

るため、GTG および CCA の双方が削除されることとなる。この相補的な motif の除去は、相補的な組み合わせが存在しなくなるまで実施する。

3mer の事例では、一回目の削除で全ての相補的組み合わせが無くなったが、実際の事例では、多数回この作業が繰り返された。想定している事例では、一方の 3mer を候補から削除した後、仮想 PCR は表 8 のように変更され、PCR がおきることが予想される数は 12 件に減少した。3mer の事例では、ATG-ATG および ATG-TGC の組み合わせについて 5 件の PCR が予想されるため、これらが選択されることとなる。実際の実施時には、完全に対象ゲノム上での重複がない 6 塩基 motif が選択された。

最終的に選択された motif の 5' 側の配列を手作業で調べ、実際の対象配列上に存在する塩基の多様性をすべて満たす degenerate プライマーを作出した。

2) 平成 20 年度

CoCoMo Ver3 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った。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、使用されなかっ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℃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⑩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、高い順位のを鋳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。

6. 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 Pajek (<http://pajek.imfm.si/doku.php>, 平成 19 年～20 年) または GraphViz (<http://www.graphviz.org/About.php>, 平成 20 年～21 年) を使用した。

平成 21 年度には、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。

3) degenerate 塩基の高速相同性検索(平成 21 年度)

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(図 3A)。

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 3B)。この記載方法を用いた場合、混合塩基は二つ以上のの桁が 1 である数値として表現される(図 3B)。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 3C)。

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. 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 4)

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 DNA Works2 プログラムに基づいて設計された(図 2, <http://mcl1.ncifcrf.gov/dnaworks/dnaworks2.html>)。本研究では、多数のゲノム断片を合成する必要があったため、OE-PCR のためのオリゴマーを設計するプログラムを、Hoover と Lubkowski の方法(Hoover and Lubkowski, 2002)に従い作成した。

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、OE-PCR 反応に用いられた。

② OE-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。

6. ウイルス感染実験、RNA の抽出および cDNA の合成

1) ウイルス感染細胞の作成(平成 19 年度)

平成 19 年度においては、感染モデルとして SARS コロナウイルスを使用した。SARS コロナウイルス感染細胞の作成には Vero E6 細胞と、Frankfurt 1 株として分離された SARS コロナウイルスを用いた。Vero E6 細胞は通常 75 cm² のフラスコ内で、0.2 mM/ml の L-glutamine、100U/ml のペニシリン、10 μg/ml のストレプトマイシン、5%(v/v)の牛胎子

血清(FBS)を添加した DMEM (Sigma-Aldrich)中で経代培養され、5% CO₂ 大気中に 37°Cで維持された。細胞は 25 cm² フラスコにいったん分けられ、100%集密に達するまで培養された。その培養液はウイルス感染の前に、2%の FBS を含んだ DMEM に置き換えられた。その後、SARS コロナウイルス Frankfurt 1 株の感染は、感染効率(m.o.i)が 10 の細胞内で行なわれた。実験は感染症研究所の村山分室の P4 施設で実施された。

2) 感染細胞からのウイルスゲノムを含んだ RNA の精製

RNA の抽出には、Isogen (Nippon gene) を用いて定法に従った。すなわち、Isogen に浸漬した細胞 200 μ l に Isogen 600 μ l を加え、ボルテックスをかけた後、12,000 rpm で 15 分遠心し、上精を採取した。上精に 500 μ l のイソプロピルアルコールを加え、13,000rpm で 10 分遠心、沈殿を 500 μ l の 75%エタノールで洗浄し、13,000rpm で 5 分間遠心した。上精を捨てた後、ペレットを完全に乾燥させ、RNase free water (DDW)30 μ l に溶出し、RNA として用いた。

