

ウイルスだけでなく DNA ウィルスも検出できるように RDV 法を改良することを目的としている。古くから細胞間で特異的に発現する遺伝子を検出する場合には、サブトラクション法が使われてきた。しかし、この方法は最終産物をプラスミドにクローニングして大腸菌を形質転換しなければならない。カルタヘナ条約により、未知のウイルス遺伝子のクローニング実験は文部科学大臣の承認を得たうえでレベル 3 の実験室でおこなわなければならない。しかし、本研究班のように緊急性を要するウイルスにおいては、できるだけ早いウイルス遺伝子塩基配列の決定が望まれる。一方、RDV 法は新しいダイレクトシーケンス法を特徴としているので、上記のような組み換え DNA 申請の必要がない。そこで、本研究ではサブトラクション法と RDV 法を組み合わせることにより、迅速かつ網羅的なウイルス検出方法の確立を試みた。

ウイルス分離された検体の場合であっても、未知のウイルスを検出することは容易ではない。このように、いつでも RDV 法による新規ウイルスの検出ができるように、常日頃から RDV 法を実施することは重要と考えられる。本研究ではウイルス分離後の検体から 2 つの新規ウイルスを検出したので、あわせて報告する。

B. 研究方法

RDV 法バージョン 3.1 : RDV 法は網羅的なウイルス検出法であるが、バージョン 1 では 6 万通りの PCR を実施しなければならなかった。そこで、アンプリコンを作成するためのアダプターや PCR のプライマーを工夫することによ

り、256 通りの PCR で網羅性を保てるように改良した方法をバージョン 3.1 として発表した (Watanabe et al., 2008)。さらに工夫を重ねることにより、64 通りでおこなうバージョン 3.1 を開発してきた (Watanabe et al., 2010) (図 1)。

新規ウイルスの検出 : 紀伊半島において、山口大学・農学部・前田健博士、田辺市・ふるさと自然公園センター・鈴木和男博士らのご協力を得て、野生のコウモリやイノシシを捕獲した。コウモリは脾臓からウイルス分離を試み、イノシシは血清からウイルス分離を試みた。コウモリの採取に関しては、東京大学・農学部・渡辺俊平博士、明石博臣博士らのご協力を得た。新規ウイルスの検出には、RDV バージョン 3.1 を用いた。

サブトラクション RDV 法の開発 : サブトラクションは和光純薬の DsDD cDNA Subtraction kit を用いた。ウイルス感染細胞および培養上清から抽出した DNA 群から非感染細胞および培養上清から抽出した DNA 群を差し引くために、RDV バージョン 3.1 を用いて DNA ライブラリーを作成した。その後、サブトラクションキットを用いて得られた DNA 断片を RDV 法のダイレクトシーケンス技術で解析した (図 2)。

(倫理面への配慮)

患者配列を特定する作業は行わず、個人情報に結びつく配列解析は一切行っていない。

C. 研究成果

1) RDV 法バージョン 3.1

図1に方法の概略を示した。cDNAライブラリーを作成する過程で、従来の平滑末端処理後にアダプターを付加する方法から、左右で異なる制限酵素で切断した後に別々のアダプターを結合させる方法に改良した。このことにより、制限酵素処理時に選択性が増し、最終PCR産物(2次ライブラリー)からダイレクトシーケンスをおこなうときのプライマーセットを64通りまで減少させることに成功した。操作性は従来のRDV法と変わりがなく、2日間で解析を完了できるが、検出感度はスタート時に10万分子の標的DNAを必要とするので、従来のRDV法より約10倍感度が悪いことがわかった。しかしながら、64通りのPCRだけで理論上網羅的解析を行うことができるので、このような技術の中で最も簡便な方法であると考えられる。

2) コウモリの新規ヘルペスウイルス

コウモリ (*Miniopterus fuliginosus*) の脾臓から初代培養細胞の樹立を試みた際に、3回目の継代で細胞障害が現れたので、培養上清からDNAを抽出し、ヘルペスウイルスを広く検出できるプライマーセットを用いてPCRをおこなったが、バンドは得られなかった(山口大学・前田健博士)。そこで、RDV法バージョン3.1を用いて解析したところ、Blast検索で *Tupaia herpesvirus* (GenBank No. AF281817) に相同性のある遺伝子が含まれていることが明らかになった。DNA genome walkingなどの手法を用いて、glycoprotein B領域の遺伝子配列を決定し、系統上を作成したところ、ベータヘルペスウイルス群に属すること

が明らかになったので、*Bat betaherpesvirus-2* と命名した(図3)。さらに同じ地域で疫学的調査をおこなったところ、50匹のコウモリ中4匹の脾臓由来DNAでPCR陽性であった。

この結果から、RDV法バージョン3.1は簡便ながらも新規ウイルスを検出できることが明らかになった。

3) イノシシの新規ラブドウイルス

イノシシの血清をヒトスジシマカ由来の培養細胞C6/36を用いてウイルス分離を試みたところ、接種後2週間培養した時点で細胞障害は認められなかったが、陰性対象に比べて細胞の増殖が遅くなっていた(山口大学・前田健博士)。この培養上清をサル腎臓由来のVero細胞に接種したところ、3日目に細胞が円形になるような細胞障害が観察された。培養上清からRNAを抽出して、日本脳炎ウイルスのRT-PCRを実施したが陰性であった。そこで、RDV法を用いてウイルス遺伝子の検出を試みたところ、ラブドウイルスに近縁であることが明らかになった。RDV法により得られた遺伝子断片(T13-3 fragment)をアミノ酸に変換して、代表的なラブドウイルスとのアライメントをおこなった(図4)。このウイルスを暫定的に、*Nishimuro rhabdovirus* と命名した。

4) サブトラクションRDV法

和光純薬のサブトラクションキットでは、左右に別々のアダプターが結合されたアンプリコンを作成する必要がある。それゆえ、RDV法バージョン3.1の2次ライブラリーをサブトラクションキットの出発材料として用いることが可能である。また、このキ

ットでは最終産物をクローニングしなければならぬので、RDV法のダイレクトシーケンスを組み合わせた(図2)。このサブトラクションRDV法が有用であるかを確かめるために、ウナギからRDV法により検出された新規DNAウイルス(ウイルスの検出については平成22年度文部科学省研究C基盤、代表・東海大学・小野信一博士で報告予定)を用いた。2つのDNAフラグメントがウイルスDNAの遺伝子配列と一致した。

D. 考察

本研究の目的は、ウイルス分離ができない、もしくは同定手段のないウイルスを迅速に検出する方法を確立することである。検体は血清や髄液、臓器、あるいはウイルス分離に成功したがウイルスを同定できない培養上清などが想定される。血清・髄液・培養上清については、従来のRDV法で対応可能である。しかし、臓器などの検体は、宿主遺伝子が大量に含まれるので遺伝子情報のないウイルス核酸だけを選択的に取得することは困難である。

本研究では、まず、ウイルス分離されたが同定のできなかつた培養上清から、RDV法を用いて2種類の新規ウイルスを検出することに成功した。コウモリから検出された新規ベータヘルペスウイルスと、イノシシから分離された新規ラプトウイルスである。両ウイルスとも外見からは健康な動物から分離されたので、宿主動物に対して病原性があるかについては明らかではない。しかし、種を越えて感染するときには病原性が顕れるウイルスも

