった処理をすることは難しい。従って細胞組織加工製品に使われる生物由来原料に関するウイルス安全性を出来るだけ高めることが必要になってくる。またそれと同時に中間製品、最終製品、またフィーダー細胞等について簡便で、高感度、短時間で幅広いウイルスを低コストで行えるウイルス試験法も重要になってくる。ICH Q5Aに記載されている試験法や安全性評価法は、生産規模の大きな製品を想定したものである。細胞組織加工製品に応用できるものも少なくないが、必ずしも細胞組織加工製品に最適な試験法とは限らないものがある。特にウイルス試験法は、in vitro 試験、in vivo 試験ともに試験法に時間がかかる。また最終製品は時間をおかずに治療に使うことが想定される。そのため柔軟性に富んだ安全性の確保が必要になってくる。ウイルス試験法の選択肢も増やしていく必要が出てくるだろう。その意味でも今回検討したNGSを利用したウイルス検出法は、実用化に向けて重要な選択肢の一つとなるだろう。

##### (2) フィーダー細胞のウイルス試験法

国内の細胞組織加工製品製造に使われているフィーダー細胞3T3-J2のウイルス試験についてみていくとマスターセルバンクでは、延長S+L-アッセイ、延長XCプラークアッセイ、電子顕微鏡観察、逆転写酵素活性試験、Vero細胞等を用いた細胞変性誘発試験、赤血球吸着及び施血球凝集反応、モルモット、ニワトリ受精卵等を用いたin vivo 試験、マウス抗体産生試験、ウシ鼻甲介細胞等を用いたウシ由来迷入ウイルス試験が実施されている。またワーキングセルバンクもほぼ同様の試験が行われた。これに加えて、X線照射法により作製し、フィーダー細胞についてレトロウイルス否定試験が行われている。モルモットなど小動物をつかったin vivo試験では判定に28日を要する。またウシ由来迷入ウイルス試験ではCPEの出現の有無を14日にわたって調べるなど試験には長い時

間とコストを要する。本研究で検討したNGS法を用いたウイルス試験法は、ウイルスの感染性を重視したバイオアッセイではないため、ウイルスを取り扱う設備を必要としない。そのためウイルス非感染が明らかなマスターセルバンクがあれば、その細胞と比較することで、ワーキングセルバンクのRNA-seqデータの中に外来性のウイルスが含まれているのかを短時間で判定することができるだろう。次世代シーケンサーの機能は近年驚く程向上しており、コストの大幅なダウンも期待されるため、新規のウイルス試験法としてまず補助的に利用することを考えるべきだろう。実用化に向けては再現性や検出限界等に関する一層の検討が必要になる。

### (3) 他のウイルス試験法との比較

ICH Q5A「ヒト又は動物細胞株を用いて製造されるバイオテクノロジー応用医薬品のウイルス安全性評価」に記載されているウイルス試験法のうち、in vivo試験法、in vitro試験法についても概ね時間がかかる試験法が多い。これはフィーダー細胞のところでも述べたように、ウイルス試験法がウイルスの感染性を重視したバイオアッセイであるためである。ウイルス試験法には核酸増幅法があるが、この方法は、感度が高く、ごく短時間で結果が出る反面、結果とウイルス感染性が必ずしも一致しない場合がある。これは核酸増幅法がウイルスの核酸を標的にしているため、鋳型となる一部の核酸断片が存在すると陽性となる場合やすでに不活化しているウイルスも検出してしまうためである。今回検討したNGS法を使ったウイルス検出法では、その感度が核酸増幅法と比べてどうかという点に関しては、詳細な検討を行わなかった。しかし、核酸増幅法では標的ウイルスを増幅するプライマーは正確でなくてはならず、RNAウイルスなど多様性をもつウイルスに対して注意が必要である。今回の実験では、カリシウイルスを検出したが、特定の株Urbanaしか含まないウイルスデータベースでもF4株を検出することができた。このことは、知られていない未知のウイルスのシーケンスでもホモロジーが高いドメイン構造から検出できる可能性があることを示している。従って、ウイルスデータベースを一層充実させることで検出できるウイルスの範囲がさらに広がる可能性があるものと考えられた。加えて大きな特徴は、ウイルス由来シーケンスと判定されたリードをウイルスゲノムにマッピングすると、全ゲノムにわたる転写の確認が出来るばかり

でなく、発現レベルを推定することができる。これに対して核酸増幅法では、特定の領域の転写量を知ることができるが、ゲノム全体を見ることができない。

短所として挙げられるのは、高額なシーケンサーと試薬類である。ここ2、3年の動向を見ていると、より小型化、よりサンプル数の少ないパーソナルな解析システムが使えるようになってきている。よって試験法を取り巻く環境は将来的には好転していくものと考えられる。また解析過程に於ける問題点として、宿主由来のシーケンスではあるが、ウイルス様シーケンスに似ていたり、単純な繰り返し配列のためにウイルス由来のシーケンスだと判定されてしまう場合がある。そのため効率よく偽ウイルスリードを除くことができるように改善が今後とも必要である。この過程には、専門知識をもった技術者を必要とするが、解析プログラムの改良によって、効率よく且つ専門知識なしに、ウイルス感染の有無を判断できるようなパイプライン実現が可能であると考えられた。

## E. 結論

細胞加工製品のウイルス安全性を確保するために細胞ストック、フィーダー細胞、中間製品、最終製品等のウイルス感染の有無を高感度、網羅的に検出できるRNA-seqデータ解析のためのパイプラインを作ることができた。しかし検出限界と検出できるウイルスの範囲を向上させるためにはクリアすべき課題が残っている。

