

機関番号：16401

研究種目：基盤研究（C）

研究期間：2008～2010

課題番号：20570223

研究課題名（和文） 複合体結合能の予測と数理モデルに基づくウイルス適応度ランドスケープの構築

研究課題名（英文） The fitness landscape of a virus inferred based on a mathematical model of the population dynamics

研究代表者

渡部 輝明 (WATABE TERUAKI)

高知大学・教育研究部医療学系・講師

研究者番号：90325415

研究成果の概要(和文): ウイルス変異は免疫回避を促す一方で宿主受容体への結合能力低下(またはその逆)といった対立した効果をもたらすことがある。この宿主内における多面的な変異が如何に宿主集団における免疫回避に影響を与えるかを調べた。ウイルスゲノムに生じた変異と適応進化との因果関係を調べるために数理モデルを開発し、集団遺伝学の理論を用いて変異の宿主集団における固定確率を求めた。高い確率で定着する変異はワクチンがどれくらいの期間有効であるかを決定づけるものである。SARS コロナウイルスの適応進化をシミュレートし、ワクチン接種を受けた宿主集団に高い確率で定着し得る変異が起こり得ることを示した。

研究成果の概要(英文): We examined how pleiotropic mutations that have antagonistic effects (i.e., antibody evasion vs. receptor binding) on viral replication within hosts can impact viral immune escape in the host population. To investigate the causal relation between a mutation in a viral genome and adaptive evolution of a viral population, we developed a mathematical model. Using population genetic theory, we evaluated the probability that a mutant is fixed in a host population. The mutations that can be fixed with high probabilities may determine how long a vaccine remains effective. We simulated the adaptive evolution of coronavirus, the etiological agent of severe acute respiratory syndrome, and showed that some of mutations may have high fixation probabilities in the vaccinated host population.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,700,000	510,000	2,210,000
2009年度	1,000,000	300,000	1,300,000
2010年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,700,000	1,110,000	4,810,000

研究分野：生物学

科研費の分科・細目：生物科学・進化生物学

キーワード：分子進化・タンパク質複合体結合能

1. 研究開始当初の背景

近年多発している新興感染症の病因となるRNAウイルスはヒト以外の動物を自然宿主としている。ウイルスの変異やヒトの活動などが原因となりヒトに偶然感染したことで

ヒトの病原体となったと考えられる。ウイルスがヒトを宿主として適応するためにはヒトの細胞受容体への結合を可能とする必要があり、ヒトへ適応したウイルスは細胞受容体への結合能を維持しつつ生体防御機構か

らの攻撃をかわしていかなければならない。このような感染現象におけるウイルス側とヒト側の分子間での複雑な相互作用を記述するため、正常細胞、感染細胞、ウイルス、そして免疫グロブリンの4者で宿主内動態を表現する方程式系が1996年にNowakらによって提案されている。しかし彼らの方程式系において相互作用の強さを表すパラメータを決定する方法はなく、解析に任意性が残る結果となっていた。本研究では相互作用の強さを分子間の結合親和性と関連させて、ウイルスのタンパク質分子の変異に伴う結合親和性の変化を方程式系のパラメータに反映させる。そしてウイルスの宿主集団内での適応進化を解析する。

2. 研究の目的

感染現象におけるウイルス側とヒト側の分子間での複雑な相互作用を記述する数理モデルを基にウイルスの宿主への適応度を定義し、ウイルス変異の予測を行うことが本研究の目的である。タンパク質構造予測の手法と数理生物学的な手法を用いて、正常細胞、感染細胞、ウイルス、免疫グロブリンの4者からなる宿主生体内でのウイルスライフサイクルを適切に表現した動態方程式を解き、抗原性の変異したウイルスがワクチン接種を受けた宿主集団内で従来型のウイルスと置き換わるための条件を探る。