あるので、本研究で検出されたこれらの新規ウイルスもその可能性を否定できない。今後、これらの新規ウイルスのPCRの検出系を確立しておくことは重要である。

RDVバージョン3.1は、ウイルスの網羅的検出法の中でもっとも簡便な方法である。RDV法では、目的に応じて様々なバージョンがあるが、バージョン3.1は今後、不明検体のスクリーニングに用いる予定である。本研究では、このバージョンはサブトラクションの出発材料として用いられた。バージョン3.1の検出感度が10万分子と低いので、サブトラクションRDV法の感度も高くないと考えられる。今後は、感度の改良が課題である。

E. 結論

本研究において、RDV法を用いて2種類の新規ウイルスの検出に成功した。また、ウイルス感染臓器から未知のウイルスを検出することを想定して、サブトラクションRDV法を開発した。

F. 健康危険情報

なし。

G. 研究発表

1. 論文発表

- 1) Shumpei Watanabe, Ken Maeda, Kazuo Suzuki, Naoya Ueda, Koichiro Iha, Satoshi Taniguchi, Hiroshi Shimoda, Kentaro Kato, Yasuhiro Yoshikawa, Shigeru Morikawa, Ichiro Kurane, Hiroomi Akashi, and Tetsuya Mizutani. Novel Betaherpesvirus in bats. *Emerg. Infect. Dis.* 16. 986-988. 2010.

2. 学会発表

- 1) 西條 政幸、福士 秀悦、水谷 哲也、緒方 もも子、倉根 一郎、森川 茂. 3分節 RNA の塩基配列に基づく中国新疆ウイグル自治区におけるクリミア・コンゴ出血熱ウイルスの分子疫学と進化. 第 58 回日本ウイルス学会、徳島、2010 年 11 月
- 2) 渡辺 俊平、Masangkay Joseph S、永田 典代、森川 茂、水谷 哲也、福士 秀悦、大松 勉、上田 直也、伊波 興一朗、谷口 怜、藤井 ひかる、津田 峻平、加藤 健太郎、遠矢 幸伸、久和 茂、吉川 泰弘、明石 博臣. フィリピンにおけるコウモリコロナウイルスの検出および飼育食果コウモリを用いたウイルス感染実験. 第 58 回日本ウイルス学会、徳島、2010 年 11 月
- 3) 岩田 奈織子、永田 典代、辻 隆裕、長谷川 秀樹、佐藤 由子、横田 恭子、宇田 晶彦、水谷 哲也、西條 政幸、森川 茂、佐多 徹太郎. SARS-CoV 感染動物モデルを用いた UV 不活化 SARS-CoV の副反応について. 第 58 回日本ウイルス学会、徳島、2010 年 11 月
- 4) 酒井宏治、田丸精治、前田健、永田典代、網 康至、岩田奈織子、鈴木忠樹、水谷哲也、福士秀悦、須崎百合子、緒方もも子、長谷川秀樹、西條政幸、山田靖子、倉根一郎、森川茂. カニクイザルで致死感染症を起こしたイヌジステンパーウイルスのサル及びイヌでの病原性の解析. 第 58 回日本ウイルス学会、徳島、2010 年 11 月
- 5) 伊波 興一朗、中内 美奈、谷口 怜、福士 秀悦、水谷 哲也、緒方 もも子、西條 政幸、久和 茂、倉根 一郎、森川 茂. アルゼンチン出血熱の実験室診断法の患者血清を用いた評価. 第 58 回日本ウイルス学会、徳島、2010 年 11 月
- 6) 水谷哲也、酒井宏治、本道栄一、福士秀悦、緒方もも子、西條政幸、倉根一郎、森川茂、前田健. コウモリから分離された新規アデノウイルスのゲノム配列の決定および系統学的解析. 第 58 回日本ウイルス学会、徳島、2010 年 11 月
- 7) 酒井宏治、永田典代、網 康至、岩田奈織子、鈴木忠樹、水谷哲也、福士秀悦、須崎百合子、緒方もも子、西條政幸、長谷川秀樹、山田靖子、倉根一郎、森川 茂. カニクイザルで致死感染症を起こしたイヌジステンパーウイルスの性状と実験感染サルでの病原性の解析. 第 150 回日本獣医学会 帯広 2010 年 9 月
- 8) 谷口怜、佐山勇輔、渡辺俊平、飯塚愛恵、福士秀悦、水谷哲也、石井寿幸、久和 茂、明石博臣、吉川泰弘、森川茂、倉根一郎. レストンエボラウイルス膜糖蛋白を標的とした抗体検出系の確立. 第 150 回日本獣医学会 帯広 2010 年 9 月
- 9) 水谷哲也、酒井宏治、本道栄一、福士秀悦、緒方もも子、西條政幸、倉根一郎、森川 茂、前田 健. コウモリから分離された新規アデノウイルスの分子学的性状決定. 第 150 回日本獣医学会 帯広 2010 年 9 月

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。

RDV ver 3.1

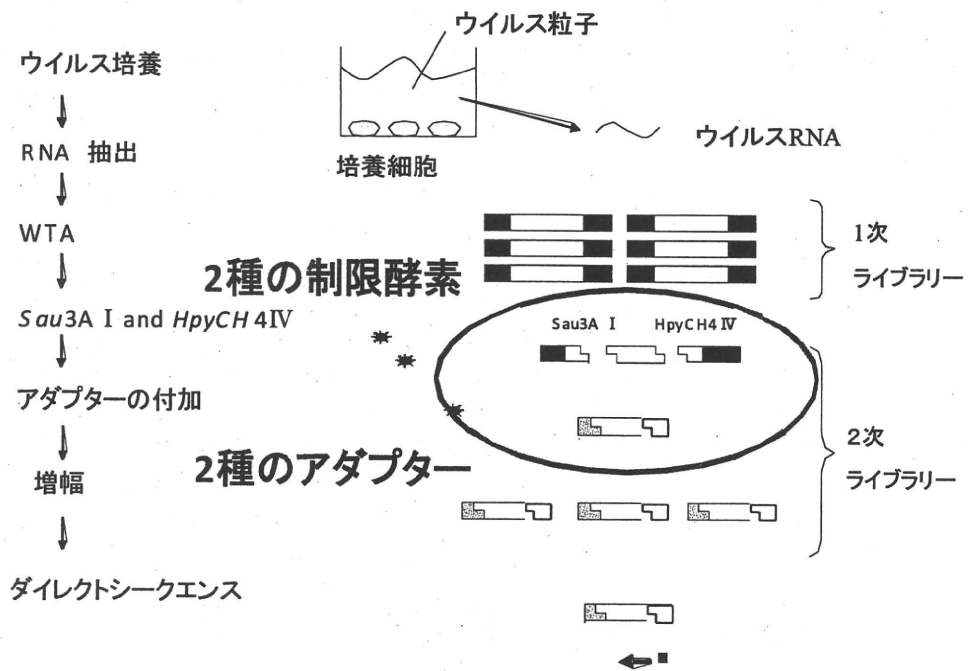


図1. RDV法バージョン3.1:

2種類の制限酵素 (Sau3AI と HpyCH4IV ; いずれも 4 塩基認識、粘着末端を形成) を用いることにより、2次ライブラリーを 64 通りの PCR で網羅性を保てるように改良した (Watanabe et al., 2010) 。 WTA: シグマ社の Whole Transcriptome Amplification kit.