3) cDNA 合成

cDNA の合成には、M-MLV Reverse Transcriptase (SIGMA-Aldrich ジャパン)および Superscript III (INVITROGEN)を用い、それぞれ 1 本鎖、2 本鎖 cDNA 合成を行った。方法は各製品の説明書に従った。すなわち、Random Primer もしくは NR9 6 μ l、10mM dNTP mix 1 μ l、RNA RNA 溶液 1 μ l、DDW 7 μ l を氷上で混合した後、70°C で 10 分間インキュベートした。その後反応物を氷上で急冷し、10 \times M-MLV Reverse Transcriptase Buffer 2 μ l、続いて 2 本鎖目の cDNA を合成するため、M-MLV Reverse Transcriptase 1 μ l、RNase Inhibitor (20 U / μ l) 1 μ l、DDW 6 μ l を氷上で混合した。37°C

で 50 分間インキュベートし、1 本鎖 cDNA を合成した。

M-MLV Reverse Transcriptase で合成した 1 本鎖 cDNA 全量 20 μ l、DDW 83.5 μ l、5 \times Second-Strand Reaction Buffer 30 μ l、10 mM dNTP mix 3 μ l、E.coli DNA Ligase (10U/ μ l) 1 μ l、E.coli DNA Polymerase I (3.5 U/ μ l) 11.5 μ l、E.coli RNase H (2 U/ μ l) 1 μ l を混合し、16°C で 2 時間インキュベートした。その後 T4 DNA Polymerase (5 U/ μ l) 2 μ l を加え、16°C で 5 分間反応を続けた。反応物を氷上に置き、0.5 M EDTA 10 μ l を混合し、2 本鎖 cDNA を合成した。合成終了後、フェノールおよびクロロホルム抽出により反応を停止し、続けてエタノール沈殿を行った。すなわち、合成終了後の溶液に同量のフェノールを加えた後ボルテックスをかけ、12,000rpm で 20 分間遠心後、水層を新たなチューブに分取した。クロロホルムについても同様の操作を施した後、溶液の 1/10 量の 3M 酢酸ナトリウムと 0.1 μ l のグリコーゲン (Roche Applied Science) を加え、さらに 2.5 倍量の 99.5%エタノールを加えて転倒混和した。室温で 13,000rpm、15 分間遠心後、チューブの底に核酸の沈殿を確認し、上精を完全に廃棄した。さらに 70%エタノールを 1 ml 加えた後同様の操作を行い、ペレットを風乾させ、30 μ l の DDW に溶解させた。DDW に溶解させた 2 本鎖 cDNA を SpinClean PCR Purification Kit (M-Biotec) で精製した。

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 \times GoTaqMix (Mg²⁺ plus) 12.5 μ l、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 μ l 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 づつ下げつつ 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。

7. 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. 研究結果および考察

B.

1. 平成 19 年度

1). 結果

① SARS コロナウイルスの検出

最初の適用事例として、CoCoMo アルゴリズムを SARS コロナウイルスに適用することを検討した。SARS コロナウイルスには、多様性が知られているが、クラスター状に保存領域が存在するために、本アルゴリズムの適用例として適切であることが予想された。

設計に先立ち、コロナウイルス科のゲノム配列の相同性を調べ、Pajek を利用してグループに分類した(図 5)。SARS コロナウイルスはコロナウイルス中では、島状のクラスターを形成していたため、設計の対象とした(表 9)。

鋳型としては、SARS コロナウイルス感染細胞抽出 RNA から合成した cDNA を用いた。RT-PCR では、感染細胞のゲノムが混入している状態での増幅産物を検討した。PCR の結果、

レーン 2 で PCR 産物が得られなかった事以外は全て予想されたサイズの増幅産物が見られた(図 6)。

② HIV-1 検出用 degenerate プライマーの設計

前述 2 2)に示したプログラム群を利用して、種内の変異株全体をカバーするプライマーをデザインした(図 7)。

HIV-1 については、取得されたゲノム塩基配列数が多数であり、かつ多様であったため、social network analysis で形成されたグループごとにグループ内の鋳型を共通して増幅するプライマーを設計した(表 10-12)。

③ HCV 検出用 degenerate プライマーの設計

HCV については、social network analysis(表 13,)で形成された一次グループの代表をグループとしてまとめなおした場合も、代表全体をカバーする degenerated プライマーが設計されたため、それらの degenerated プライマーを種全体をカバーするプライマーとみなした。