3. 研究の方法

図1にこれから説明していく研究方法全体の流れを示した。

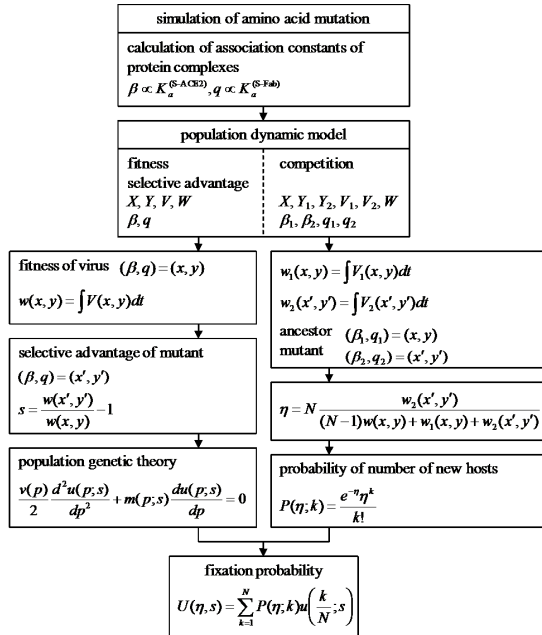


図1.

変異ウイルスは数理モデルにおける正常細胞とウイルス粒子、ウイルス粒子と抗体の相

互作用の強さを決める2つのパラメータに従来型とは異なる値を与えることで導入する。それらパラメータはタンパク質複合体の結合定数に比例することを仮定している。正常細胞とウイルス粒子では受容体分子とウイルス表面のスパイクタンパク質との結合定数に、ウイルス粒子と抗体とではスパイクタンパク質と抗原結合部位との結合定数に比例させる。変異ウイルスの出現はスパイクタンパク質の結合領域におけるアミノ酸配列を変異させることでシミュレートする。アミノ酸配列変異はタンパク質複合体の結合能に影響し、数理モデルにおいてウイルスを特徴づけるパラメータを変化させる。

(1) ウイルス動態の数理モデル

数理モデルは正常細胞 (X)、感染細胞 (Y)、ウイルス (V)、そして免疫グロブリン (W) の4者からなり、以下の微分方程式に従う。

$$\begin{aligned} \dot{X} &= \lambda - dX - \beta XV \\ \dot{Y} &= \beta XV - aY \\ \dot{V} &= kY - uV - \beta XV - qVW \\ \dot{W} &= rV - hW - qVW \end{aligned} \quad \text{式 1.}$$

正常細胞は λ で一定に産生され、 βXV で感染細胞となる。感染細胞は kY でウイルス粒子を産生する。正常細胞、感染細胞、ウイルス粒子の寿命はそれぞれ $1/d$ 、 $1/a$ 、 $1/u$ である。これらのことからウイルスの基本再生産率は $R_0 = \beta \lambda k / (adu)$ で与えられる。また抗体の寿命は $1/h$ で表される。各パラメータはこれまでに発表されている実験事実を基に決定し、式1の連立微分方程式を数値計算の手法を用いて解く。

(2) ウイルス変異を扱うための拡張

変異ウイルスは感染直後に発生すると仮定し、式1の数理モデルを以下のように拡張する。

$$\begin{aligned} \dot{X} &= \lambda - dX - \beta_1 XV_1 \\ \dot{Y}_1 &= \beta_1 XV_1 - aY_1 \\ \dot{V}_1 &= kY_1 - uV_1 - \beta_1 XV_1 - q_1 V_1 W \\ \dot{W} &= rV_1 - hW - q_1 V_1 W \end{aligned} \quad \text{式 2.}$$

$$\begin{aligned} \dot{X} &= \lambda - dX - \beta_1 XV_1 - \beta_2 XV_2 \\ \dot{Y}_i &= \beta_i XV_i - aY_i \\ \dot{V}_i &= kY_i - uV_i - \beta_i XV_i - q_i V_i W \\ \dot{W} &= rV_1 + rV_2 - hW - q_1 V_1 W - q_2 V_2 W \end{aligned} \quad \text{式 3.}$$

式3.