サブトラクション RDV

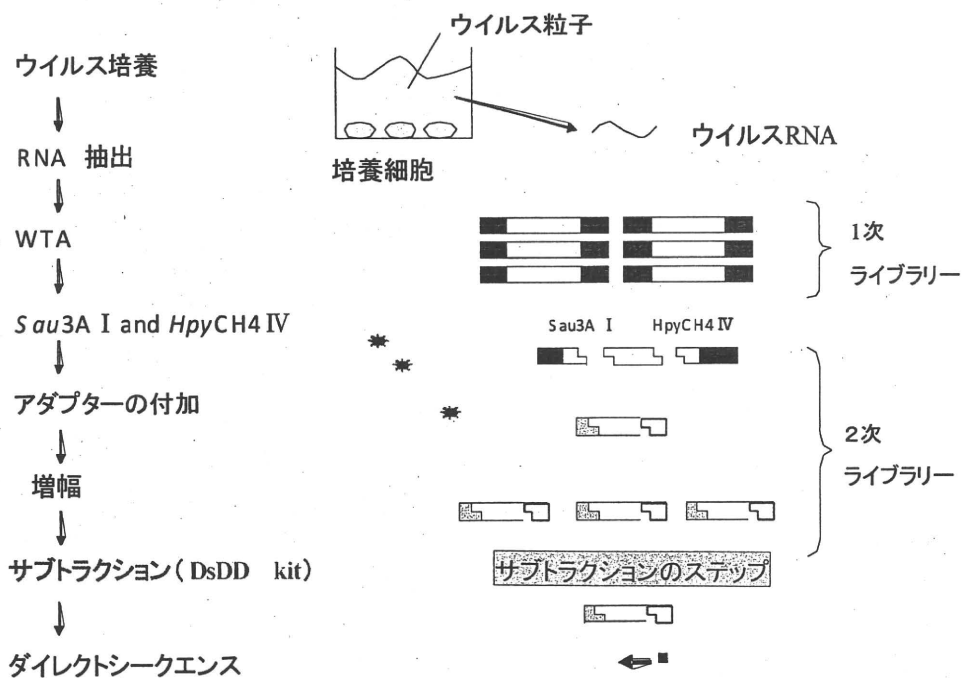


図2. サブトラクション RDV 法

RDV ver3.1 の工程において、2次ライブラリー作成後に和光純薬社のサブトラクションキットを用いる。その後、RDV 法のダイレクトシーケンスによりウイルスの遺伝子断片の塩基配列を決定する。

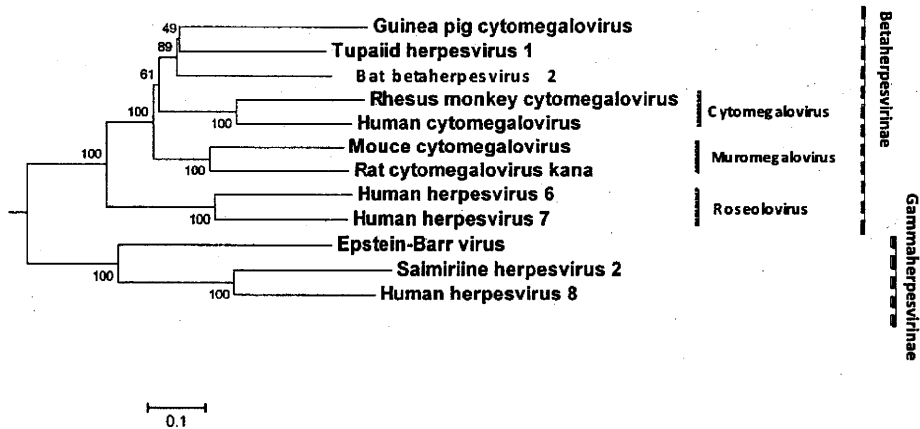


図3. 新規コウモリヘルペスウイルスの系統樹

詳細は Watanabe らの論文を参照のこと。

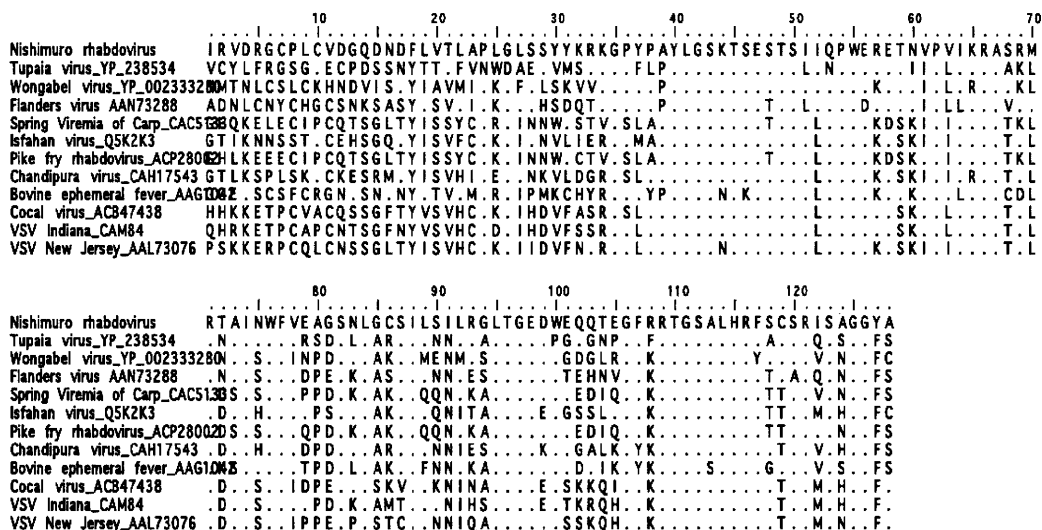


図4. 新規イノシラブドウイルス (Nishimuro rhabdovirus) のアミノ酸アライメント RDV 法で得られた遺伝子断片をアミノ酸に変換し、ラブドウイルスのアミノ酸とのアライメントをおこなった。本研究で得られた遺伝子領域はポリメラーゼ遺伝子の一部と考えられる。

現在、国内で分離・同定できないウイルス性出血熱等の診断等の
対応方法に関する研究

分担研究課題：新種、新興ウイルスが出現した場合の迅速なウイルス同定法の開発

研究分担者：遠藤 大二（酪農学園大学獣医学部放射線学教室）

研究要旨：近年、新種ウイルスの出現事例は多数報告されているが、ほとんどの事例がいずれかのウイルス科に分類されている。分担者らは、2009年度までに、ウイルス科内の不特定のウイルスを検出可能なPCR法を可能にするプライマー設計方法を開発してきた。本年度は、この設計プログラムを改善し、様々なウイルス科に適用するため、プログラムを全面的に改変することによりプライマーの自動設計を可能にするとともに、PCR増幅条件を改善することによる検出感度の向上を図った。その結果、データベースを設計用に調整することにより、当該プライマー設計プログラムを任意のウイルス科で実施可能にした。また、ホットスタート耐熱ポリメラーゼの利用とアニリング温度の調整により、 $10^0 \sim 10^1$ コピーの人工合成されたウイルスゲノム断片を鋳型についてそれぞれ30種中26種および18種のウイルスでPCR産物が確認された。本方法は新興ウイルスを含む多様なウイルス検出対策としての有効性が示唆された。

