

表 4. 人工合成オリゴヌクレオシド配列(続き)

5:E_X75657	ACTATCAAGGTATGTTGCCCGTTTGTCTCTAATTCAGGATCATC AACCACCAGTACGGGACCCCTGCCGAACCTGCACGACTCTTGCTCA AGGAACCTCTATGTTCCCTCATGTTGCTGTTCAAACCTTCGGAC GGAAATTGCACTTGATTCCCATCCCATCATCATGGGCTTCGGAA AATTCCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTTT ACTAGTGCCATTTGTTTCAGTGGTTCGCCGGGCTTTCC
6:F_X69798	ACTATCAAGGTATGTTGCCCGTTTGTCTCTACTTCCAGGATCCAC GACCACCAGCACGGGACCCCTGCAAAACCTGCACAACCTTGCACA AGGAACCTCTATGTTCCCTCCTGTTGCTGTTCAAACCTTCGGAC GGAAACTGCACTTGATTCCCATCCCATCATCCTGGGCTTTAGGAA AATACCTATGGGAGTGGGCCTCAGCCCGTTTCTCCTGGCTCAGTT TACTAGTGCAATTTGTTTCAGTGGTTCGCGTGGGCTTTCC
7:G_AF160501	ACTATCAAGGTATGTTGCCCGTTTGTCTCTGATTCCAGGATCCTC GACCACCAGTACGGGACCCCTGCAAAACCTGCACGACTCCTGCTCA AGGCAACTCTATGTATCCCTCATGTTGCTGTACAAAACCTTCGGAC GGAAATTGCACTTGATTCCCATCCCATCATCTGGGCTTTTCGCAA AATACCTATGGGAGTGGGCCTCAGTCCGTTTCTCTGGCTCAGTTT ACTAGTGCCATTTGTTTCAGTGGTTCGTAGGGCTTTCC
8:H_AY090454	ACTATCAAGGTATGTTGCCCGTGTCTCTACTTCCAGGATCTAC AACCACCAGCACGGGACCCCTGCAAAACCTGCACCACTCTTGCTCA AGGAACCTCTATGTTTCCCTCCTGTGCTGTACAAAACCTTCGGAC GGAAATTGCACTTGATTCCCATCCCATCATCTGGGCTTTTCGGAA AATACCTATGGGAGTGGGCCTCAGCCCGTTTCTCTGGCTCAGTTT ACTAGTGCAATTTGTTTCAGTGGTTCGCTAGGGCTTTCC

PvB19核酸合成領域

Genotype_Accession No.	5'-	Sequence	-3'
1: FJ591158		GCAGCAGTGGTGGTAAAAGCTCTGAAGAAGTCAAGTAAAAGCAGC TTTTTAAACCTCATCACCCAGGCGCCTGGAACACTGAAACCCCG CGCTCTAGTACGCCCATCCCCGGGACCAGTTCAGGAGAATCATTT GT	
2: GM703964		GCAGCAGTGGTGGTAAAAGCTCTGAAGAAGTCAAGTAAAAGCAGC TTTTTCAACCTCATCACTCCAGGCGCCTGGAACAGTAAAACCCCG CGCTCTAGTACGCCCGTCCCCGGGACCAGTTCAGGAGAATCATTT GT	
3: GM703962		GCAGCAGTGGTGGTAAAAGCTCTGAAGAAGTCAAGTAAAAGCAGC TTTTTCAACCTCATCACTCCCGGCGCCTGGAACAGTAAAACCCCG CGCTCTAGTACGCCCGTCCCCGGGACCAGTTCAGGAGAATCATTT GT	
4: GM703832		GCAGCAGTGGTGGTAAAAGCTCTGAAGAAGTCAAGTAAAAGCAGC TTTTTCAACCTCATCACCCAGGCGCCTGGAACACTGAAACCCCG CGCTCTAGTACGCCAGTCCCCGGGACCAGTTCAGGAGAATCATTT GT	
5: DQ333427		GCAGCAGTGGTGGTAAAAGCTCTGAAGAAGTCAAGTAAAAGCAGC TTTTTCAACCTCATCACTCCAGGCGCCTGGAACAGTAAAACCCCG CGCTCTAGTACGCCCGTCCCCGGGACCAGTTCAGGAGAATCATTT GT	
6: DQ357065		GCAGCAGTGGTGGTAAAAGCTCTGAACAAGTCAAGTAAAAGCAGC TTTTTAAACCTCATCAACCCAGGCGCCTGGAACACTGAGACCCCGC GCTCTAGTACGCCCATCCCCGGGACCAGTTCAGGAGAATCATTTG G	

WNV核酸合成領域

Genotype_Accession No.	5'-	Sequence	-3'
1: FJ483549		GACAGCGGCTGTGGTATGGCATGGAGATTAGACCACAGAGACAT GATGAAAAGACCCCTTGTGACGTACAAGTGAATGCTTACAACCGT GATATGATTGATCCTTTTCAGCTGGGCCTTCTGGTCTGTCTTGG CCACCCAGGAGGTCCTTCG	
2: NC_001563		GAGAATGGCTGTGGTATGGAATGGAATTCGACCTACGCGGCAC GACGAAAAGACCCCTCGTGAATCGAGAGTGAATGCATACACGCC GACATGATTGATCCTTTTCAGTTGGGCCTTATGGTCTGTCTTGG CCACCCAGGAGGTCCTTCG	

表 4. 人工合成オリゴヌクレオシド配列 (続き)

3. FJ425721_D17763	GAGAATGGCTGTTGGTATGGAATGGAAATTCGACCTACGCGGCAC GACGAAAAGACCCTCGTGCAATCGAGAGTGAATGCATACAACGCC GACATGATTGATCCTTTTCAGCTGGGCCTTCTGGTCGTGTTCTTGG CCACCCAGGAGGTCCTTCG
4. FJ159131	GAATCGGGTTGTTGGTATGGGATGGAGATACGCCACTTCGCCAC GATGAGAAAACCCCTGGTGCAATCAAAGGTTACAGCTTACAATGCT GACATGATTGATCCTTTTCAGCTGGGCCTTCTGGTAGTGTTCCTTGG CCACCCAGGAGGTCCTTCG
5. FJ159130	GAATCGGGTTGTTGGTATGGGATGGAGATACGCCACTTCGCCAC GATGAGAAAACCCCTGGTGCAATCAAAGGTTACAGCTTACAATGCT GACATGATTGATCCTTTTCAGCTGGGCCTTCTGGTAGTGTTCCTTGG CCACCCAGGAGGTCCTTCG
6. DQ256376	AGAAGTGGATGTTGGTATGGCATGGAGATCAGGCCACAGAAGCAT GATGAAAAGAACTCTCGTTCAGTCACAGGTGAATGCCTTACAATGCT GACATGATTGACCCTTTTCAGTTGGGCCTTCTGGTCGTGTTCTTGG CCACCCAGGAGGTCCTTCG
7. DQ176636	GAGAGCGGATGTTGGTATGGAATGGAAATTAGACCGCTAAGGCAT GATGAGAAAACCTCTTGTGCAATCGAGAGTGAATGCGTACAATGCT GACATGATTGATCCTTTTCAGTTGGGCCTTCTGGTTGTGTTCTTGG CCACCCAGGAGGTCCTTCG
8. AY765264	GATAGTGGATGTTGGTATGGAATGGAAATCAGACCTTTGAAGCAT GATGAGAAAGACGTTGGTTCAATCTAGGGTGAAGCGCCTACAAATCT GATATGATTGATCCTTTTCAGCTGGGCCTTCTGGTAGTGTTCCTTGG CCACCCAGGAGGTCCTTCG
9. AY688948	GAGAATGGCTGTTGGTATGGAATGGAAATTCGGCCCCTGCGGCAT GACGAAAAGACCCTCGTGCAAGTCGAGAGTAAATGCATACAACGCC GACATGATTGATCCTTTTCAGTTGGGCCTTCTGATCGTGTTCCTTGG CCACCCAGGAGGTCCTTCG
10. AY603654	GATAGCGGTTGTTGGTATGGTATGGAGATTAGACCACAGAGACAT GATGAAAAGACCCTTGTGCAAGTCACAAGTGAATGCCTTACAACGCT GACATGATTGATCCTTTTCAGTTGGGCCTTCTGGTCGTGTTCTTGG TCACCCAGGAGGTCCTTCG