変異ウイルス出現まで ($0 \leq t < \hat{t}$)、拡張さ

れた方程式 (式 2) は添え字を除いて拡張前の方程式と同じである。変異ウイルスの発生時間 \hat{t} をウイルス粒子の寿命程度 ($1/u$) にとり、変異ウイルス発生以後 ($t \geq \hat{t}$) の動態を式 3 の連立微分方程式で記述する。添え字 i は 1 か 2 をとる。変異ウイルスは発生に続き宿主細胞に感染し、感染した細胞 (Y_2) は変異ウイルス粒子 (V_2) を産生する。

(3) 宿主内での適応度

式 1 でパラメータの組 $(\beta, q) = (x, y)$ で特徴づけられるウイルス集団について、その宿主内での適応度を変数 V の時間積分によって定義する。

$$\text{式 4. } w(x, y) = \int_0^{\tau} V(x, y; t) dt$$

宿主間の伝染力は宿主内のウイルス量に直接影響されていることを仮定している。感染宿主が医療機関による医療処置を受けるまで ($t < \tau$) の期間で時間積分を行い有効なウイルス量とした。パラメータ (β, q) による 2 次元平面上で、ウイルス適応度ランドスケープは w の値の変化で与えられる。高い値の w を持つウイルスは新たな宿主へ効率よく感染出来る。

(4) 変異ウイルスの宿主集団への定着

変異ウイルスの宿主集団への固定確率を計算するためには選択的優位性と初期頻度を定義する必要がある。選択的優位性は変異ウイルスと従来型の適応度の比で定義される。

$$\text{式 5. } s = \frac{w(x', y')}{w(x, y)} - 1$$

ここで $(\beta, q) = (x', y')$ は変異ウイルスを特徴づけている。頻度が p で与えられる変異ウイルスの固定確率は以下の微分方程式を境界条件 $u(0) = 0$ 及び $u(1) = 1$ のもとで解くことにより与えられる。

$$\text{式 6. } \frac{v(p)}{2} \frac{d^2 u(p)}{dp^2} + m(p) \frac{du(p)}{dp} = 0$$

$m(p)$ と $v(p)$ はそれぞれ以下の様に与えられる。

$$\text{式 7. } \begin{aligned} m(p) &= \frac{sp(1-p)}{1+sp} \\ v(p) &= \frac{p(1-p)}{N} \end{aligned}$$

N は宿主集団の規模を表している。全ての宿主はワクチン接種を受けているとし、変異ウイルスが出現するまでは従来型のウイルスに感染しているものとする。宿主個々が感染と治癒を経て随時入れ替わっていると

も、変異ウイルスが宿主集団に定着するまでは宿主集団の規模は一定であると仮定している。変異ウイルスの固定確率は以下の様に得られる。

$$\text{式 8. } u(p) \equiv u(p; s) = \frac{1 - e^{-(2N-1)\log(1+sp)}}{1 - e^{-(2N-1)\log(1+s)}}$$

(5) 変異ウイルスの初期頻度

変異ウイルスは宿主内で従来型のウイルスと競合する関係にある。この競り合いの結果によって変異ウイルスが新たな宿主へ感染し得るかどうかが決まる。式 3 において従来型ウイルス (V_1) と変異ウイルス (V_2) についての時間積分を考える。

$$w_1(x, y) = \int_0^{\tau} V_1(x, y; t) dt$$

式 9.

$$w_2(x', y') = \int_0^{\tau} V_2(x', y'; t) dt$$

変異ウイルスが出現した宿主以外の宿主では従来型のウイルスのみが存在している。変異ウイルスの頻度は w_2 と全ての宿主でのウイルス量の総和との比で表される。このことから次の世代で変異ウイルスに感染する宿主の数 (k) の分布がポアソン分布を用いて以下の様に与えられる。

$$P(\eta; k) = \frac{e^{-\eta} \eta^k}{k!}$$

$$\eta = N \frac{w_2(x', y')}{(N-1)w(x, y) + w_1(x, y) + w_2(x', y')}$$

式 10.