A. 研究目的

近年、病原性の新種ウイルスの出現は、年間1件以上報告されているが、アレナウイルスのLuJoウイルスに見られるように、ほとんどの事例ではすでに決定されているウイルス属の中間に出現することはあっても二種以上のウイルス科の中間に出現するような事例はほとんどない。また、プニヤウイルスの様にウイルス科内に多様なウイルス属を含む場合には特定のウイルス属内でウイルス属をまたぐことなく新種のウイルスが報告されている。このような結果から、ウイルス科内のウイルスを網羅できればほとんどの事例で新種・新興ウイルスの検出が可能であることが示唆された。その前提に基づき、分担者らは、アレナウイル

ス科およびハンタウイルス属を網羅的に検出可能なそれぞれのための3組のプライマーセットと設計プログラムを開発し、人工合成遺伝子を用いて実用の可能性を示した。

上述のウイルス科を網羅するプライマーとして、分担者らは、縮重プライマー(degenerateプライマー)を用いたが、degenerateプライマーは、構成するdegenerate塩基の多様性から、完全に自動的な設計は難しく、プログラムをコントロールできる人員が条件の判断や数値の設定を行う必要があった。一方、近年、一定期間におけるウイルス遺伝子の報告数は増加傾向にあり、今後は次世代シーケンサーなどの普及に伴いさらに増大することが予想される。増加する塩基配列データの中には新

規の病原性ウイルスの同定のために重要となるデータが含まれる可能性があり、これらを活用してその時点での最善の検査用プライマーを設計することが、新種・新興ウイルス検出には最も効果的になることが予想される。これらのことから、2009年度までに開発されたプログラムを基に、より機動的にかつデータ量に関わらず自動的な最適プライマーの設計を可能にするプログラムの修正が必要であると結論した。また、前年度の方法では検出感度が低かったため、反応条件を検討して検出感度を高めることも目標とした。

これらのプログラムの改善と反応条件の検討を通じ、新種ウイルスを高い可能性で検出可能な PCR プライマーの設計・提示方法を開発することを、本研究の目的とする。

B. 研究方法

1. 塩基配列の記載方法

塩基配列および degenerate 配列は、Nomenclature Committee of the International Union of Biochemistry (NC-IUB) による Nomenclature for Incompletely Specified Bases in Nucleic Acid Sequences, Recommendations 1984 (<http://www.chem.qmul.ac.uk/iubmb/misc/naseq.html>) に従った。

計算中、degeneracy が莫大な数となる場合があったため、degeneracy の指標として degeneracy の 2 を底とする対数を算出し、degeneracy 指数として用いた。

2. プライマー設計プログラム

1) 開発環境

プライマー設計プログラム (Co-ordination of common motifs: CoCoMo

アルゴリズム)を稼働するためのオペレーティングシステムとしては、Linux (Ubuntu10.04 または 10.10)を用いた。

開発言語としては、Ruby 1.87 を、同言語のバイオインフォマティクス用の拡張モジュールとしては BioRuby を、データベースとの接続モジュールとしては MySQL/Ruby を使用した。データを格納・処理するためのデータベースには MySQL 5.1 (<http://dev.mysql.com/>)を使用した。

2) 連携プログラム

塩基配列間の相同性を算出するためには、NCBI の開発した Blast2 プログラム (<ftp://ftp.ncbi.nih.gov/blast/>)を使用した。複数のウイルスの塩基配列を整列させるためには、MAFFT プログラム (MAFFT version 6, <http://align.bmr.kyushu-u.ac.jp/mafft/software/>)を用いた。多数の配列を相同性を指標としてグループ化するためのプログラムとしてはクラスタリングソフト MCL プログラム (<http://www.micans.org/mcl/>) を利用した。

3) 対象ウイルスデータ

各ウイルスの既知のゲノム塩基配列データは、GenBank (<http://www.ncbi.nlm.nih.gov/>)よりダウンロードし、BioRuby を用いたプログラムにより MySQL のテーブルに格納した。

4) CoCoMo アルゴリズムの構成

i) CoCoMo アルゴリズム

CoCoMo としては、前年度報告したプログラムを改変したアルゴリズムを作出した。

塩基配列のアライメントを元に degenerate プライマーを作出した場合、1kb の鋳型について、1,000 本以上の degenerate プライマーが予測されることが多い。一般的に、degenerate プライマーの数が増えると、相互作用を想定する必要から、それらの評価のための計算時間は級数的に増大する。2009 年度までに分担者が作出した方法では、計算を軽量にするために、プライマー候補の除去を繰り返した。また、計算の爆発的増大を防

ぐために、少数のウイルス配列から予測された共通 degenerate プライマーについて、多数のウイルス配列との相同性を調べ、相同性が一定以上のプライマーを適用可能プライマーとして記録する方法も検討された。

候補プライマーの除去を繰り返す方法も少数のグループのプライマーから普遍的プライマーを再選択する方法のいずれについても、計算途中での候補の微妙な調整がなければ、プライマー候補が計算途中で枯渇し計算が停止することが頻発した。一方、計算停止を防ぐために候補プライマー数を増大させると、反対に計算が爆発的に増大し、計算が全く停止しなくなり結果を得ることができなくなった。

そこで、報告者は、degenerate 塩基を含む塩基配列を 4 桁の二進数に変換し、相同性や degeneracy などのプライマー評価に必要な数値を算出するプログラムを作成することにより計算全体を高速化した。さらに、前年度までの方法のうち、計算停止の原因となった箇所について検討を進めた結果、計算の基本となるウイルス塩基配列のセットの調整方法を見直し、プライマー作出の基本単位をウイルス遺伝子とし、適度な相同性距離にあるウイルス遺伝子集団をプライマー算出の基本データとして再構成することと、計算上もっとも時間を要したアライメントをアミノ酸配列で実施することにより、プライマー予測プログラムの途中停止を殆どの場合防ぐことに成功した。これらの試みに基づいて下記の手順が構成された(図 1)。

- (1) ウイルス遺伝子の収集と均質データベースの作成(図 2)。
- (2) ウイルス遺伝子産物アミノ配列の相同性の算出。
- (3) ウイルス遺伝子産物のクラスター化(図 2)。
- (4) クラスター内ウイルス遺伝子産物アミノ配列のアライメント(図 3)。
- (5) ウイルス遺伝子産物アミノ配列のアライメントを元にしたクラスター内塩基配列アライメントの作成(図 3)。
- (6) クラスター内共通塩基配列の作成と

degeneracy の算出(図 4)。

- (7) degeneracy 条件を満たすプライマー候補 1 の選択(図 4)。
- (8) プライマー候補塩基配列の末端 degeneracy を元にしたプライマー候補 1 の方向の決定(図 4)。
- (9) プライマー候補 1 の仮想 PCR(図 5)。
- (10) 仮想 PCR で 100-500 塩基の PCR 産物を生じることが予想されるプライマー候補 1 の選択(プライマー候補 2、図 5)。
- (11) プライマー候補 2 についてリボゾーム RNA へのハイブリダイズの評価して点数化
- (12) プライマー候補 2 についてヘアピン構造の可能性を評価して点数化
- (13) プライマー候補 2 についてセルフダイマーの可能性を評価して点数化
- (14) プライマー候補 2 について仮想 PCR でヘテロダイマーを形成する可能性を評価して点数化
- (15) (11)-(14)の評価点を合計してプライマー候補 2 の総合評価を算定。上位のプライマーをプライマー候補 3 とする。