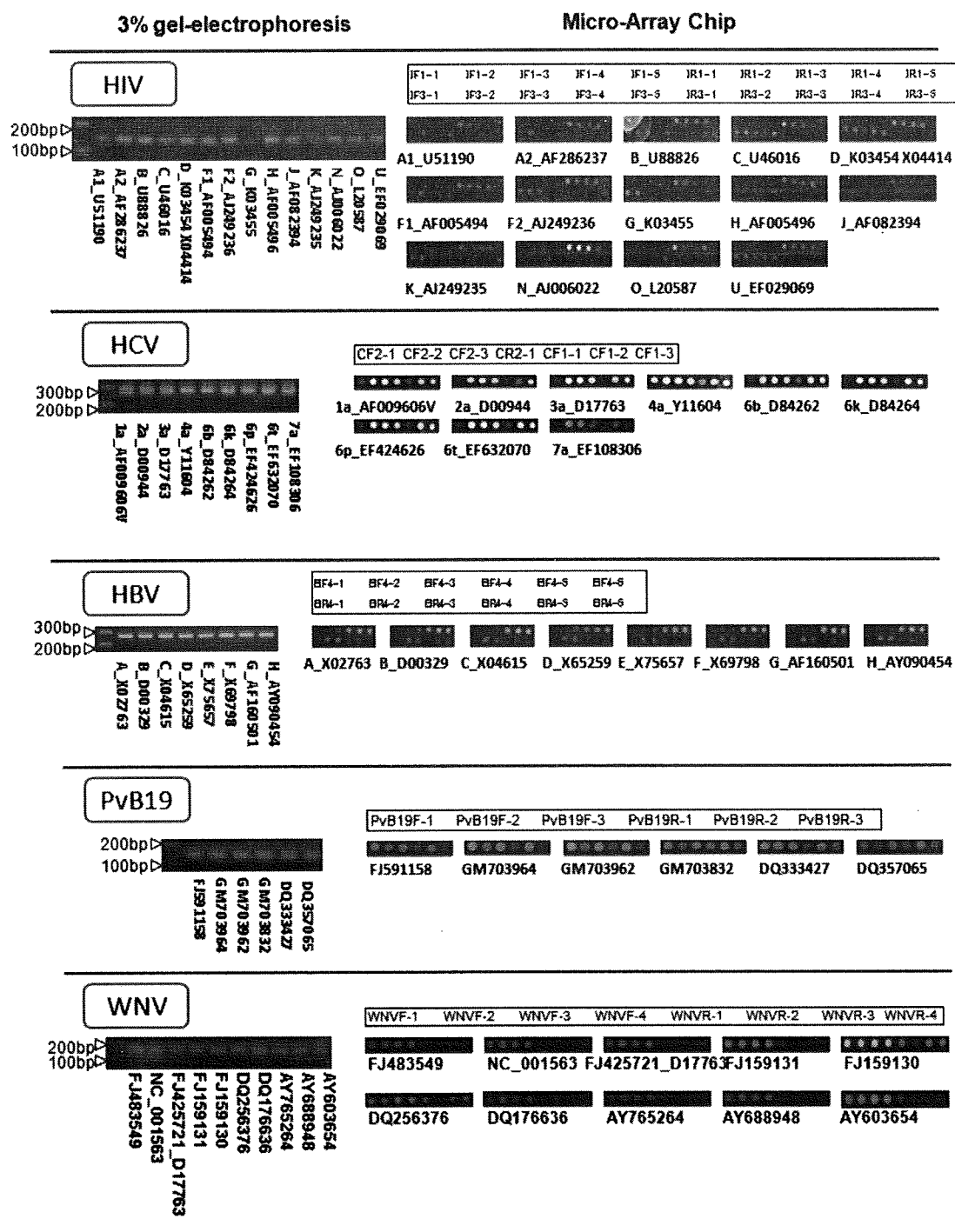


図1.人工合成オリゴヌクレオシドを用いた変異株に対するバリデーション
HIV、HCV、HBV、PvB19、WNV の各ウイルス変異株の遺伝子配列を基に作製した合成オリゴヌクレオチドを鋳型として用い、Digenerate プライマーと検出用プローブの変異株に対する有効性を評価した。左：Digenerate プライマーを用いて PCR を行い、3%-アガロースゲル電気泳動により増幅産物を確認した。右：PCR により Cy5 標識した増幅産物を DNA Chip 上に滴下し、シグナルの検出を行った。