式 8 と式 10 を組み合わせると、変異ウイルスの固定確率が以下の様に得られる。

$$\text{式 11. } U(\eta, s) = \sum_{k=1}^N P(\eta; k) u\left(\frac{k}{N}; s\right)$$

(6) タンパク質複合体結合能の指標

タンパク質 **A** と **B** で構成される複合体の結合能は結合定数で測られる。

$$\text{式 12. } K_a = \frac{[A \cdot B]}{[A][B]}$$

ここで $[A]$ 、 $[B]$ と $[A \cdot B]$ はそれぞれ平衡状態にあるタンパク質 **A**、**B** とそれらの複合体の濃度である。この結合定数はそれぞれのタンパク質の尤度から構成される尤度比と比例関係にあることを示すことが出来る。

$$\text{式 13. } K_a \propto \frac{P(A \cdot B)}{P(A)P(B)}$$

アミノ酸配列の変異に対して主鎖構造は安定であると仮定して、式 13 の尤度比を構造

が与えられたときの配列の分布として書き直すことが可能である。

$$K_a \propto \frac{P(\text{seq}_A, \text{seq}_B | \text{str}_{A+B})}{P(\text{seq}_A | \text{str}_A)P(\text{seq}_B | \text{str}_B)}$$

式 1 4.

式 1 4 の右辺は Protein Data Bank より得られる構造データをもとに申請者等が以前開発した配列-構造適合度を用いて計算することが出来る。

(7) パラメータ設定と構造データ

ワクチン接種者を 100 万人規模で想定し、その内の 1% が感染しているとする ($N = 10000$)。正常細胞の寿命を 100 日 ($1/d = 100$ days) とし、抗体の寿命は 20 日 ($1/h = 20$ days) としている。ウイルス粒子の寿命はウイルスにより大きく異なるがここでは 2 時間 ($1/u = 2$ h) とし、感染細胞については 2 日 ($1/a = 2$ days) としている。

SARS コロナウイルスのスパイクタンパク質と宿主細胞受容体 (ACE2) の複合体構造データは X 線結晶構造解析により測定されており、Protein Data Bank で公開されている (PDB code: 2AJF)。同様に抗体との複合体も 2 種類公開されており、抗体 80R との複合体 (PDB code: 2GHW) と抗体 m396 との複合体 (PDB code: 2DD8) を用いる。

4. 研究成果

(1) 宿主内動態

図 2 A にウイルスの宿主内動態を示している。

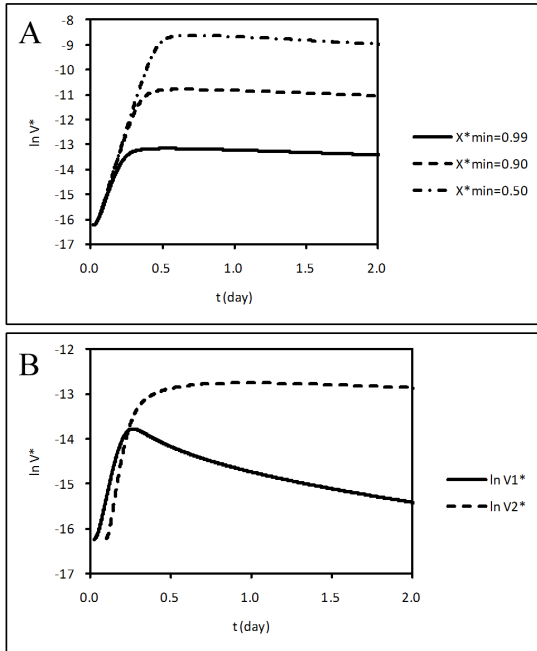


図 2.