5) 対象ウイルス遺伝子のクラスター表示

ウイルス遺伝子間の相同性全体を図示化するためには、Pajek(<http://pajek.imfm.si/doku.php>)を用いた。

6) プライマー検証のためのウイルスゲノム断片の合成

プライマー検証用 DNA 断片は、GenScript 社(<http://www.genscript.com/>)に依頼した。

7) degenerate プライマーによる PCR

プライマー検証用 DNA 断片を鋳型とした PCR は、HotStartGoTaq Green PCR Mix (プロメガ)によって実施された。反応条件としては、低アニーリング温度と平均的アニーリング温度の二段階増幅条件を用いた。

すなわち、2×GoTaqMix (Mg₂+ plus) 12.5 μl、200 pmol の各 degenerate プライマー、プライマー検証用 DNA 断片に滅菌

蒸留水を 25 μ l となるように混合し、反応溶液とした。反応は 95°C 2 min の後、[94°C 30 sec — 46°C 30 sec — 72°C 30 sec] の反応を 15 回おこない、続けて、[94°C 30 sec — 55°C 30 sec — 72°C 30 sec] の反応を 25 回行った。最後に 72°C 5 min の反応を一回行い 8°C で保温した。

8) マイクロチップゲル電気泳動

メーカーの手順書に従い、PCR 反応後の反応溶液を機器にセットし、分析した。本報告では、100~400 塩基の PCR 産物が分析対象となったため、分離緩衝液として DNA-500 Separation buffer (島津) を、マーカーとして DNA-500 marker for MultiNA (島津) を用いた。

C. 研究結果

本年度は、改善された設計プログラムが提示したプライマー配列での PCR 検出感度を向上させるため、短い断片上で共通性の高いプライマーを設計し、試験用の DNA 断片を用いて PCR 条件を改善した。そのためのプライマー設計の対象として、前年度にクラスターを報告したアレナウイルス科ウイルスの L セグメントの polymerase 遺伝子の中央部約 200 塩基を選択した。また、ブニヤウイルス科ハンタウイルス属の M セグメント上の糖タンパクプリカーサー (glycoprotein precursor: GPC) 遺伝子の中央部および 3' 側の保存領域を選択した。続けて、本年度改善された PCR 設計プログラムの稼働を確認するため、アレナウイルス Polymerase の 1kbp 断片またはハンタウイルスの S-segment ORF 全長を対象としてプライマーを用いて設計の妥当性を検討した。

1) 短い共通領域での degenerate PCR 条件の検討

(1) アレナウイルス

アレナウイルス科ウイルスの L セグメントの polymerase 遺伝子の中央部約 200 塩基に注目した。鋳型配列の対象としてこの領域周辺のアミノ酸配列を元に平準化された 19 種のアレナウイルスの polymerase 遺伝子断片を用いた(表 1)。ク

ラスターを探索する領域を徐々に変えながら、比較的多数のウイルスが含まれたクラスターを形成した 2 つの領域についてプライマーが設計された。それぞれの領域について 10 種のアレナウイルスの遺伝子データ(表 1)から塩基配列を収集し、PCR 試験に用いる人工遺伝子断片を合成した。これらの人工遺伝子をコピー数 $10^9 \sim 10^2$ となるように希釈し、アレナウイルスの網羅的検出を目的として設計された degenerate プライマー(表 2)により PCR を実施した(図 6)。その結果、 10^4 、 10^2 または 10^1 コピーの感度でそれぞれ 20 通りの組み合わせのうち、6、4 および 10 通りで PCR が確認された。

(2) ハンタウイルス

CoCoMo アルゴリズムによるプライマー設計にあたっては、M フラグメントの保存領域の相同性を種間で比較し、多くの種を含む二つのクラスターを設定した。クラスターの設定にあたっては、PCR 検出の感度を検討する実験を計画したため、PCR 産物を 200 塩基程度の領域に絞る(表 3)、網羅性を優先する指標としてプライマーを設計した(表 4)。

続いて、PCR 試験に用いる人工遺伝子断片を合成し、コピー数 $10^9 \sim 10^0$ となるように希釈し、ハンタウイルス属共通検出用プライマーにより PCR を実施した(図 7)。その結果、 10^1 または 10^0 コピーの感度でそれぞれ 10 通りの組み合わせのうち、2 および 8 通りで PCR が確認された。

(3) 1kbase レベルの設計対象配列に関する degenerate プライマー設計

アレナウイルスの塩基配列データは、遺伝子名称が種によって異なる場合があるため、annotation のみでは機械的に遺伝子を特定することが難しかった。そのため、設計に先立ち、ウイルス蛋白質全体について相同性や内包性を健闘し、データの冗長性を除いた(図 8)。設計プログラムは L-SegmentORF 全体を対象として実施したが設計プログラムが停止することはなかった。さらにそのデータベースを対象としてクラスターを作出し、クラスターごとにクラスター内の遺伝子を代

表するウイルス種について L-Segment の 3300~4000 塩基の 700 塩基程度の断片を合成した(表 5)。さらに、設計されたプライマー中その断片を増幅する可能なプライマーを設計した(表 6, PCR 成功例のみのセットを示す)。L-Segment 全体では 400 本、1196 組のプライマーセットが予想され、そのうち 34 組が合成断片上にフォワード・リバース双方のプライマーがハイブリダイズすることが予想された。本年度はそれらのうち 10 組について試験を実施し 4 組で PCR 反応産物が観察された。特に二組については、今回合成した鋳型についてはこれら二組のうちいずれかで増幅されることが示唆された(図 9)。

D. 考察

多数の遺伝子から共通する塩基配列を検出する作業は、計算技術の向上とデータベースの高度化から、2000 年ごろには日常的に実施可能なことであると予想されていた可能性が高い。遺伝子を数学的に取り扱う研究では、人為的に選択されてきた相同遺伝子のアライメント全体を俯瞰するという人為的操作の対極として、遺伝子の塩基配列を 10 塩基程度のパターンの集合としてあつかい、共通配列の選択を集合の論理演算として取り扱う試みが行われていた。しかしながら、遺伝子の配列パターンは元集合のサイズがある程度以上になると級数的に増大するため、「単純に」全てのパターンを網羅しつつ共通かつ最適な塩基配列パターンを選択することは、難しく、さらに公共データベース上の既知の配列の増大により、このような共通配列と最適プライマーの選択は、現実的にはいかなるコンピュータでも不可能なことが示されつつある。

一方、近年、データと可能性が莫大である問題について、最善とは証明できなくても、きわめてそれに近い解答を作出できるアルゴリズムという計算手法が実績を上げている。データベースの高性能化と低価格化にともない、数学的に厳密な証明が無い場合でも、大量のデータで、アルゴリズムから予測された方法の的確さを帰納的に証明可能になってきたこと

が背景にある。データベースの普及と分析は、同時に、熟練した職人の技術の確かさを証明することにもつながった。結果として、経験からの確かな勘得た技術者の技術の喪失を防ぐために、彼らの考え方、視点およびカンをアルゴリズムやロジックとして抽出する作業が多くの企業で進められている。この仕組みの確立により、たとえば従来専門家のみが可能にしていたバスの配車が、コンピュータにより可能になり、さらに専門家が扱える最大台数を越えて最適の配車が行えるため配送効率や安全を向上させることが可能になっている。