GenBank で HCV のデータを検索したところ、28,988 件の塩基配列の登録があり、全てをダウンロード、MySQL データベースに格納後、サイズおよび相同性に基づいてグループに分類したところ、241 のグループに分かれた(表 18)。全ゲノムのサイズと一致するグループには 235 件の塩基配列が分類されたため、このグループについて、相互の相同性を求めた後、social network 分析に基づいて 9 個のグループに分割した。続けて、それぞれのグループから代表配列選択して代表グループを形成した後、代表グループを網羅する primer(HCV ゲノム共通 degenerate primer)を設計した(表 14)。それらの一部を primer として、標準 HCV ウイルス由来 cDNA を鋳型として PCR を実施したが、確実な増幅は 12 件中の 1 件のみで観察された(図 8)。

2). 考察

① CoCoMo アルゴリズムの適用範囲

本研究では、CoCoMo アルゴリズムによって SARS コロナウイルス、HIV-1 および HCV ウイルスを検出するための primer を設計した。いずれの場合についても、基本的には、①対象塩基配列の入手、分類および対象塩基配列グループの選定→②塩基配列間の相同性と social network 解析に基づくグループ化→③対象塩基配列上の 6 塩基 motif の選定→④PCR の位置にある相補的 6 塩基 motif の削除→⑤選定されたオリゴマーの 5' 領域 15 塩基の選択と共通 degenerated 配列の決定、という処理順序で degenerated primer を抽出することができた。これらのステップのうち、③以降は塩基配列グループによって大きな差はなく、一旦対象グループが決定されれば自動的に進められることが示唆された。ただし、相補オリゴマーの削除ステップには多大な計算資源が必要とされるため、実際には処理が可能な各塩基配列の長さや塩基配列数は限定されている。その限界内にグループを設定するための①での分類グループの選定と②での効率的なグループの設定が設計された primer の適用範囲と検出感度に大きく影響することが示唆された。

② SARS コロナウイルスへの応用

SARS コロナウイルス感染細胞から鋳型を作成することで、感染細胞のゲノムが混入している状態での自動化して設計された primer の有効性を検討した。PCR の結果、33 サンプル中 1 サンプル以外は全て予想された増幅産物が得られた (図 6)。このサンプルに PCR がかからなかった原因としては、おそらく設計と実験に用いた株の差や、または変異があった可能性が考えられる。実験全体としてはほぼ全てのサンプルに PCR がかかったことから、この primer 設計方法は実際のウイルスに対しても有効であるということが示唆された。

② HIV-1 および HCV 用 degenerated primer の合成

HIV-1 および HCV に対しての primer 設計では、それまでの事例に比べて遺伝子データベース上に大量のデータが蓄積されているため、基本的には計算機上での primer 設計には有利な状況であると考えられた。ただし、実際の設計段階において、6 塩基 motif を基本とする CoCoMo アルゴリズムでは計算機の機能限界が問題となった。そのため、適切なグループを構成し、ある程度限定されているが一方では良くゲノムグループ全体の配列情報を代表しているグループの作成が degenerate primer 設計上は重要であることが示唆された。HIV-1 においては、塩基配列を限定する過程で、ゲノム全体の相同性比較から HIV-1 を 3 種のグループに分け、それぞれについて degenerate primer を設計した。この設計 primer は

比較的効率良い PCR 結果を出していることから、有効な方法と考えられる。しかしながら、現在 social network 分析に使用した Pajek が図示可能なデータが 200 程度に限定されている。HCV では、図示に合致する数に塩基配列データを限定することが難しかったため、図示に代わる方法として仮想空間上での分類を行うソフト(create_group_sns)を利用した(図 7)。しかしながら、設計された degenerate primer は期待された効率を示していないことから、social network 分析用ソフトの変更を含むグループ化のための方法の変更が必要と考えられる。

3). 小括

特定の種の塩基配列データを大量にデータベース上から入手し、グループ化と motif 分析を組み合わせることにより、効率良く degenerate primer を設計可能なアルゴリズム (CoCoMo アルゴリズム) およびプログラム群が整備された。合成オリゴマーを用いた実験と 3 種のウイルスでの適用試験では、CoCoMo アルゴリズムによって設計された degenerate primer は一定の効率を示した。これらの結果