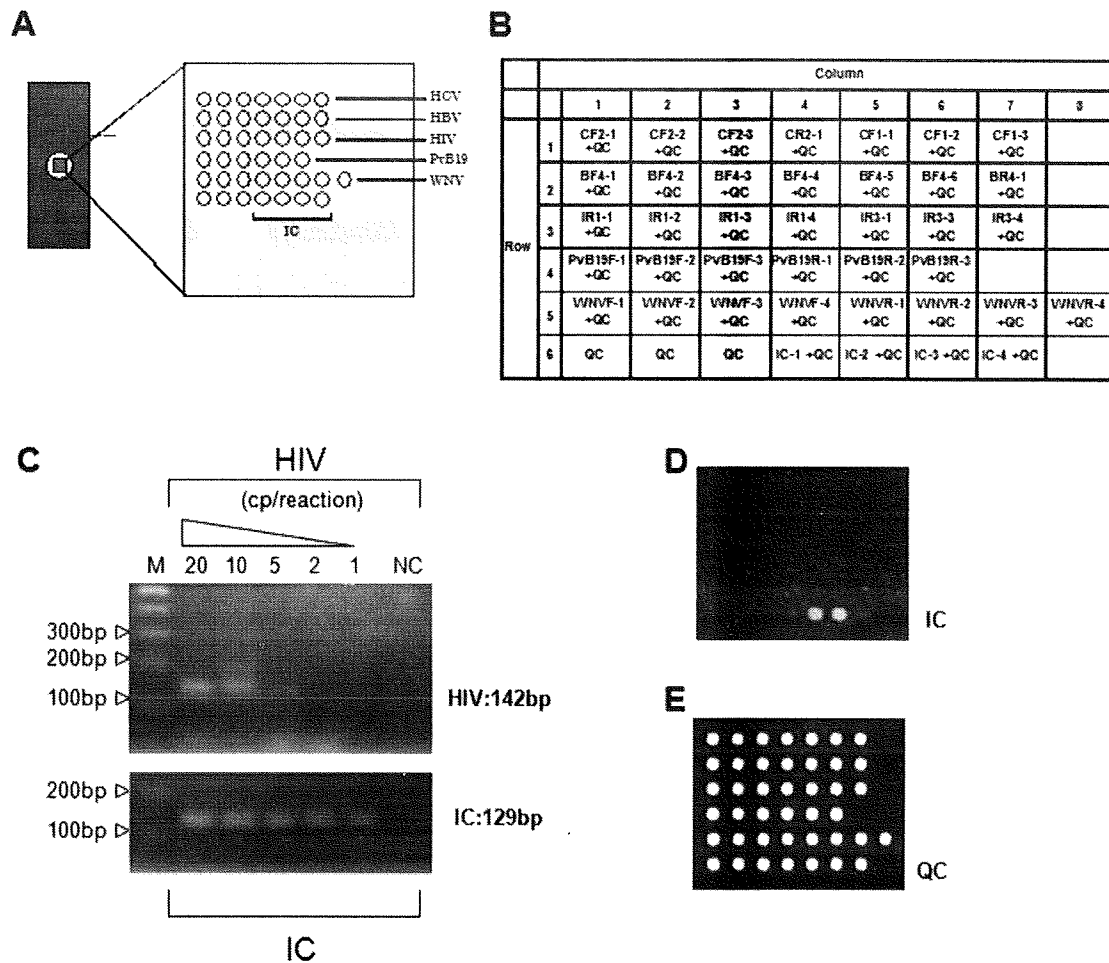


図 2. スポット配置とプローブ ID、及び精度管理用コントロール

A: スライドガラス上に固定されたウイルス核酸検出用プローブの配置と、B: プローブ ID。
 3mm×3mm 角のチップに約 100 個のスポットを置くことができる。C: Internal Control (IC) の使用例として段階希釈された HIV のウイルス核酸を鋳型に検出用 Digenerate プライマーと IC 用プライマーを用いて PCR を行い、3%-アガロースゲル電気泳動により増幅産物を確認した。
 D: Internal Control 用スポットの確認。IC 用プライマーを用いて PCR により Cy5 標識した増幅産物を DNA Chip 上に滴下し、シグナルの検出を行った。E: Quality Control (QC) 用プローブの確認。Cy3 で標識された合成オリゴヌクレオチドを用いてスポットの品質管理を行った。

HIV		HCV		HBV		PvB19		WNV	
cp*		IU*		cp*		IU*		cp*	
20		10		10		10		20	
10		5		5		5		10	
5		2		2		2		5	
2		1		1		1		2	

*Copies or International units /reaction

図3. マイクロアレイによるウイルス核酸の検出

標準検体から調整したウイルス核酸の希釈系列を用いて各ウイルスの検出用プローブの検出限界をマイクロアレイにより測定した。Digenerate プライマーを用いた PCR によって Cy5 標識されたウイルス核酸を DNA Chip 上に滴下し 50°C、2 時間ハイブリダイズを行った。

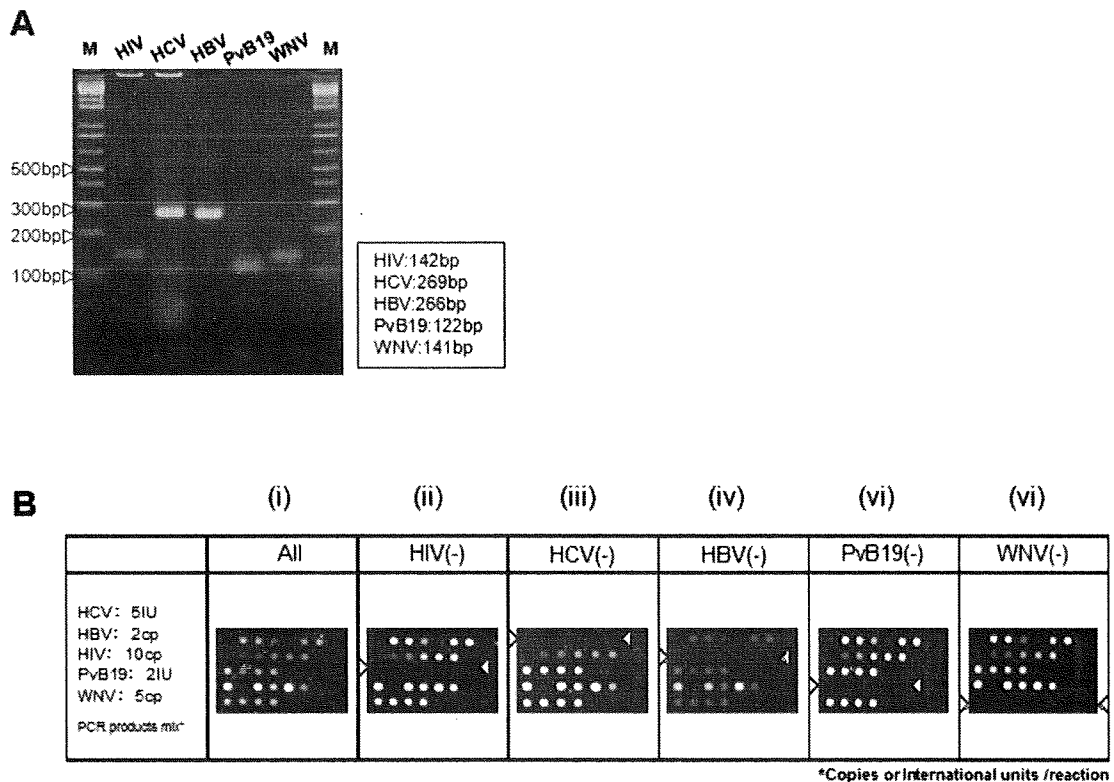


図 4.複数ウイルス核酸の同時検出

各ウイルス検出用プローブがクロスハイブリすることなく複数のウイルス検出することができるかを試験した。A:500IU/cp の各ウイルス核酸を鋳型として特異的 Digenerate プライマーによる PCR を行った後、3%-アガロースゲル電気泳動により増幅産物を確認した。B: PCR により増幅された各ウイルス核酸を混ぜ合わせ、DNA Chip 上に滴下し 50℃、2 時間ハイブリダイズを行った。(i) 5 種全てのウイルスを混合、(ii) HIV を除く 4 種 (iii) HCV を除く 4 種 (iv) HBV を除く 4 種 (v) PvB19 を除く 4 種 (vi) WNV を除く 4 種。

厚生労働省科学研究費補助金（肝炎等克服緊急研究事業）
ウイルス肝炎感染防止体制の確立に関する総合研究
分担研究報告書(平成 21 年度)

ウイルス種・株間共通プライマーを用いた信頼性の高いウイルス
検出方法の開発

研究協力者 遠藤 大二 酪農学園大学獣医学部放射線学教室

研究要旨

本研究は、輸血血液の安全維持に用いる、多数のウイルス種および株を同時に高感度・高信頼性で検出する方法の開発を目的とする。報告者は平成 20 年度において、ウイルスを共通して検出するための degenerate プライマー設計プログラム (Co-ordination of common motifs; CoCoMo)を開発した。当該設計プログラムは、従来にない degenerate プライマー設計能力を示したが、共通プライマーを選択するためのウイルス塩基配列数に限界があるため、多数のウイルス種用の共通プライマーを設計するためには改善が必要であった。本年度においては、新興再興ウイルス感染症に対応することを目的として、CoCoMo プログラムによる設計プライマーの関連ウイルスへの適用を高速に検索するプログラム群を開発した。実施例として、ヒトヘルペスウイルスおよびヒトアデノウイルスを選択し、開発プログラム群により、複数種を共通して増幅するプライマーを設計した。共通プライマーは、人工的に合成されたウイルス遺伝子断片を増幅したことから、degenerate プライマー設計と適用性の高速探索を連携した方法は、多ウイルス種・多ウイルス株の同時増幅プライマー設計に有効であることが示唆された。

A. 研究目的

1. 研究背景

輸血の血液安全については、血液を介して感染する恐れのあるウイルスの混入に対するコントロールが最大の課題となっている。日赤では、平成 11 年から安全対策として B 型肝炎ウイルス (HBV)、C 型肝炎ウイルス (HCV)、ヒト免疫不全ウイルスタイプ 1 (HIV1) に対するウイルス核酸増幅検査 (NAT) を導入し、ウイルス検査の感度を向上させる対策を実施している。しかし、NAT 検査のウイルス検査の感度限界は十分といえず、推計上は年間 13 件の HBV が検査をすり抜けて感染すると予測されている。同