本研究では、ウイルス検出を多数行ってきたウイルス専門家の研究班にバイオインフォーマティクス技術を持つ分担者が寄与することにより、人為的に行われてきたウイルスプライマーの選定作業を、システムティックに明示的な論理とコンピュータ上で実現可能な計算に置き換えることにより、増大する遺伝子とウイルス研究の知見をウイルス検出に最大限生かすための試みである。近年の「経験的な手法に基づいたプログラム」手法の成功から、このようなアプローチが、新種ウイルス検出のための効果的な方法を提供することが示唆される。

分担者らは、これまでマニュアルに行われてきたプライマー設計の多くの部分をコンピュータ内の計算に行うことを通じて、従来よりも網羅性の高いプライマーの設計プログラムを作出してきた。コンピュータプログラムでは、専門家が目で感じることができる「保存領域の大体の場所」を特定することが難しいため、その作業の代役として可能な小プログラムを連結させていくという方法で改善を進めてきた。前年度は、ウイルス科を 3 組のプライマーセットで網羅可能な程度まで設計技術を向上させることができた。

しかしながら、検出感度は網羅性の向上にともない退行し、昨年度設計したプライマーは 10^7 コピーのウイルス遺伝子断片を必要とした。本年度は、基本的な方向を変更することなく、さらにヒトの経験に近づくための改善を進めた。具体的には、基本データの平準化、遺伝子ホ

モロジーに基づくクラスターの自動化、アミノ酸配列を利用した塩基配列アライメントの作成と共通塩基と degeneracy の算出にもとづくプライマーの予測およびプライマーと判定を可能な限り相対評価に基づくことである。これらの改変により、従来しばしばおきていたプログラムの計算バーストや候補配列の消失にともなう停止の頻度が極めて低くなった。見直された全ての過程において、個々のプログラムが改善されていることと計算上負荷の少ない指標が選択されたことにより、計算は適切に分割され、無理で巨大な計算を行う危険性を回避することができるようになった。本研究で最初に適用された 200 塩基程度の小さな断片でも今年のプログラムは相対的に優れたプライマーセットを算出した。続けて行われたウイルスのセグメントレベルでの分析においても、プライマー計算途中のプログラム停止は起きなかった。これらの結果から、今年度のプログラムはプライマー設計を改善していく上で、ウイルスの PCR 検出を多数経験してきた専門家の知恵を導入するのに見合ったプログラムが構築できたことが示唆される。

実際の PCR においては、検出感度が求められるが、CoCoMo アルゴリズムでは 2009 年度はかなり低い検出感度を示した。本年度は、degenerate プライマーでしばしば使われている PCR プロトコルを使うことにより、検出感度を $10^3 \sim 10^0$ という実用レベルに引き上げた。このような基盤の上に、現在単純な加算と相対評価として用いている塩基配列上の指標をプライマー設計に生かすことが必要となる。分担者は、今年度において多様な条件のプライマーと鋳型を調整し、最初の試みをスタートした。プライマーと鋳型配列の一致率、GC 比、プライマー末端の T_m 値など多様な指標について、PCR 増幅性との相関を調べ、利用可能な指標をプライマー設計に生かしていくことが今後求められる。このような多様な指標についての経験の積み重ねを数値化し、再利用可能にしていくための手法として Bayesian Classifier が知られているため、

来年度においてはこの指標を活用し、設計されたプライマーの PCR 反応を起こす確率を高める。

E. 健康危険情報

該当なし。

F. 研究発表

1. 論文発表

遠藤大二「修復遺伝子をクローニングするためのプライマー設計」放射線生物研究 45 巻 2 号 170-182

2. 学会発表

遠藤大二 「多様性の高い複数の塩基配列について degenerate プライマーを設計する CoCoMo ウェブサーバ」 日本獣医学会学術集会、2010 年 3 月、東京

G. 知的財産権の出願・登録状況

1) オープンソースとして公表予定のプログラムコード：省略

2) インターネット上で利用可能なプログラム限定公開サイト(2011 年度および 2012 年度においては研究班内での利用を予定): <http://cocomo.ver2.asia>

図1 ウイルス科を網羅するためのプライマー設計方法の概要

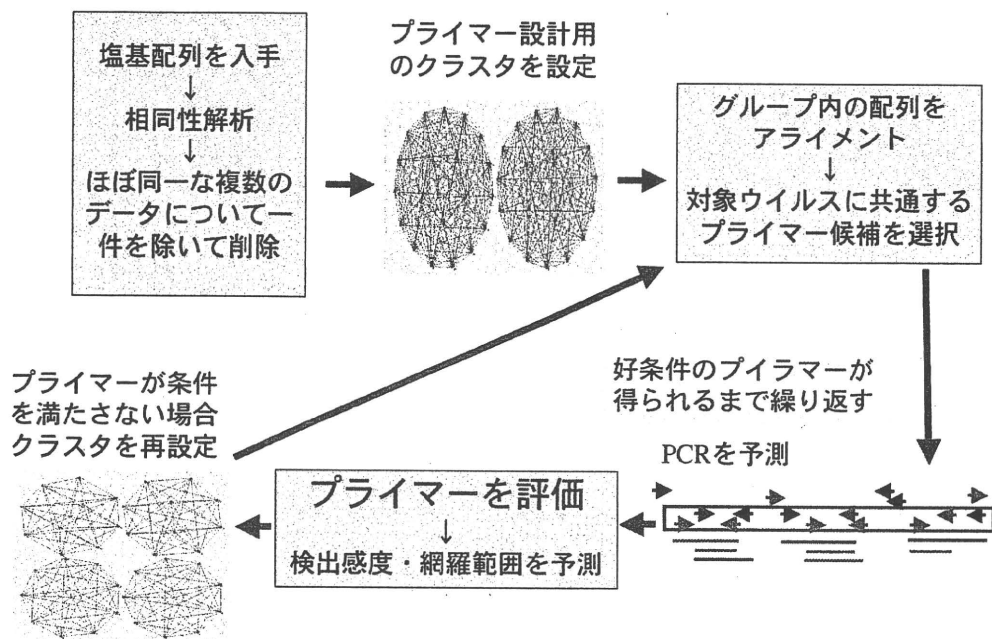
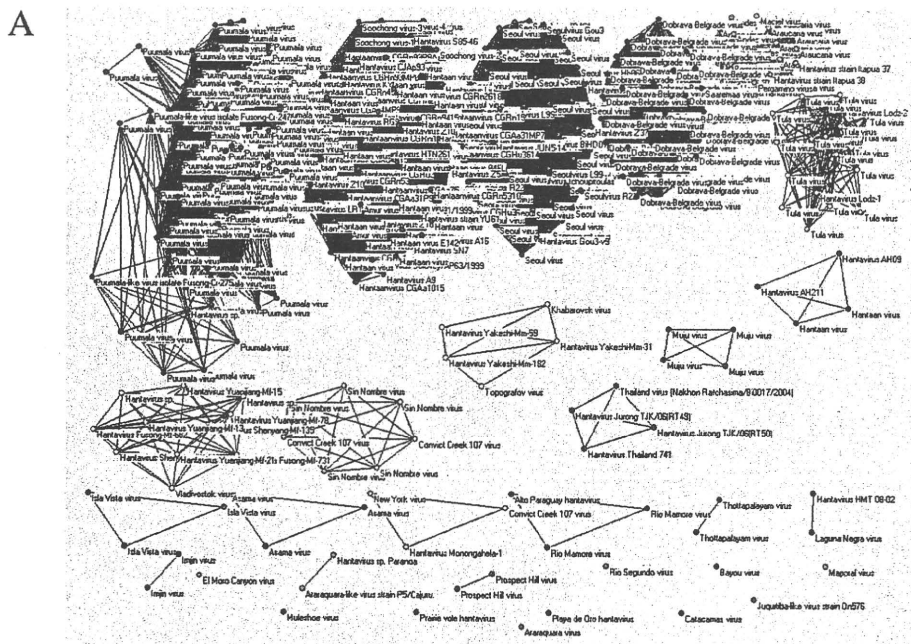
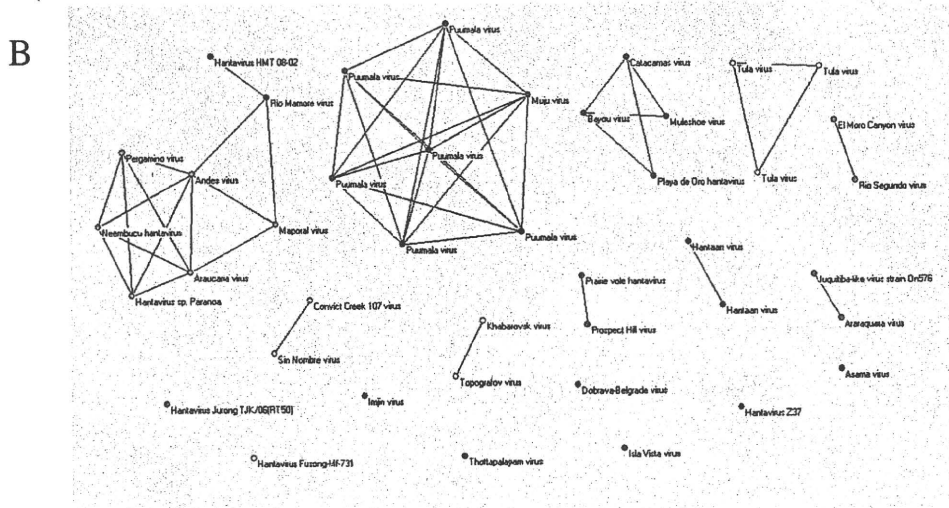