から、CoCoMo アルゴリズムの実用性が示されたが、同時にプログラムの改善と個々の適用種に応じた方法を変更する必要性が示唆された。

2. 平成 20 年度

1). 結果

① HCV

Genotype 1-6 に分類される Subtype ウイルスゲノム 24 件の塩基配列データ(表 15)を元にプライマーを設計した。ギャップ motif の局在に基づく CoCoMo アルゴリズムによる分析結果として 150 種のプライマーが予測された(表 16)。degeneracy が低く、増幅産物のサイズが比較的小さいプライマーセットを選択し、degenerate プライマーの鑄型となるゲノムフラグメントを予測したところ(表 17)、相同性の高さから 8 個のグループに分けることができた(表 18)。各グループについて、OE-PCR によりゲノムフラグメントを合成するため、オリゴマーを設計した(表 19)。

OE-PCR の結果全ての合成ゲノムフラグメントについて、バンドが確認された(図 9)。これらのバンドの PCR 産物を定量後、 $10^9 \sim 1$ copy となるように純水で希釈し、degenerate プライマーにより増幅試験を実施したところ、全ての Genotype について、予想されたサイズのバンドが 1 copy の鑄型についても観察された(図 10)。

② HBV

Genotype A~H に分類される 8 件の塩基配列データ(表 20)を元にプライマーを設計した。ギャップ motif の局在に基づく CoCoMo アルゴリズムによる分析結果として 1000 種以上のプライマーが予測された(表 21)。degeneracy が低く、増幅産物のサイズが比較的小さいプライマーセットを選択し、degenerate プライマーの鑄型となるゲノムフラグメントを予測した(表 22)。それらの配列に基づき各 Genotype について、OE-PCR により

ゲノムフラグメントを合成するため、オリゴマーを設計した(表 23)。

OE-PCR の結果全ての合成ゲノムフラグメントについて、バンドが確認された(図 11)。これらのバンドの PCR 産物を定量後、 $10^9 \sim 1$ copy となるように純水で希釈し、degenerate プライマーにより増幅試験を実施したところ、全ての Genotype について、予想されたサイズのバンドが 1 copy の鑄型についても観察された(図 12)。

③ HIV-1

Genotype A~H に分類される 8 件の塩基配列データ(表 24)を元にプライマーを設計した。ギャップ motif の局在に基づく CoCoMo アルゴリズムによる分析結果として 1000 種以上のプライマーが予測された(表 25)。degeneracy が低く、増幅産物のサイズが比較的小さいプライマーセットを選択し、degenerate プライマーの鑄型となるゲノムフラグメントを予測した(表 26)。それらの配列に基づき各 Genotype について、OE-PCR によりゲノムフラグメントを合成するため、オリゴマーを設計した(表 27)。

OE-PCR の結果、全ての合成ゲノムフラグメントについて、バンドが確認された(図 13)。これらのバンドの PCR 産物を定量後、 $10^9 \sim 1$ copy となるように純水で希釈し、degenerate プライマーにより増幅試験を実施したところ、Genotype A~G についてはゲノムフラグメント 10^5 copy まで、Genotype H についてはゲノムフラグメント 10^9 copy まで、予想されたサイズのバンドが観察された(図 14)。

④ HEV

GenBank 上に登録されている Genotype 1-4 HEV、Avian HEV および Swine HEV のゲノム塩基配列をダウンロード後、ホモロジーの高い配列をグループ化したところ 15 件の配列に集約された(表 28)ため、これら

の塩基配列を元に degenerate プライマーを設計した。ギャップ motif の局在に基づく CoCoMo アルゴリズムによる分析結果として 1000 種以上のプライマーが予測された(表 29)。

続けて、degeneracy が低く、増幅産物のサイズが比較的小さいプライマーセットを選択し、degenerate プライマーの鋳型となるゲノムフラグメントを予測した(表 30)。それらの配列に基づき各 Genotype について、OE-PCR によりゲノムフラグメントを合成するため、オリゴマーを設計した(表 31)。