様に HCV や HIV1 では NAT 導入後の 10 年間にそれぞれ 4 例、1 例のすり抜けによる感染例が既に報告されている。

これらの調査推計に加え、ウイルス学的研究からも検査感度を向上させる必要性が示唆されている。すなわち、HBV、HCV および HIV1 はいずれも高頻度で突然変異を繰り返しており、現在使用されている遺伝子検査用のプライマーが結合する塩基配列に変異が生じないという保証はできない。

また、これらの主要な対象とされるウイルスに加え、E 型肝炎ウイルス (HEV)、ヒトパルボウイルス B19 などの感染事例が報告され、感

染検査の対象を拡大する必要性が示唆されている。さらに、地球温暖化やグローバル化により、1997 のニパウイルス脳炎、1999 ウェストナイル熱／脳炎、2001 E型肝炎、2003 重症急性呼吸器症候群(SARS)、2003 高感染性鳥インフルエンザ、そして2008年から2009年には豚由来の新型インフルエンザなど、多様なウイルスの出現が報告されており、国内にも侵入の事例や可能性があり、輸血安全のための対策の必要性が示されている。

これらの状況を勘案すると、NATが高感度ではあるといっても、そのみに依存することは適切とは言えず、検出対象を拡大しつつ高い検出感度を持つ検査方法の開発が必要である。また、輸血血液検査は検体数が膨大であるため、新たに開発される検査のコストは制限される必要がある。検査の開発体制については、安全体制にイニシアティブを持つという点で、国外企業の開発にのみ依存することなく、国内の研究組織も独自に開発を進めることが望ましい。このような社会的要請に基づき、本研究では、多様なウイルスに対応しつつ高感度であるウイルス検出方法の開発に取り組んで来た。

検出範囲を広げるための方法として、すべての塩基と水素結合をするイノシン酸に加えて、二重鎖を形成する際に強い水素結合を形成する人工塩基が開発されているが、それらは非特異的結合を促進するため、プライマーによるPCR検出の特異性を低下させる。従って、検出範囲を広げるためのプライマーを設計するためには、ウイルス検出のための特異性を維持しつつ複数株または複数種のウイルスへ適用範囲を広げるプライマー設計を実施する必要性が示された。

特異性を維持しつつ、適用範囲を広げるための方法としては、複数の塩基配列を整列させて共通の塩基配列をそのままプライマーとして採用し、ウイルス株間で変化している塩基については、対象配列間での複数の塩基を混合塩基一文字で示した塩基(degenerate 塩基)をプ

ライマーとして設定する方法が取られる。

たとえば、二種の鋳型の塩基配列が

ATGGGCAGGGA

ATGGCCAGGGA

であった場合、5番目の位置の塩基をGとCの混合塩基として

ATGG{CG}AGGGA

という配列のプライマーが使用される。このようなプライマーは degenerate 塩基を利用することから、degenerate プライマーと呼ばれる。HIV1、HCV および HBV など、多様な塩基置換を生じているウイルスを鋳型として PCR 法を実施する場合には、この degenerate プライマーを使用することにより、特異性を失わずに多くの変異株を同時に検出することが可能となることが知られていた。

複数のウイルスに対する degenerate プライマーを設計する場合、既知の塩基配列をデータベース上から取得し、配列を整列(アライメント)させ、株間での塩基置換または変異が少ない領域を特定し、その領域内にプライマーを設計する。このようなプライマー設計作業は、プライマーが鋳型に結合する際の degenerate 塩基の位置の影響や、degenerate 塩基を最小にとどめるための配列の選択など、複雑な因子が多数関係する。そのため、単純なプログラムでの解決が難しいため、基本的には研究者の目視に依存して、プライマー配列が決定されてきた。また、実際には、多くの領域が候補に挙がるため、多数の候補配列から実験的にプライマーを絞り込む作業が行われることが通例であった。多くの事例では、該当するウイルスの塩基配列の報告数が限られていたため、目視での設計で十分であった。

しかしながら、既知の配列の増加に伴い、さらに高度な設計技術が必要とされるようになってきた。1990年代以来、プライマー設計を自動化する試みが報告されている。

自動化されたプライマーの設計においては、鋳型塩基間で共通する配列を最大にし、

degenerate 塩基の使用範囲を最低にすることが、多数のウイルスに対応しながら検出感度を高めるためには必要とされる。このようなプライマー設計を単純な計算を用いた場合は、途中に必要とされる塩基配列は膨大となり、どのような高速コンピュータを用いても 100 年を越すような膨大な計算時間を要する場合もあることが判明している。数学的には、このような膨大な時間を要する計算について、計算時間を短縮して近似解を得るために、アルゴリズムと呼ばれる計算手法が知られている。最適な degenerate プライマーの設計についてもアルゴリズムが取り入れられたため、プライマーの自動設計は設計アルゴリズムと呼ばれている。

既知のプライマー設計アルゴリズムでは、プライマーのうち 3' の短い配列を設定した後に、残りの 5' 領域を設計する方法が、効率的な共通 degenerate プライマーに用いられていた。本研究でも平成 19 年度に 3' 末端の短い共通配列を選択することによるプライマー設計方法を開発した。当該プライマー設計方法は、対象ウイルス間に共通する短い塩基配列 motif を調整することによりプライマーを設計するため、CoCoMo アルゴリズム (Co-ordination of common motifs) と称することとした。

CoCoMo アルゴリズムでは、相同性が比較的低い対象ウイルス群でのプライマー設計効率を向上させるため、検索する motif に、共通性を無視する塩基をギャップとして挿入し、共通 motif 選択に当たっては、塩基配列間での共通塩基配列を整列させる既知のプログラムを導入し、さらにプライマーの評価モジュールを作成してプログラムに組み込んだ。結果として、前年度よりも広い範囲のウイルスに共通するプライマーを算出することが可能となった。

設計されたプライマーの実用性を検討するためには、増幅対象となる塩基配列を持つ DNA を用いた PCR 反応が実施された。本研究で対象とするウイルスは多岐にわたっており、既

知のウイルス全てを鋳型として検証することは難しかったため、連結オリゴマー PCR (OE-PCR) 法を用いてプライマーの鋳型となるウイルスゲノム領域を合成方法を導入した。

前年度に制作されたプログラムは、HBV、HCV および HIV1 では効率的なプライマーの設計能力を示した。しかしながら、ヒトアデノウイルスやヒトヘルペスウイルス 5~7 のように、単一種内に多様性が存在する上、検出対象が複数のウイルス種にまたがる場合には、プライマーの設計範囲を越えるという現象が予備的に見られた。また、CoCoMo アルゴリズムに導入する鋳型塩基配列数が 20 本を越える場合には、必要となる計算時間が 10 日間を越え、コンピュータが計算上の暴走に入らないようにコントロールすることに多大な計算資源を割かれる問題が見られるようになった。

平成 21 年度においては、CoCoMo アルゴリズムでのプライマー設計を最大に生かすために、プライマー設計過程を 2 つのステップに分けた。第 1 のステップでは CoCoMo アルゴリズムを適用するための、最適なウイルス塩基配列のグループを多数のウイルス塩基配列から選択した。この選択過程においては、前年度以上に多くのウイルス配列を対象として最適なグループを設定する必要があったため、多様な対象について高速でグループ化を実施する既知のプログラムを導入した。

第 2 のステップでは、最適化によって得られた多数のプライマー候補について、設計対象以外の広いウイルス塩基配列への適用可能性を高速に検索した。ここでは、degenerate 塩基配列と一般の配列間での相同性を極めて高速に算出することが必要となったが、既存の方法では計算時間がかかりすぎるため、独自の方法を開発した。