図2 遺伝子データの平準化



(説明は図1B下部に記載)

図2 遺伝子データの平準化



ハンタウイルス属ウイルスのSegment S オープンリーディングフレーム(ORF)翻訳産物の相同性に基づいたデータの平準化。赤い点でデータベース上の一件のデータを、青い線で相同性を示す。A: 2011年1月31日にGenBankに掲載されているハンタウイルス属ウイルスのSegment S ORF全データについてアミノ酸配列の相同性80%以上の場合遺伝子間に線をひいて作成されたクラスタを示す。B: Aのデータについてアミノ酸配列が95%以上相同な遺伝子データは同一とみなして代表の一遺伝子のみ示した図。線は80%以上の相同性を示す。AおよびBの図中で線分での接続が無く島状に独立している点の集合はクラスタを示している。相同性の算出にはBLAST2を、表示にはPajekを使用した。

図3 クラスタ内塩基配列のアライメント作成

A株 ATGGCAACTATGGAAGAAATCCAGAGAGAAATC
 B株 ATGGCATCTATGCAAGAAATCCAAAGAGAAATC
 C株 ATGCCAATGGCCGAAATCCCGAGAGAACTC


A株 MATMERIOREI
 B株 MASMOEIOREI
 C株 M-MAEIPREL

A株 ATGGCAACTATGGAAGAAATCCAGAGAGAAATC
 B株 ATGGCATCTATGCAAGAAATCCAAAGAGAAATC
 C株 ATG---CCAATGGCCGAAATCCCGAGAGAACTC

クラスタ内の遺伝子データから、塩基配列のアライメントを作成する手順を示す。遺伝子の塩基配列(上段)から、翻訳産物のアミノ酸配列にデータを転換し、そのアミノ酸配列についてMAFFTプログラムを用いてアライメントを作成した(中段)。さらにアミノ酸配列のアライメントから、各アミノ酸の元になるコドンを再構成することにより塩基配列のアライメントを作成した(下段)。赤枠と矢印で塩基配列アライメント作成方法の事例を示した。塩基配列アライメント作成の際には、アミノ酸配列アライメントでのギャップは3塩基分のギャップとなった。

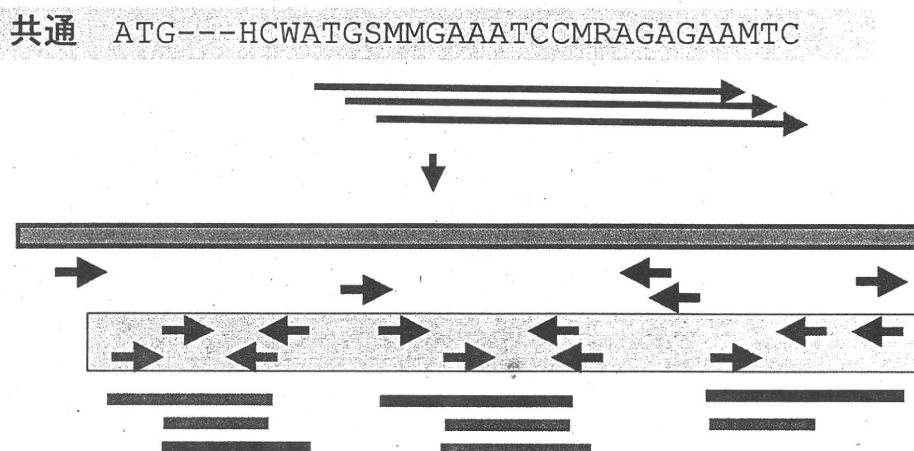
図4 塩基配列アライメント上でのプライマー候補の選択

A株	ATGGCAACTATGGAAGAAATCCAGAGAGAAATC
B株	ATGGCATCTATGCAAGAAATCCAAAGAGAAATC
C株	ATG---CCAATGGCCGAAATCCCGAGAGAACTC
共通配列	ATG---HCWATGSMMGAAATCCMRAGAGAAMTC
Degeneracy	111---312111222111111122111111211



塩基配列のアライメントからプライマー候補1を選択する手順。図3下段の塩基配列アライメントから、塩基配列アライメント上で共通配列をdegenerate配列として作成された(中段)。プライマー候補は共通配列から以下の条件を満たす配列として選択される。条件:①ギャップを含まない、②いずれかの末端6塩基にはdegenerate塩基を1個のみ含む、③残り15塩基のdegeneracyの積算値が2048以下、④全長21塩基。

図5 アライメントに基づくプライマー候補2の選択



アライメントの共通配列から選択されたプライマー候補1(上段)の方向と遺伝子上の位置に基づき、仮想PCRを実施し、100~500塩基対のPCR産物を生じるプライマー候補1のみをプライマー候補2とした(下段)。