OE-PCR の結果、全ての合成ゲノムフラグメントについて、バンドが確認された(図 15)。これらのバンドの PCR 産物を定量後、 $10^9 \sim 1$ copy となるように純水で希釈し、degenerate プライマーにより増幅試験を実施したところ、ID 1, 150, 335, 756, 1142, 1567, 1568 および 2095 の鋳型についてはゲノムフラグメント 10^5 copy まで、ID 335, 533 および 1731 の鋳型についてはゲノムフラグメント 10^7 copy まで、予想されたサイズのバンドが観察された(図 16)。

⑤ West Nile virus

GenBank 上に登録されている West Nile virus のゲノム塩基配列をダウンロード後、ホモロジーの高い配列をグループ化したところ 10 件の配列に集約された(表 32)ため、これらの塩基配列を元に degenerate プライマーを設計した。ギャップ motif の局在に基づく CoCoMo アルゴリズムによる分析結果として 1,000 種以上のプライマーが予測された(表 33)。

⑥ Human Parvovirus B19

GenBank 上に登録されている Human Parvovirus B19 のゲノム塩基配列をダウンロード後、ホモロジーの高い配列をグループ化したところ 7 件の配列に集約された(表 32)ため、これらの塩基配列を元に degenerate プライマーを設計した。ギャップ motif の局在に基づく

CoCoMo アルゴリズムによる分析結果として 1000 種以上のプライマーが予測された(表 33)。

2). 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発展させ、相同性の比較的低いウイルスグループにも適用可能なプログラムを開発した。プログラムの新興・再興感染症への適用の可能性を検討するために、HCV, HBV, HIV-1, HEV, West Nile virus および Parvovirus B19 についてそれぞれ共通プライマーを設計した。前年度の方法では、6 塩基 motif について配列および位置が一致ものを探索するが、HIV-1 や HEV などでは配列とおよその位置が一致する motif は見いだされずプライマーは設計されなかったのに対し(結果示さず)、本年度は、対象ウイルスグループのすべてについて共通プライマーが設計された。

degenerate プライマーを motif の局在を基に予測するプログラムは、前年度に初期バージョンが開発され、相互にホモロジーが高いウイルス集団に対して有効な設計が行われた (CoCoMo Ver1)。degenerate プライマーの設計プログラムは、一般的に対象配列数が増えるごとに、また、それぞれの長さが増大するごとに、級数的に計算時間が増大するが、motif の局在に基づく CoCoMo Ver1 は、多数の配列に対しても、計算時間の増大を抑えて短時間でのプライマー設計を可能にした。このことから、motif に基づくプライマー設計という方法は、有効性の高い方法であることが示唆された。

本年度においては、前年度に開発した方法 (CoCoMo Ver1) を改善して、アミノ酸配列を、効率よくプライマー設計につなげるアルゴリズムを検討した。その結果、コドンとの同調が可能な 1 塩基のギャップを含む 4 塩基の motif が、有効な探索用 motif として考案された (CoCoMo Ver3)。

本年度改善されたプログラムは、前年度に

比べて、多様性の高いウイルス群に対する設計効率が高かった。この効率の高さは、ウイルス間共通のギャップ motif が多数存在することと関係していると考えられる。本年度用いたギャップ motif は 2 組の 2 塩基の固定塩基の motif に、1 塩基の自由塩基を加えたものであるが、2 塩基一致の後 1 塩基は相異を許容するという探索方法は、コドンの縮重のパターンと一致しており、アミノ酸コードの変化を伴わないサイレントな変異の影響を受容するデザインとなっている。そのため、アミノ酸配列が一致する蛋白質レベルでの保存領域で、多くプライマー設計用の motif が検出されたものと考えられる。

ギャップ motif は、2 アミノ酸残基のウイルス間の一致に伴って検出されてしまうため、多数の同一配列のギャップ motif が記録され、適切な選択を行わなければ、プライマーの予測には莫大な計算時間が必要になる。そのため、本年度は、motif 選定過程にホモロジーに基づくアライメント(整列)過程を導入した。複数の候補から、精度と計算時間の短さから MAFFT プログラムが選定された。このプログラムの導入により、多数のギャップ motif 一致点から、ウイルス塩基配列上の相同領域を示すギャップ motif 一致点を選択することが可能となった。続けて、仮想 PCR を実施することにより、プライマーとして有効な配列が効率良く選択された。