このように、プライマー設計を 2 つのステップに分割することにより、前年度に比較して広い範囲の対象ウイルス株・種について共通プライマーが設計可能であることが示された。

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/na_seq.html)に従った。

塩基の縮重性 (degeneracy) は表 1 に従い、積算縮重値 S_d は塩基配列内の全ての degenerate nucleotide の積算縮重値 (degeneracy) を積算した。例えば、ATRYBNCC の場合、degeneracy は $2 \times 2 \times 3 \times 4 = 48$ とした。計算中、degeneracy が莫大な数となる場合があったため、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

CoCoMo アルゴリズムを実施するためのコンピュータとしては、64bit CPU (Athlon64) および 2~4G バイトのメモリーを搭載した DOS/V 仕様パーソナルコンピュータを用いた。プログラ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 (Ubuntu9.04 または 9.10) 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。鋳型上の 5 塩基の motif の出現頻度を累計するためには、平成 19 年に報告者が開発した gas ソフトを用いた。この gas プログラムをコンパイルおよび稼働するために、gcc C コンパイラを使用した。

元遺伝子データ、処理過程の一時データ、予

測プライマー塩基配列、位置情報および方向等の付随情報を格納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 mcl (<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

1) CoCoMo アルゴリズム本体

CoCoMo アルゴリズムでは degenerate プライマーを設計するために下記の順で処理を実施した。

① ウイルスゲノム配列のアライメント (図 1A)

テーブルに格納されたデータを、MAFFT プログラムによりアライメントした。アライメント上では、相同性を持つ配列が再構成された配列上で同じ位置にくるように配置され、相同位置までの塩基配列が不足する場合には、“—” の記号 (ギャップ) が挿入され、位置の調整が行われた。結果として、“—” の記号を加えて 5' 端からカウントしたアライメント上での塩基の位置は共通配列上の位置を示すように調整された。

② 共通ギャップ motif の検索 (図 1B)

探索対象とする motif は、前年度は連続する 6~7 塩基としたが、このように規定して、旧世界アレナウイルス共通 motif を探索した場合、対象ウイルスに共通して存在し、かつアライメント上での共通位置に局在する 6 塩基 motif は非常に少なく、それらの motif を起点として設計されたプライマーはほとんど見出すことができなかつた。そのため、アライメント上の同一位置に存在しながら多数のウイルスに共通して存在する motif として、一塩基を検索対象から外したギャップモチーフを用いることとした。このモチーフは ATNGC のように 4 塩基は motif としてウイルスゲノム上での位置を探索する対象とするが、中央の 1 塩基は N に示すようにどのような塩基でも探索上許容することとした。すなわち、ATNGC というギャップモチーフを探索した場合、ウイルスゲノム上の ATCGC, ATAGC, ATGGC および ATTGC の配列が探索対象となる。

このギャップ motif について、全ての対象ウイルスゲノム上での位置を特定しその位置を比較した。すべての対象ウイルス上に存在し、アライメント上での位置が 20 塩基以内であるようなギャップ motif を選択し、共通ギャップ motif とした。

③ degenerate flanking 配列の算出(図 1C)

前ステップにおいて抽出された共通ギャップ motif について、その 5' 側の 16 塩基を各ウイルスゲノムから抽出し、対象ウイルスで共通する degenerate 配列と degeneracy を算出した。このステップで degeneracy が 1,024 を超える場合には、当該ギャップ motif はプライマー予測対象から外した。

このステップで degeneracy が 1,024 以下と予測された flanking 配列とギャップ motif の連結した配列をプライマー候補 1 とした。

④ プライマー候補 1 の仮想 PCR(図 1D)

個々のプライマー候補 1 については、ギャップ記号が挿入されているアライメント上での位置が記録されているが、元の対象配列では

その位置が記録されていない。そのため、全てのプライマー候補 1 について元配列上の位置と方向を記録した。

続いて、位置の記録を基に、各プライマー候補 1 の 3' 方向 100~600 塩基の範囲について探索プライマーと反対方向のプライマー候補 1 が存在した場合、二本のプライマー候補を PCR 産物を生成するセットとして記録した。PCR プライマーセットの記録にあたって、セットの候補に使われたプライマー候補 1 を別に保存し、使用されなかつたプライマー候補は候補から棄却した。結果として、プライマー候補 1 の一部のみが次段階のプライマーとしてリストに残った。

⑤ PCR に寄与すると予測されたプライマー候補 1 の伸長(図 1E)

PCR に寄与すると予測されたプライマー候補 1 について、再度標的ゲノム上での 5' 隣接配列を 5 塩基取得してプライマー候補に連結し、プライマーの全長を 26 塩基とした。このことにより、プライマー候補はいずれも 63°C 以上のアニーリング温度を持つことが算出され、後にプライマーのアニーリング温度を 63°C に統一する作業が可能となった。この段階のプライマーをプライマー候補 2 とした。

⑥ 忌避配列を含むプライマー候補 2 の棄却(図 1F)

CCCCC などの同一塩基の連続または CTCT などのヒトゲノム上に多数存在するリピート配列を含むプライマー候補 2 は、PCR に実際に使用した場合に、非特異的 PCR 産物を生成する可能性が高いため、プライマー候補から棄却した。

⑦ プライマー候補 2 による仮想 PCR(図 1G) 対象ウイルスゲノム上のプライマー候補 2 の位置は、プライマー候補 1 の時点で算出されているが、この位置を基に、④で示した仮想 PCR を実施した。④と同なじく、仮想 PCR の対象となったプライマー候補のみをリストに残し、他を候補から棄却した。

⑧ プライマー候補 2 の塩基長の調整(図 1H)

プライマー候補 2 について、最近接塩基対法 (Nearest Neighbor method) に基づいてアニーリング温度を算出した (http://www.genosys.jp/whatsnew/tm/tm_syosail.html)。プライマー候補 2 が degenerate プライマーである場合は、可能なプライマー塩基配列全てについてアニーリング温度を算出し、平均値を当該プライマーのアニーリング温度とした。アニーリング温度が 64°C を越える場合、各プライマーの 5' 端の 1 塩基を除いた配列について、アニーリング温度を再計算した。5' 端からの塩基の除去は、アニーリング温度が 64°C となるまで繰り返された。結果として、長さは異なるがアニーリング温度が同一のプライマーセット候補がリストに残り、これをプライマー候補 3 とした。

⑨ プライマー候補 3 の評価 (図 1I)

プライマー候補 3 について、セルフダイマーの形成とヘアピン構造の形成について、塩基配列内の相補配列を探索することにより予測した。セルフダイマーの形成とヘアピン構造については、アニーリング温度も算出し、セルフダイマーについては 20°C 以上ヘアピン構造については 16°C 以上が予測された場合、それらの構造が安定で、PCR に使用した場合に副次産物が大量に形成されることが予測されるため、候補から棄却した。これらの棄却の対象とならなかったプライマー候補をプライマー候補 4 とした。

⑩ プライマー候補 4 による仮想 PCR (図 1J)

プライマー候補 4 について④に述べた方法で計算上の PCR 産物の形成を予測した。プライマー候補 4 は、いずれかの PCR 産物の形成に関与することが⑧で明らかであるため、PCR を形成する二本のプライマー候補 4 をプライマーペア候補とした。

⑪ プライマーペア候補の評価 (図 1K)

リストされた各プライマーペア候補についてプライマーダイマーの形成とそのアニーリング温度を予測した。アニーリング温度が 20°C

を越える場合、当該ペアをプライマーペアリストから棄却した。この時点でリストに残ったプライマーおよびプライマーペアを最終的なプライマーセットとした。

各プライマーセットについては、プライマーの配列と forward および reverse プライマーの degeneracy, ゲノム上の平均位置、PCR 産物の平均サイズ、および forward および reverse プライマーの degeneracy の合計値を記録した。