基本再生産率が $R_0 = 100$ であり、パラメー

タ q について 3 つの値を適用している。それぞれの値で正常細胞の最低量が $Xd/\lambda|_{\min} = 0.99, 0.90, 0.50$ のように決まっている。 Xd/λ で示される量は、ウイルス感染が無い状態での正常細胞量に対する比に相当している。感染初期ではウイルス量は対数尺度で線形に増加しており、抗体の中和能力 (q) には依存していない。図 2 B に変異ウイルスと従来型とが競り合っている様子が示されている。従来型の基本再生産率が $R_0 = 100$ で変異ウイルスでは $R_0 = 200$ の場合に相当し、パラメータ q は図 2 A において正常細胞の最低量が $Xd/\lambda|_{\min} = 0.99$ の場合で示している。変異ウイルスが従来型を抑制している様子がわかる。

(2) 宿主内での適応度ランドスケープ

式 1 で示されている数理モデルを解き、適応度ランドスケープを得た (下図、抗体 80R との場合を示した)。

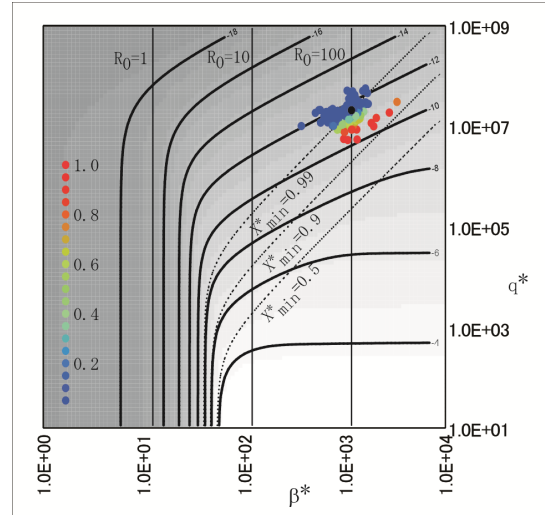


図 3.

得られた適応度ランドスケープは数理モデルに含まれるパラメータの値によらず共通の特徴を持っている。適応度は変数 V の時間積分 (式 4) で定義されているが、感染期間を考慮して時間 τ で打ち切っている (ここでは $\tau = 10$ days)。更に低い基本再生産率 ($R_0 \sim 1$) ではウイルス動態は感染初期において抗体の中和能力には依存していない。そのため基本再生産率が低い領域 (β の値の小さい領域) では適応度ランドスケープは抗体の中和能力 (q の値) には依存していない。

選択強度の分布を観るため任意に 1000 通りの 1 残基置換を起こした。それぞれの適応度ランドスケープ上での位置と宿主集団への固定確率を色で図 3 に示した。基本再生産率が $R_0 = 100$ で正常細胞の最低量が $Xd/\lambda|_{\min} = 0.99$ となる位置を従来型としてシミュレートしている。幾つかの変異は 1 に近い固定確率を示している。