本年度は、アニーリング温度の算出についても高速のプログラムを整備した。その結果、設計されたプライマー候補についての熱力学的な評価が可能となり、一般的なプライマー設計プログラムと同等の性能をプログラム内に持つに至った。

今回、CoCoMo Ver3 プログラムにより設計されたプライマーについては、HCV および HBV では 1 copy までの検出感度が示された(図 10 および 12)。それに対して HIV-1 および HEV では、感度は $10^7 \sim 10^5$ にとどまった(図 14 および 16)。これらの双方のウイルスについ

て、プライマーは多数予測されたため、今後も継続して感度の高いプライマーを探索することにより、感度の高い検出法を設計できる可能性はある。それらの試験データは、感度と degeneracy や鋳型配列との関係などを推測するための基礎データとして有用であることが期待される。

3). 小括

HEV, West Nile virus および ParvoB19 virus については、報告されている Genotype や subtype が必ずしもプライマー設計に有効ではなかった、本年度は、それらの設計のためにホモロジーが高い分離株をひとつのグループとして、プライマーの設計を行った。結果として、PCR 試験が行われた HEV では良好な結果が得られなかったが、今後の検討によって改善が得られた場合、ホモロジーに基づくグループ分けとその自動化が、プライマー設計上重要な方法になることが予測される。

3. 平成 21 年度

1). 結果

当該年度は、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 17)。

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 36)。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 37)。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 38)。

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 18)。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 19)。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 39)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 20)。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 40)。

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 20)。

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 41)。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 42)。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: GGCATGTAYGCMTCBGC, 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(図 21)。

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 21)。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 $10^7 \sim 10^9$ 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 22)。

2). 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。

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 3)。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 3)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增大したため、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検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、今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(図 19)。

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か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 22)。

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。

3). 小括

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。

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增大したため、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検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。

D. 総括

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して実施された。平成 19 年度においては、プライマー設計のための方法の基本構築を設計した。対象ウイルス間共通プライマーとしては、複数の塩基を同時に規程する degenerate プライマーとすることを決定し、その設計のためには対象ウイルス塩基配列間で共通する motif を起点として、degenerate 配列を設定するためのプログラム群を開発した。このプログラム群で設計された SARS コロナウイルスと HCV での実施試験では、効果的な増幅が観察された。平成 20 年度においては、前年度開発されたプログラムを改善し、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)

として統合的な利用を可能にした。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 HBV, HCV, HIV1, HEV, WNV および Parvovirus B19 ウイルスで示した。このプログラムには、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ったため、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を平成 21 年度に実施した。当該年度においては、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。

これら 3 年間にわたる技術開発により、実用性の高い、プライマー設計プログラムが開発された。

E. 健康危険情報

該当なし。

F. 研究発表

1. 論文発表

1). Daiji Endoh, Tetsuya Mizutani, Shigeru Morikawa, Isao Hamaguchi, Koji Sakai, Kazuya Takizawa, Yuuichi Osa, Mitsuhiko 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

G. 知的財産権の出願・登録状況

1. 著作権を設定し、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
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 塩基配列分析事例における 3mer 頻度

3mer	頻度	3mer	頻度
ACA	1	GCC	1
ATC	1	GGA	1
ATG	3	GGC	1
CAC	1	GTG	1
CAT	3	TCC	1
CCA	2	TGC	1
GAT	1	TGG	2
GCA	1	TGT	1

表 3 塩基配列分析事例における 3mer の位置

位置	センス鎖	位置	相補鎖
1	ATG	1	CAT
2	TGG	2	CCA
3	GGA	3	TCC
4	GAT	4	ATC
5	ATG	5	CAT
6	TGT	6	ACA
7	GTG	7	CAC
8	TGG	8	CCA
9	GGC	9	GCC
10	GCA	10	TGC
11	CAT	11	ATG