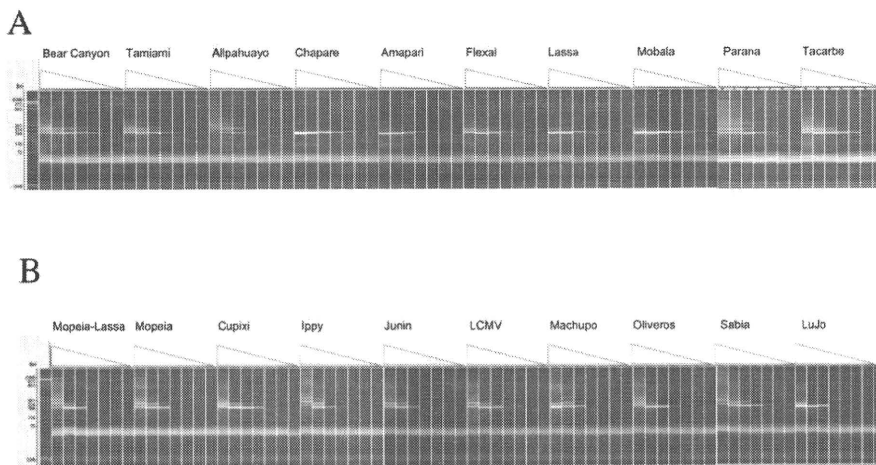
表1 アレナウイルス用プライマーの検証対象としたウイルス種およびセグメントLの塩基配列Accession番号

Accession	種名	略称	分類
NC_010249	Allpahuayo virus	ALLV	New world arenaviruses
NC_010251	Amapari virus	AMAV	New world arenaviruses
NC_010255	Bear Canyon virus	BCNV	New world arenaviruses
NC_010759	Flexal virus	FLEV	New world arenaviruses
DQ489718	Junin virus	JUNV	New world arenaviruses
NC_012777	LuJo virus	LuJo	Unclassified arenaviruses
NC_005079	Machupo virus	MACV	New world arenaviruses
NC_010250	Oliveros virus	OLVV	New world arenaviruses
NC_010761	Parana virus	PARV	New world arenaviruses
NC_006313	Sabia virus	SABV	New world arenaviruses
NC_010702	Tamiami virus	TAMV	New world arenaviruses
NC_004292	Tacaribe virus	TCRV	New world arenaviruses
NC_010563	Chapare virus	CHAPV	unclassified New world arenaviruses
NC_007906	Ippy virus	IPPYV	Old world arenaviruses
NC_004297	Lassa virus	LASV	Old world arenaviruses
DQ868488	Lymphocytic choriomeningitis virus	LCMV	Old world arenaviruses
NC_007904	Mobala virus	MOBV	Old world arenaviruses
NC_006572	Mopeia Lassa reassortant 29	MOPLAS	Old world arenaviruses
DQ328875	Mopeia virus	MOPV	Old world arenaviruses

表2 アレナウイルス科L-Segment内の共通性が高い領域を増幅するプライマーセット

クラスター名称	Forward primer	Reverse primer	L-segment内の領域	平均的増幅産物
L4	GSCCCCA YTTKGAAT G	AAGAACARGTWGGWA GYAATAGRGA	3529-3745	216
L5	TCATCWCCCARAC AWARAA YCT	CATWYACYTCTAGTGAT GAYCAAAT	2944-3159	195

図6 アレナウイルスdegenerateプライマーによる増幅試験



アレナウイルスL-segment, polymerase遺伝子中央部の保存領域で設計されたdegenerateプライマーによるPCR産物のマイクロチップ電気泳動像。鑄型としては、各レーンの上にウイルス種名を示し、鑄型としては人工合成された各ウイルスの該当領域のDNA断片を用いた。三角形の底辺から 10^7 , 10^6 , 10^4 , 10^3 , 10^2 , 10^1 および 10^0 コピーの鑄型が反応に用いられていることを示す。PCR反応にはHotStartGoTaqGreen Mixtureを用いた。プライマー濃度および温度条件は本文に示す。左端のレーンにマイクロチップ電気泳動装置が算出したマーカーの位置を示す。

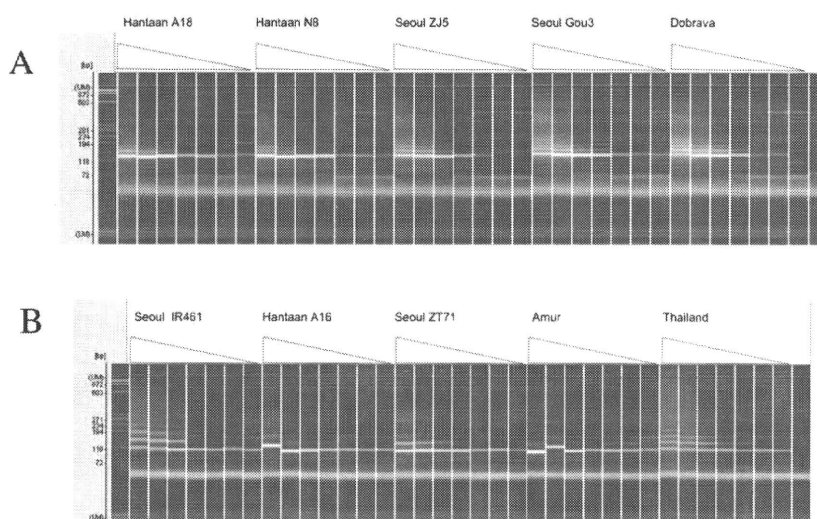
表3 ハンタウイルス用プライマーの検証対象としたウイルス種

Accession	種名	Strain
GQ120966	Hantaan virus	A18
EF077656	Hantaan virus	N8
FJ811839	Seoul virus	ZJ5
AB027521	Seoul virus	Gou3
GU904035	Dobrava-Belgrade virus	
AF458104	Seoul virus	IR461
AF288645	Hantaan virus	A16
EF117248	Seoul virus	ZT71
EF371454	Amur virus	JilinAP06
TIDG1G2A	Thailand virus	

表4 ハンタウイルス属M-Segment内の共通性が高い領域を増幅するプライマーセット

クラスター名称	Forward primer	Reverse primer	L-segment内の領域	平均的増幅産物
M2	ARYGCAARGAYCC WGAYGGGDAT	CCWACMCCWGARC CCCA	2859-3020	141
M3	WAARGGBTCTYAT GGTTTGYGA	GGYTCACARTGKGTR AARCA	1657-1776	119

図7 ハンタウイルスdegenerateプライマーによる増幅試験



ハンタウイルスM-segment, glycoprotein遺伝子中央部の保存領域で設計されたdegenerateプライマーによるPCR産物のマイクロチップ電気泳動像。鑄型としては、各レーンの上にウイルス種名を示し、鑄型としては人工合成された各ウイルスの該当領域のDNA断片を用いた。三角形の底辺から 10^7 , 10^5 , 10^4 , 10^3 , 10^2 , 10^1 および 10^0 コピーの鑄型が反応に用いられていることを示す。PCR反応にはHotStartGoTaqGreen Mixtureを用いた。プライマー濃度および温度条件は本文に示す。左端のレーンにマイクロチップ電気泳動装置が算出したマーカーの位置を示す。