(3) 適応進化に寄与する主なアミノ酸残基

図4に宿主集団に高い確率で定着し得る変異を起こし得るアミノ酸残基を示した。

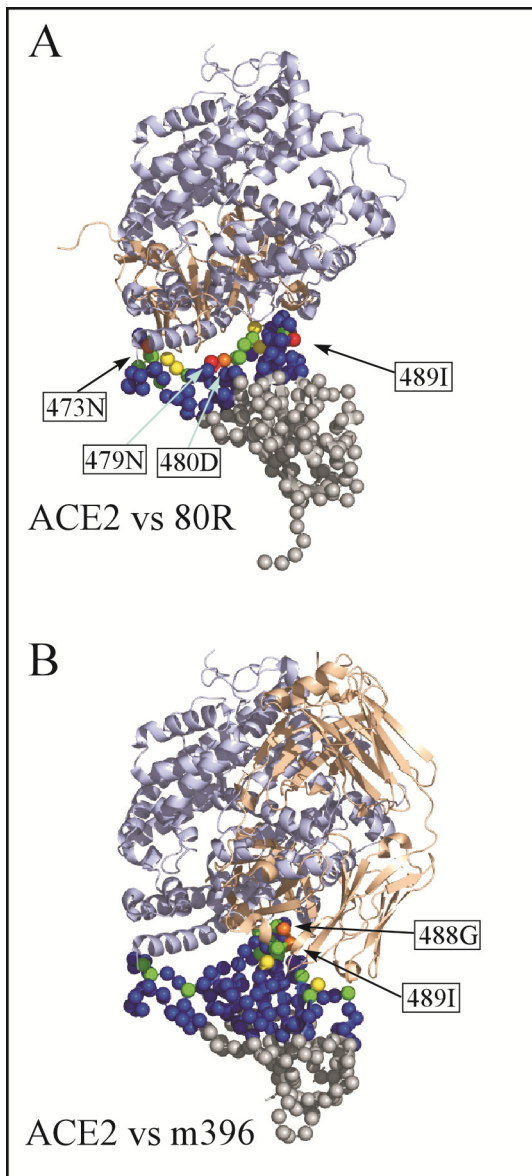


図4.

色付きの球は該当するアミノ酸残基の α 炭素原子の位置を示している。各残基で20通りの変異(同じアミノ酸への変異も含む)における固定確率の平均を色で示している。赤は平均固定確率が0.6以上0.8未満、橙は0.4以上0.6未満、黄は0.2以上0.4未満、緑は0.1以上0.2未満、青が0.1未満を示している。シミュレーションにおける従来型は図3の場合と同様である。図4Aは抗体80Rの場合を示しており、ベージュ色で示した構造が抗体80Rの抗原結合部位である。薄紫色で示した宿主細胞受容体ACE2の構造を重ね合わせて示している。図4Bは抗体m396との

場合を示している。

赤や橙で示している残基では高い固定確率を持つ変異が起こり得る状況にあり、幾つかの変異では0.9を超える固定確率を示している。これらの残基における突然変異率を考慮することで、宿主集団に定着する変異がどの程度の時間的な尺度で起こり得るのかを見積もることが出来る。そのことはワクチン開発において有用な知見を与えるものと考えられる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計1件)

- ① Teruaki Watabe and Hirohisa Kishino, Structural considerations in the fitness landscape of a virus, *Molecular Biology and Evolution*, 査読有り, 27, 2010, 1782-1791

[学会発表] (計6件)

- ① 渡部輝明、タンパク質にかかる多様化圧の空間分布とベイズ推定、統計関連学会連合大会、2010年9月7日、早稲田大学(東京都)
- ② 渡部輝明、局所尤度と階層ベイズによる多様化圧の空間マッピング、日本進化学会、2010年8月3日、東京工業大学(東京都)
- ③ 渡部輝明、タンパク質共進化と分子進化速度の時空間変動、第10回日本蛋白質科学会年会、2010年6月16日、札幌コンベンションセンター(札幌市)
- ④ Teruaki Watabe, Structural consideration on the fitness-landscape of virus, Annual Meeting of the Society for Molecular Biology and Evolution (SMBE 2009), 2009年6月4日, The University of Iowa (Iowa City, IA, U.S.A.)
- ⑤ 渡部輝明、構造データベースに基づく複合体結合能の予測とウイルス適応度マップの構成、統計関連学会連合大会、2008年9月9日、慶応大学(日吉)
- ⑥ 渡部輝明、構造データベースに基づく複合体結合能の予測とウイルス適応度マップの構成、日本進化学会、2008年8月22日、東京大学(駒場)

6. 研究組織

(1) 研究代表者

渡部 輝明 (WATABE TERUAKI)

高知大学・教育研究部医療学系・講師

研究者番号：90325415

(2)研究分担者

岸野 洋久 (KISHINO HIROHISA)

東京大学・農学生命科学研究科・教授

研究者番号：00141987