表 4 塩基配列分析事例における相互に相補的な 3mer の位置

位置	センス鎖	位置	相補的 3mer	位置	相補的 3mer
1	ATG	1	CAT	2	CCA
2	TGG	2	CCA	3	TCC
3	GGA	3	TCC	4	ATC
4	GAT	4	ATC	5	CAT
5	ATG	5	CAT	6	ACA
6	TGT	6	ACA	7	CAC
7	GTG	7	CAC	8	CCA
8	TGG	8	CCA	9	GCC
9	GGC	9	GCC	10	TGC
10	GCA	10	TGC	11	ATG
11	CAT	11	ATG		

表 5 塩基配列分析事例における PCR 位置に存在する 3mer

センス鎖		反対側 3mer の位置および配列					
位置	配列	6	7	8	9	10	11
1	ATG	ACA	CAC	CCA	GCC	TGC	ATG
2	TGG		CAC	CCA	GCC	TGC	ATG
3	GGA			CCA	GCC	TGC	ATG
4	GAT				GCC	TGC	ATG
5	ATG					TGC	ATG
6	TGT						ATG

表 6 塩基配列分析事例における相互に相補的な 3mer の組み合わせと仮想 PCR での利用 PCR 数

センス側 3mer		相補側 3mer		センス側 3mer		相補側 3mer	
配列	利用 PCR 件数	配列	利用 PCR 件数	配列	利用 PCR 件数	配列	利用 PCR 件数
ATG	6	CAT	0	ATG	6	CCA	3
TGG	5	CCA	3	TGG	5	TCC	0
GGA	4	TCC	0	GGA	4	ATC	0
GAT	3	ATC	0	GAT	3	CAT	0
ATG	6	CAT	0	ATG	6	ACA	1
TGT	1	ACA	1	TGT	1	CAC	2
GTG	0	CAC	2	GTG	0	CCA	3
TGG	5	CCA	3	TGG	5	GCC	4
GGC	0	GCC	4	GGC	0	TGC	5
GCA	0	TGC	5	GCA	0	ATG	6
CAT	0	ATG	6	CAT	0		

表7 塩基配列分析事例における相互に相補的な 3mer のうち一方の削除

センス側 3mer		相補側 3mer		センス側 3mer		相補側 3mer	
配列	利用 PCR 件数	配列	利用 PCR 件数	配列	利用 PCR 件数	配列	利用 PCR 件数
ATG	6	GAT	0	ATG	6	GCA	3
TGG	5	GCA	3	TGG	5	TCC	0
GGA	4	TCC	0	GGA	4	ATC	0
GAT	3	ATC	0	GAT	3	GAT	0
ATG	6	GAT	0	ATG	6	ACA	1
TGT	1	ACA	1	TGT	1	CAC	2
GTG	0	CAC	2	GTG	0	GCA	3
TGG	5	GCA	3	TGG	5	CCC	4
GGC	0	CCC	4	GGC	0	TGC	5
GCA	0	TGC	5	GCA	0	ATG	6
GAT	0	ATG	6	GAT	0		

表8 塩基配列分析事例における選択された 3mer による仮想 PCR

センス鎖		反対側 3mer の位置および配列		
位置	配列	7	10	11
1	ATG	CAC	TGC	ATG
2	TGG	CAC	TGC	ATG
3	GGA		TGC	ATG
4	GAT		TGC	ATG
5	ATG		TGC	ATG