⑫ プライマーペアの順位づけ

①から⑪の過程を通じてリストに残ったプライマーセットについて、degeneracy の合計値の低い順に順位を付けた。続いて、同一の合計 degeneracy を持つプライマーセットについては、PCR 産物の平均サイズが低い順に順位を付け、高い順位のものをつ型のゲノム断片を合成し、検証実験を行う合成プライマー候補とした。

2) degenerate 塩基の高速相同性検索

CoCoMo アルゴリズムを適切な 6 本以下のウイルス配列に対して適用した場合、1,000 本以上の degenerate プライマーが予測される。CoCoMo アルゴリズムは、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すため、予測されたプライマーセットには多くのウイルスに適用可能であるものが含まれることが予測される。一方、多数の対象ウイルス配列を全て CoCoMo アルゴリズムの対象とした場合には、計算時間が莫大となり、結果が得られない。そのため、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を検討した。

この方法は理論的には、プライマー選択の考え方に合致しており、実際に少数のウイルス配列ではうまく稼働する。しかしながら、多数のウイルス配列を対象として多数の

degenerate プライマーを検討する必要があるのに対し、degenerate 塩基を含む配列とウイルス配列の相同性検索には長大な計算時間がかかる(図 2A)。

そこで、報告者は、塩基配列を 4 桁の二進数に変換し、degenerate 塩基と通常の塩基とのその相同性を高速で算出するプログラムを開発した。塩基配列を 4 桁の二進数にした場合、通常の塩基はいずれかの 1 桁のみが 1 ビットである数値として表現可能である(図 2B)。この記載方法を用いた場合、混合塩基は二つ以上の桁が 1 である数値として表現される(図 2B)。この記載方法に基づけば、degenerate 塩基と通常塩基のホモロジーは、degenerate 塩基での 1 となっているビットの数で、degenerate 塩基と通常塩基で一致するビットの数を割った値となる(図 2C)。

この方法により相同性を算出する場合、4 桁の二進数は 1 桁の 16 進数で表現できる。たとえば、ATGCCAT という塩基配列は、二進数で表現すると(0001), (1000), (0100), (0010), (0010), (0001), (1000)、16 進数で表現すると、1842218 という数値で示すことができる。この表現方法では、degenerate を含む全ての塩基配列は 16 進数の連続として表現可能であり、ホモロジーの算出は 16 進数の論理演算として表現できる。このような 16 進数の演算は文字列での演算に比べて、コンピュータの内部演算との親和性が高いため、高速になることが知られている。本研究では、このような計算プログラムの採用により、degenerate 塩基のホモロジーを多数の鋳型ウイルス配列に対して検索することが可能となった。

5. 対象ウイルス塩基配列のグループ化

対象ウイルスを分類してグループ化するためには、第 1 にウイルス塩基配列間の相同性を BLAST2 プログラムによって、全ての組み合わせについて求めた。続けて、それらの相同性を図示することにより、図上のデータを目視に

よってグループ化した。図示には GraphViz (<http://www.graphviz.org/About.php>) を使用した。また、図示に加え、自動的分類を mcl (<http://www.micans.org/mcl/>) に補助的に利用した。

6. プライマー検証のためのウイルスゲノム断片の合成

① DNA 合成用連結オリゴマー設計プログラム(図 2)

設計されたプライマーの増幅効率を検討するためには、プライマーの鋳型となる種々のウイルス断片が必要となるが、実験室では入手および取り扱いが困難なウイルスも存在するため、連結するオリゴマーをアニーリングおよび PCR することにより 100~400 塩基程度の DNA 断片を合成した(OE-PCR 法)。OE-PCR 法に用いる 19~60 塩基のオリゴマーは、基本的には Hoover と Lubkowski の方法(Hoover and Lubkowski, 2002)に従い、DNA Works2 プログラムを用いて設計された(図 3, <http://helixweb.nih.gov/dnaworks/>)。

すなわち、合成対象となる塩基配列の 5' 端からアニーリング温度が 70°C と予測される長さの塩基配列を第一オリゴマーとして設定し、続けて、第一オリゴマーの 3' 端からアニーリング温度が 60°C となるように第二オリゴマーとの重複領域を設定した。続けて第一オリゴマーと相補的で第一オリゴマーとの重複部分を含めた配列長が 55 塩基程度のオリゴマーを設計し、第二オリゴマーとした。さらに、この第二オリゴマーの 5' 端からアニーリング温度が 60°C となるように第三オリゴマーとの重複領域を設定した。このオリゴマーを追加する作業を対象塩基配列の 5' 端まで繰り返した。

設定されたオリゴマーは、全て依頼合成され(北海道バイオシステム)、OE-PCR 反応に用いられた。

②OE-PCR 反応

OE-PCR は、3' →5' exonuclease 活性を持

つ耐熱性 DNA polymerase (PrimeStar HS)により実施された。

すなわち、5×PrimeSTAR Buffer (Mg²⁺ plus) 10 μl、1×dNTP Mixture (2.5 mM) 4 μl、2 pmol の各 Oligomer、および PrimeSTAR HS DNA Polymerase (2.5 U/μl) 0.5 μl と滅菌蒸留水を up to 50 μl となるように混合し、[98°C 10 sec - 60°C 5 sec - 72°C 10sec]の条件で 30 サイクル反応し、一次鋳型を作成した。続けて、5×PrimeSTAR Buffer (Mg²⁺ plus) 10 μl、1×dNTP Mixture (2.5 mM) 4 μl、20 pmol 5' 端の Oligomer、20 pmol 3' 端の Oligomer および PrimeSTAR HS DNA Polymerase (2.5 U/μl) 0.5 μl と滅菌蒸留水を up to 50 μl となるように混合し、[98°C 10 sec - 60°C 5 sec - 72°C 10sec]の条件で 30 サイクル反応し、プライマー検証用鋳型を作成した。プライマー検証用鋳型は、10⁹~1 copy になるように純水で希釈され、degenerate プライマーの鋳型として使用された。

7. degenerate プライマーによる PCR

OE-PCR 反応により作成されたウイルスゲノム断片を鋳型とした degenerate プライマーによる PCR は、GoTaq Green PCR Mix (プロメガ) によって実施された。反応条件としては、タッチダウン PCR 法を用いた。

すなわち、2×GoTaqMix (Mg²⁺ plus) 12.5 μl、50 pmol の各 degenerate プライマー、OE-PCR で作成された鋳型 DNA に滅菌蒸留水を 25 μl となるように混合し、反応溶液とした。

反応は 95°C 1 min の後、[94°C 30 sec - 67°C 30 sec - 72°C 60 sec]の反応を一回おこない、続けて、[94°C 30 sec - 66°C 30 sec - 72°C 60 sec]の反応を一回行った。このように各回アニーリング温度を 1°C ずつ下げつつ 10 サイクルの反応を実施した。

続けて、[94°C 30 sec - 55°C 30 sec - 72°C 60 sec]の反応条件で 20 サイクルの反応をした。

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、プライマー設計アルゴリズムの実証対象として、ヒトサイトメガロウイルスおよびヒトアデノウイルスを選択した。いずれのウイルスについても、検出対象を複数種に設定することが可能で、対象となった種には遺伝的多型が幅広く認められたため、本方法の適用事例として適切と考えられた。結果の項に、双方のウイルスの概説を述べた後にプライマー設計および検証の結果を記載する。

2. ヒトサイトメガロウイルス

ヒトサイトメガロウイルス (HCMV, HHV-5) は、βヘルペスウイルス亜科に属するウイルスで、HHV-6 および HHV-7 と比較的近縁であると言える。HCMV は、日本人成人では約 90%のヒトが感染しているが、健康なヒトが発症することはまれであり、通常は不顕性感染として経過し、生涯体の中に潜伏感染している。しかし、後天性免疫不全症候群患者などの易感染性宿主では、肺炎・消化管潰瘍・網膜炎・脳炎など、場合によっては死に至る重篤な感染症を引き起こす。