## F. 研究発表

### 1. 論文発表

1) なし

### 2. 学会発表

- 1) 苑宇哲, 前田洋助, 川崎ナナ, 原田信志, 遊佐敬介 マウス微小ウイルスの核への侵入にはホスホリパーゼ A2 活性が必要である: 第 62 回日本ウイルス学会学術集会 (2014. 11. 10-12) 横浜
- 2) 前田洋助, 寺沢広美, 光浦智将, 中島詩織, 門出和精, 田中勇悦, 遊佐敬介, 原田信志 HTLV-1 Env 発現産生細胞における HTLV-1 受容体 GLUT1 の制御機構の解析: 第 62 回日本ウイルス学会学術集会 (2014. 11. 10-12) 横浜

## G. 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



**細胞**  
(細胞ストック, フィーダー細胞, 中間製品, 最終製品等)



**RNA 5  $\mu$ g (2~5 $\times$ 10<sup>5</sup> cells)**



**NGS による解析**



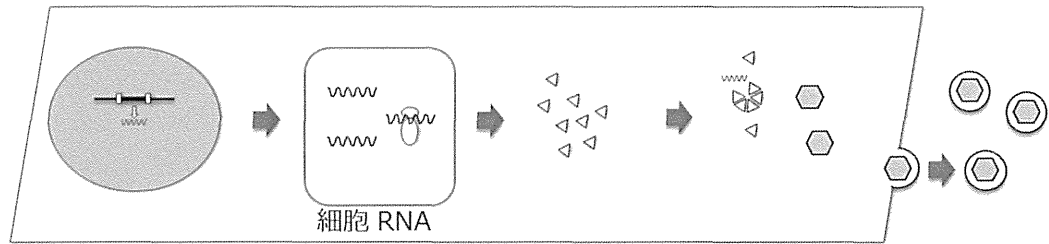
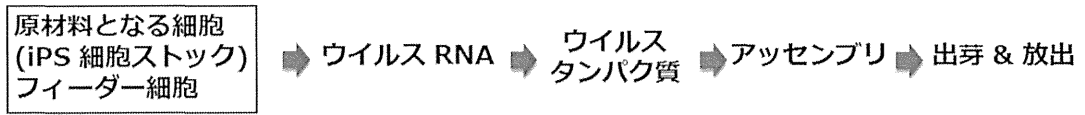
**RNA-seq data 解析**



**virus seqs?**

**Yes or No**

図1 次世代シーケンサー(NGS)を使ったウイルス検出法



例 レトロウイルス感染細胞

図2 次世代シーケンサー(NGS)を使ったウイルス検出法

<b>Samples</b>	<b>No. of reads</b>	<b>length/read (nt)</b>	<b>Bases (%&gt;=30)</b>
<b>HEK293</b>	<b>23M</b>	<b>194</b>	<b>2.310G (88.47)</b>
<b>HEK293 + FCV 12 h</b>	<b>33M</b>	<b>195</b>	<b>3.316G (88.72)</b>
<b>HEK293 + FCV 24 h</b>	<b>31M</b>	<b>194</b>	<b>30.96G (88.35)</b>

表 1 HEK293 細胞, FCV 感染 HEK293 細胞 RNA-seq データ (Illumina HiSeq 2500 使用)

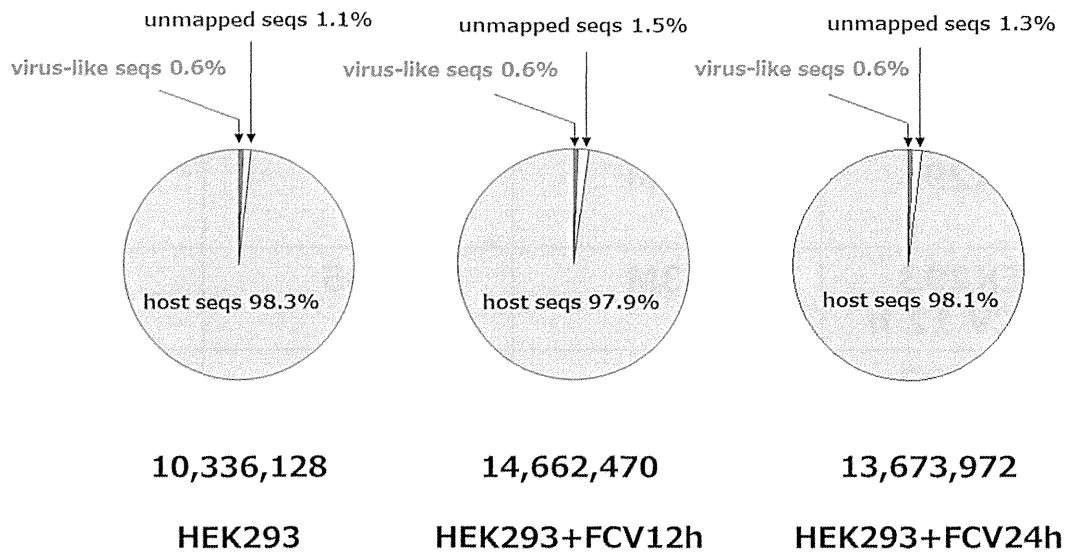


図3 HEK293 細胞, FCV 感染 HEK293 細胞の RNA-seq データに含まれるウイルス様シーケンス