表9 SARS コロナウイルスへの適用実験に用いたプライマー

レーン	primer (Forward/Reverse)		増幅産物 サイズ(bp)
	番号	配列	
1	160	5'-TCRAGAAATTCAACTCCTGGCAGCA-3'	150
	156	5'-AGGCTTTTTAGATGCCTCAGCAGCA-3'	
2	161	5'-AYAARGTGACWCTYGCTGATGCTGG-3'	550
	157	5'-TRGCAGAAGCYCTGATTTTCAGCAGC-3'	
3	168	5'-GAACTTAGATTCCCTCGAGGCCAGG-3'	390
	169	5'-GAGAATTTCCCCTACTGCTGCCAGG-3'	
4	168	5'-GAACTTAGATTCCCTCGAGGCCAGG-3'	390
	167	5'-GAATTTCCCCTACTGCTGCCAGGAG-3'	
5	164	5'-AGAGCTACCCGACGAGTTCGTGGTG-3'	310
	167	5'-GAATTTCCCCTACTGCTGCCAGGAG-3'	
6	162	5'-AGCTACCCGACGAGTTCGTGGTGGT-3'	310
	167	5'-GAATTTCCCCTACTGCTGCCAGGAG-3'	
7	163	5'-TAGATTCCCTCGAGGCCAGGGCGTT-3'	380
	167	5'-GAATTTCCCCTACTGCTGCCAGGAG-3'	
8	166	5'-GACGAGTTCGTGGTGGTGACGGCAA-3'	230
	170	5'-GAGAAGAGGCTTGACTGCCGCCTCT-3'	
9	168	5'-GAACTTAGATTCCCTCGAGGCCAGG-3'	320
	170	5'-GAGAAGAGGCTTGACTGCCGCCTCT-3'	
10	166	5'-GACGAGTTCGTGGTGGTGACGGCAA-3'	300
	167	5'-GAATTTCCCCTACTGCTGCCAGGAG-3'	
11	164	5'-AGAGCTACCCGACGAGTTCGTGGTG-3'	310
	169	5'-GAGAATTTCCCCTACTGCTGCCAGG-3'	
12	162	5'-AGCTACCCGACGAGTTCGTGGTGGT-3'	310
	169	5'-GAGAATTTCCCCTACTGCTGCCAGG-3'	
13	163	5'-TAGATTCCCTCGAGGCCAGGGCGTT-3'	380
	169	5'-GAGAATTTCCCCTACTGCTGCCAGG-3'	
14	166	5'-GACGAGTTCGTGGTGGTGACGGCAA-3'	300
	165	5'-AGAATTTCCCCTACTGCTGCCAGGA-3'	
15	168	5'-GAACTTAGATTCCCTCGAGGCCAGG-3'	390
	165	5'-AGAATTTCCCCTACTGCTGCCAGGA-3'	
16	164	5'-AGAGCTACCCGACGAGTTCGTGGTG-3'	310
	165	5'-AGAATTTCCCCTACTGCTGCCAGGA-3'	
17	162	5'-AGCTACCCGACGAGTTCGTGGTGGT-3'	310
	165	5'-AGAATTTCCCCTACTGCTGCCAGGA-3'	
18	163	5'-TAGATTCCCTCGAGGCCAGGGCGTT-3'	380
	165	5'-AGAATTTCCCCTACTGCTGCCAGGA-3'	
19	166	5'-GACGAGTTCGTGGTGGTGACGGCAA-3'	300
	169	5'-GAGAATTTCCCCTACTGCTGCCAGG-3'	
20	164	5'-AGAGCTACCCGACGAGTTCGTGGTG-3'	240
	170	5'-GAGAAGAGGCTTGACTGCCGCCTCT-3'	
21	162	5'-AGCTACCCGACGAGTTCGTGGTGGT-3'	240
	170	5'-GAGAAGAGGCTTGACTGCCGCCTCT-3'	
22	166	5'-GACGAGTTCGTGGTGGTGACGGCAA-3'	450
	156	5'-AGGCTTTTTAGATGCCTCAGCAGCA-3'	
23	163	5'-TAGATTCCCTCGAGGCCAGGGCGTT-3'	320
	170	5'-GAGAAGAGGCTTGACTGCCGCCTCT-3'	
24	189	5'-TGAGTTTGACCGTGATGCTGCCATG-3'	930
	191	5'-TCTGGTGTTACAGTRATTGCCTGTC-3'	
25	192	5'-AATTATTRAAGTCAATAGCCGCCAC-3'	260
	192	5'-GGTTTMACATATAGTGAGCCGCCAC-3'	
26	192	5'-AATTATTRAAGTCAATAGCCGCCAC-3'	260
	193	5'-GGTTTMACATATAGTGAGCCGCCAC-3'	
27	196	5'-GCAGARGGRAGCAGAGGCGGCAGTC-3'	210
	156	5'-AGGCTTTTTAGATGCCTCAGCAGCA-3'	
28	194	5'-CGCAGARGGRAGCAGAGGCGGCAGT-3'	210