HCMV、HHV-6 および HHV-7 は、それぞれ 235kbp、321kbp および 153kbp のゲノム上に 165、192 および 86 種の蛋白質をコードしていることが知られている。プライマーの設定は、蛋白質をコードしている遺伝子ごとに行うため、これらのウイルスに関する遺伝子の GenBank 上の報告配列 736 件を対象配列として、相同性によってグループを生成した。

実際のグループ生成にあたっては、対象が遺伝子に限定された点と、塩基配列では十分な相同性が見られないことがあったため、各遺伝子から予測される蛋白質のアミノ酸配列について相同性を求めた。736件から可能な2遺伝子の組み合わせ ${}_{736}C_2=736 \times 735 / 2 = 270,480$ 組を全て選択してアミノ酸配列について blast2 プログラムにより、相同性を求めた(直接の結果は省略)。アミノ酸配列上で80%以上の相同性を持つ場合に二つの遺伝子を「相同性あり」として線で結ぶこととし、相互に線で結ばれた「クラスター」の中に含まれる遺伝子を保有するウイルスをカウントした。プライマー設計の対象遺伝子グループとしては、クラスターを構成する遺伝子をコードするウイルスに HHV-5, HHV-6 および HHV-7 の全てが含まれる6つのクラスターを選択した(図3)。

選択されたクラスターは、それぞれ主として DNA packaging terminase subunit 2、single-stranded DNA-binding protein、DNA polymerase catalytic subunit、helicase-primase helicase subunit、nuclear egress lamina protein および major capsid protein をコードする遺伝子群であった(表2)。それぞれの選択クラスターについて CoCoMo アルゴリズムによりプライマーの設計を実施したところ、第6クラスターでのプライマー予測数が最大となった(表3)。プライマー予測結果から、この第6クラスターが HHV-5, HHV-6 および HHV-7 の共通プライマー予測上最適であることが示唆され、そのクラスターでの予測プライマー中最適と判断されたプライマーについて人工遺伝子断片による PCR 試験を実施することとした(表4)。

PCR 試験に用いる人工遺伝子断片を合成したところ、第2段階の PCR で単一のバンドが得られた(図5)。これらの人工遺伝子をコピー数 $10^7 \sim 10^0$ となるように希釈し、最適と予測されたプライマーにより PCR を実施した(図6)。その結果、HHV-5 では鋳型 10 コピーまでバンド

が検出されたが、HHV-6A、HHV-6B および HHV-7 ではそれぞれ 10^3 、 10^3 および 10^4 までしかバンドが検出されなかった。

3. ヒトアデノウイルス

アデノウイルス科、マストアデノウイルス属(哺乳動物に感染するアデノウイルス)のヒトアデノウイルス群のウイルスは、ヒトに感染した場合、上気道感染、咽頭炎、結膜炎、胃腸炎などを引き起こす。プール熱や夏風邪の原因ウイルスの一つ。ヒトアデノウイルスには30以上の血清型がある。

ウイルスは、径70-90nmの正二十面体構造でエンベロープは持たない。正二十面体の各頂点にペントンファイバー(penton fiber)とよばれる細い突起がついている。ゲノムは二重鎖DNAで、分子量 $2.0 \sim 2.5 \times 10^7$ 、約36,000塩基対からなる。両方の5'末端にDNAの複製開始に重要な役割をはたす蛋白質(VPg)が結合している。粒子を穏やかに処理して遺伝子を分離すると、両末端の結合タンパク質が互いに会合して、DNAが環状になっているように見える遺伝子が得られる。

ヒトアデノウイルスについては、GenBank上でゲノム配列全長が報告されている44個のゲノムデータ(表5)からDNA polymerase 遺伝子の配列を抽出して設計対象とした。設計用のグループはDNA polymerase 遺伝子の塩基配列の相同性に基づいて、blast2プログラムでの相同性80%を基準としてグループを生成した(図7)。この基準で生成されたグループは、文献上で報告されている亜種と一致したグループとなった(表6)。

A, B1, B2, C, D, E および F の5亜種から各1ゲノムを選択して CoCoMo アルゴリズムでのプライマー設計対象とした(図7)。

選択された7種のDNA polymerase 遺伝子からは、252組のプライマーペアが予測された(表7)。全ての予測プライマーペアについて鋳型への適用性を検討した。その結果、30種の

プライマーが6亜属全てについてPCRの可能性があることが示唆された(表8)。これらのうち、degeneracyが最低であり、PCR産物サイズが最小であるプライマーセット(CC-1, Forward: GGCCATGTAYGCMTCBGC, Reverse: IGGRGGGTCNGCRTC)を試験プライマーとすることとした。このプライマーについてPCR試験を行うため、増幅予測箇所を人工遺伝子断片として合成した(表9)。

PCR試験に用いる人工遺伝子断片を合成したところ、多くの第2段階のPCRで単一のバンドが得られたが、一部のウイルス型では副次バンドが形成された(図8)。副次バンドが存在している中でのPCR試験は、結果に対して有利に働くことは無いと判断されたため、副次バンドも含めて希釈を実施した。人工遺伝子はコピー数 $10^7 \sim 10^0$ となるように希釈された。最適と予測されたプライマーによりPCRを実施したところ、一部の型では鋳型10コピーまでバンドが検出されたが、一部の型ではそれぞれ 10^3 までしかバンドが検出されなかった(図9)。

D. 考察

本研究では、前年度に開発したCoCoMoプライマー設計アルゴリズムを活用し、血液ウイルスの検出を目的として設計された多数のプライマーについて適用範囲を拡大するプログラム群を整備した。前年度開発したCoCoMoアルゴリズムは、塩基配列の相同情報を元に対象配列を整列させ、さらに、配列間に散在する短い相一致領域を検出してプライマーを選定していくという機能を備えており、HBV, HCV, HIV1 および HEV ウイルスの共通プライマー設計においては、既存のプログラムに比べ高い設計能力を示した。しかしながら、このプログラムについても、設計時の適用範囲を10種以上のウイルスに拡大すると、計算時間が級数的に延長し、さらに、相同性の低い配列の混入によりプライマーの予測が得られなくなるという問題があった(2008年度報告書参照)。

そこで、本研究では、PCR対象範囲を拡大するために、プライマー設計のための高速検索ステップを追加し、CoCoMoアルゴリズムでプライマーを予測した後、プライマー適用ウイルスを探索するという方法を開発した(図2)。この第二の高速探索アルゴリズムの導入により、CoCoMoアルゴリズムで予測したプライマーから、共通性の広いプライマーペアを選択することが可能となった。

具体的な取り組みにおいては、degenerateプライマーの相同性分析には、元々長い計算時間を要したため、独自の二進数演算化(図2)によって相同性探索を高速化した。この高速化によって探索効率が増大したため、CoCoMoアルゴリズムによるプライマー設計の際に、設計対象ウイルスを限定して多数の予測プライマーを算出するという選択肢も生まれた。予備的に、①CoCoMoアルゴリズムで多数のウイルスを対象にして少数のプライマーセットを得てEstimateプログラムで適用ウイルスを探索する方法と②CoCoMoアルゴリズムで少数のウイルスを対象にして多数のプライマーセットを得た後Estimateプログラムで適用ウイルスを探索する方法を予備的に試したところ、ヒトヘルペスウイルス5, 6 および7 およびヒトアデノウイルスでは、CoCoMoアルゴリズムでウイルス種を限定した場合の方が結果として多数の共通ウイルスがEstimateプログラムで予測された。

CoCoMoアルゴリズムで少数のウイルスを対象にして多数のプライマーセットを得た後Estimateプログラムで適用ウイルスを探索する方法を活用するために、次に、少数のCoCoMoアルゴリズム対象ウイルスの設定方法を決定する必要があった。比較的低い相同性に基づいてウイルスの関係を探索する方法としては、ネットワーク図の構築が有効であることが、2007年および2008年度の本研究でも示された。今年度は、ネットワーク図の分析効率を高めるために、図示にGraphvizを、計算上のグループ

分けに mcl を導入した。これらのプログラムをウイルス間の相同性に適用させて関係を図示することにより、プライマー設計のためのグループをクラスターとして予測することが可能となった。クラスターの探索には図示された相関図を元にクラスターを主導的に検出することが必要であった。一方、Estimate プログラムの導入により拡大したウイルス数が増大した場合には目視によるクラスター形成は難しい可能性もある。そこで、本年度は自動でクラスターを検出するソフトを並行して使うことにより、再度、HBV, HCV, HIV1, HEV 等の血液安全対象ウイルス全般を検出するためのプライマーの再設計にも適用可能な準備を進めた。自動的なクラスター化により、10,000 件を越すような、莫大なウイルス塩基配列を元にしたプライマーを設計することも可能になると推定される。本年度ヒトヘルペスウイルス 5(サイトメガロウイルス), 6 および 7 の遺伝子群について遺伝子産物のアミノ酸配列を元にプライマー設計を試みた。これらのウイルス全体をカバーするプライマーペアが設計された。

同様に、ヒトアデノウイルスについては DNA polymerase の塩基配列に基づく分析によるグループ化が可能であった。結果として CoCoMo アルゴリズムの連続的使用により 6 種の亜属全てについてプライマーが設計が可能であることが示された。これらのプライマー中最適と判断されたプライマーペアについては、全候補ウイルスについて PCR が可能であることが実験的にも示された(図 9)。

ヘルペスウイルスおよびアデノウイルスについては、今回二組のプライマーペアについての増幅効率を検討した。検討プライマー数が限定された理由は、鋳型合成コストにある。鋳型の合成には、多数のオリゴマーが必要とされ、今回のような多数のウイルスを対象とする場合には、そのコストがさらに膨大になる。従って、本報告では、実 PCR 試験を実施したプライマー例が限定されることとなった。一方、今後

は、このプライマー

これらの結果から、本年度設計されたプライマー設計プログラム群により、新興ウイルスも含んだ多様なウイルスの検出に利用可能であることが示唆される。

E. 総括

本年度においては、新興再興ウイルス感染症に対応することを目的として、前年度に開発した CoCoMo アルゴリズムに対して、設計プライマーの関連ウイルスへの適用を高速に検索するプログラム群を開発した。実施例として、ヒトヘルペスウイルスおよびヒトアデノウイルスを選択し、開発プログラム群により、複数種を共通して増幅するプライマーを設計した。共通プライマーは、人工的に合成されたウイルス遺伝子断片を増幅したことから、degenerate プライマー設計と適用性の高速探索を連携した方法は、多ウイルス種・多ウイルス株の同時増幅プライマー設計に有効であることが示唆された。

F. 健康危険情報

該当なし。

G. 研究発表

1. 論文発表

1). Daiji Endoh, Tetsuya Mizutani, Shigeru Morikawa, Isao Hamaguchi, Koji Sakai, Kazuya Takizawa, Yuuichi Osa, Mitsuhiro Asakawa, Yasuhiro Kon and Masanobu Hayashi, CoCoMo-primer: a web server to design degenerate primers for virus research, Nucleic Acid Research 誌投稿中

2. 学会発表

1). 遠藤大二 「共通オリゴマー分析に基づく多種ウイルス共通プライマーの設計」 日本獣医学会学術集会、2009 年 4 月、栃木

H. 知的財産権の出願・登録状況

4. 著作権を設定し、公表予定のプログラム
コード：省略
5. インターネット上で利用可能なプログラム
公開サイト：
<http://www.geneknot.jp/cocomo>

図表

表 1 塩基配列のコードと degeneracy

コード	意味	コードの由来	degeneracy	Degeneracy 指数
G	G	Guanine	1	0
A	A	Adenine	1	0
T	T	Thymine	1	0
C	C	Cytosine	1	0
R	G or A	puRine	2	1
Y	T or C	pYrimidine	2	1
M	A or C	aMino	2	1
K	G or T	Keto	2	1
S	G or C	Strong interaction (3 H bonds)	2	1
W	A or T	Weak interaction (2 H bonds)	2	1
H	A or C or T	not-G, H follows G in the alphabet	3	1.58
B	G or T or C	not-A, B follows A	3	1.58
V	G or C or A	not-T (not-U), V follows U	3	1.58
D	G or A or T	not-C, D follows C	3	1.58
N	G or A or T or C	aNy	4	2

表2 HHV-5, HHV-6 および HHV-7 共通検出プライマー設計対象とした遺伝子群

Cluster No	ID	Organism	Product	Nucleotide length	Amino Acid length
1	14	Human herpesvirus 5	DNA packaging terminase subunit 2	2553	850
	282	Human herpesvirus 6A	DNA packaging terminase subunit 2	2181	726
	283	Human herpesvirus 6B	transport/capsid assembly	2181	726
	284	Human herpesvirus 6B	transport protein	2181	726
	285	Human herpesvirus 6	U40	2181	726
	586	Human herpesvirus 7	DNA packaging terminase subunit 2	2166	721
	587	Human herpesvirus 7	U40	2166	721
2	4	Human herpesvirus 5	single-stranded DNA-binding protein	3708	1235
	234	Human herpesvirus 6	U41 major DNA binding protein	3399	1132
	232	Human herpesvirus 6A	single-stranded DNA-binding protein	3399	1132
	235	Human herpesvirus 6B	U41	3399	1132
	233	Human herpesvirus 6B	major DNA binding protein	3399	1132
	557	Human herpesvirus 7	U41	3396	1131
	558	Human herpesvirus 7	major DNA binding protein	3396	1131
3	3	Human herpesvirus 5	DNA polymerase catalytic subunit	3729	1242
	244	Human herpesvirus 6	U38 DNA polymerase	3039	1012
	242	Human herpesvirus 6A	DNA polymerase catalytic subunit	3039	1012
	243	Human herpesvirus 6B	DNA polymerase	3039	1012
	245	Human herpesvirus 6B	U38	3039	1012
	559	Human herpesvirus 7	DNA polymerase catalytic subunit	3042	1013
	560	Human herpesvirus 7	U38	3042	1013
4	7	Human herpesvirus 5	helicase-primase helicase subunit	2871	956
	264	Human herpesvirus 6A	helicase-primase helicase subunit	2475	824
	265	Human herpesvirus 6B	helicase	2475	824
	266	Human herpesvirus 6B	U77	2475	824
	572	Human herpesvirus 7	helicase-primase helicase subunit	2463	820
	573	Human herpesvirus 7	U77	2463	820
	574	Human herpesvirus 7	helicase	2463	820
5	64	Human herpesvirus 5	nuclear egress lamina protein	1131	376
	445	Human herpesvirus 6	U37	795	264
	444	Human herpesvirus 6A	nuclear egress lamina protein	795	264
	688	Human herpesvirus 7	nuclear egress lamina protein	780	259
	689	Human herpesvirus 7	U37	780	259
6	2	Human herpesvirus 5	major capsid protein	4113	1370
	227	Human herpesvirus 6	U57 major capsid protein	4038	1345
	226	Human herpesvirus 6A	major capsid protein	4038	1345
	228	Human herpesvirus 6B	U57	4038	1345
	548	Human herpesvirus 7	major capsid protein	4038	1345
	549	Human herpesvirus 7	U57	4038	1345