

インフルエンザウイルスの進化における共変異の変化の解析

谷口剛¹ 伊藤公人² 五十嵐学² 村上悌治² 高田礼人² 原口誠¹

北海道大学, 大学院情報科学研究科¹ 人獣共通感染症リサーチセンター²

概要

インフルエンザウイルスにおける抗原変異の規則性を発見するために、アミノ酸残基の共変異を解析する。共変異とは、タンパク質を構成するアミノ酸残基のうち、複数の位置のアミノ酸が共に置換する現象である。従来からアミノ酸残基の共変異を解析する手法がいくつか提案されていたが、それらの手法では進化の過程における分岐や時間的關係が考慮されていなかった。そこで、これらの問題を解決するために、進化系統解析によって得られる系統樹を利用する手法を提案する。過去40年間のH3N2亜型インフルエンザウイルスのHAタンパク質を対象とし、共変異の検出を行い、その結果を示す。また、共変異は時代と共に変化するため、共変異の変化を検出するための手法を提案する。

An Analysis of Correlated Mutations in the Hemagglutinin of Influenza Viruses

Tsuyoshi TANIGUCHI¹, Kimihito ITO², Manabu IGARASHI², Teiji MURAKAMI²,
Ayato TAKADA², Makoto HARAGUCHI¹

Graduate School of Information Science and Technology¹, Hokkaido University,
Research Center for Zoonosis Control², Hokkaido University

Abstract

The influenza viruses undergo antigenic drift to escape from antibody-mediated immune pressure. In order to predict possible structural changes of their molecules in future, it is important to analyze the patterns of amino acid substitutions in the past. In this paper, we present a new method to extract the sets of residue positions which were involved in correlated mutations. We also discuss a method to detect changes of correlation among co-evolving residues.

1 はじめに

毎年世界中でインフルエンザが流行し、発熱、急性肺炎等の重篤な疾病を引き起こしている。インフルエンザウイルスの遺伝子は突然変異を起こしやすく、毎年ごく僅かな変異ウイルスが人の免疫システムから逃れて生き残り、その翌年、抗原性が少し異なる変異ウイルスとして流行を繰り返す。一般に、ウイルス感染症の予防には、薬品によって不活化したウイルスをワクチンとして注射する手法が有効である。しかし、インフルエンザの場合ウイルスの抗原性が変化し続けるため、ワクチンによる効果が長続きしない。タミフル等の抗ウイルス剤も使用されるが、薬剤耐性ウイルスへの変異も報告されている。これらの理由から、決定的な予防法の開発は難しいとされている。

インフルエンザウイルスは、抗原変異によって流行を繰り返す。一般に、ウイルスに感染した個体の体内では、ウイルスの抗原領域を特異的に認識する抗体が産生される。抗体はウイルス表

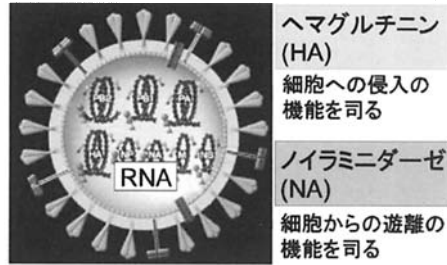


図 1: インフルエンザウイルスの構造模式図

面のタンパク質に結合することで、ウイルスが細胞に侵入する機能を阻害して、細胞でウイルスが複製されることを抑制する。一方、集団においてウイルスは咳やくしゃみによって別の個体に感染する。この感染を繰り返すうちに、突然変異によって、抗体が認識する部位のアミノ酸が偶然別のアミノ酸に置き換わったウイルスが現れる。この変異ウイルスが以前の感染時に産生された抗体によって病原体として認識されなければ、一度インフルエンザに罹ったことがある個体にもまた感染する。このように、ウイルスが抗体との結合から逃れていくためのアミノ酸の変化を抗原変異と呼ぶ。

インフルエンザウイルスは赤血球凝集素 (HA) とノイラミニダーゼ (NA) の二種類のスパイクタンパク質を粒子表面に持つ (図 1)。HA および NA はウイルスの表面に突出しているため、抗体による免疫システムの標的となりやすい。インフルエンザの抗原変異は、HA および NA 上のアミノ酸が突然変異によって別のアミノ酸に置換され、タンパク質の部分的構造が変化し、抗体によって認識されなくなることに起因する。HA と NA はそれぞれ宿主細胞への侵入と細胞からの遊離を司る。HA と NA はこれらの機能を保持する必要がある、抗原変異におけるアミノ酸置換には何らかの制限があると考えられる。そのため、インフルエンザウイルスの抗原変異におけるアミノ酸置換に、ある種の規則性が潜在する可能性がある。

本研究では、インフルエンザウイルスの抗原変異における規則性を解析するために、複数の残基位置のアミノ酸が共に置換する「共変異」の現象に着目する。インフルエンザウイルスのゲノムは 1980 年代に解読されている。現在、NCBI のデータベースには約 3 万本の塩基配列が登録され、異宿主動物間のウイルス伝播経路の解明や、過去の抗原変異における抗原決定部位の推定に活用されている。しかし、これらの遺伝子情報に基づき、インフルエンザウイルスのアミノ酸残基の共変異を解析した研究は少ない。Ghedini らは、1998 年から 2004 年までの間に分離された H3N2 亜型のインフルエンザウイルス 207 株のアミノ酸配列を比較し、いくつかの共変異を報告している [3]。また、本稿の共著者らは、1968 年から 2005 年に分離された 2183 株のインフルエンザウイルスを対象に、HA における共変異を解析している [4]。

アミノ酸残基の共変異を解析するために、従来では、相互情報量 [5] や、相互情報量を正規化した指標 [6] を用いた手法が提案されている。これらの解析では、アミノ酸配列の多重配列アライメントを行い、アライメントにおける n 番目のアミノ酸残基を、二十種類のアミノ酸 A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V を値として持つ確率変数 X_n として表す。その上で、確率変数 X_i と X_j との関係解析し、 i 番目と j 番目のアミノ酸残基の共変異を抽出する。しかし、これらのアプローチでは、進化の過程におけるインフルエンザウイルス株の分岐や時間的關係が考慮されていない。

本研究では、進化の過程における分岐と祖先関係を考慮して共変異を解析するために、進化系統解析を応用する手法を提案する。インフルエンザウイルスの遺伝子データのように、祖先のアミノ酸配列が明らかにされている場合、進化系統解析によって推定される系統樹を利用すると、アミノ酸配列間の祖先関係が明らかになる。また、二つのアミノ酸配列において、同一位置のアミ

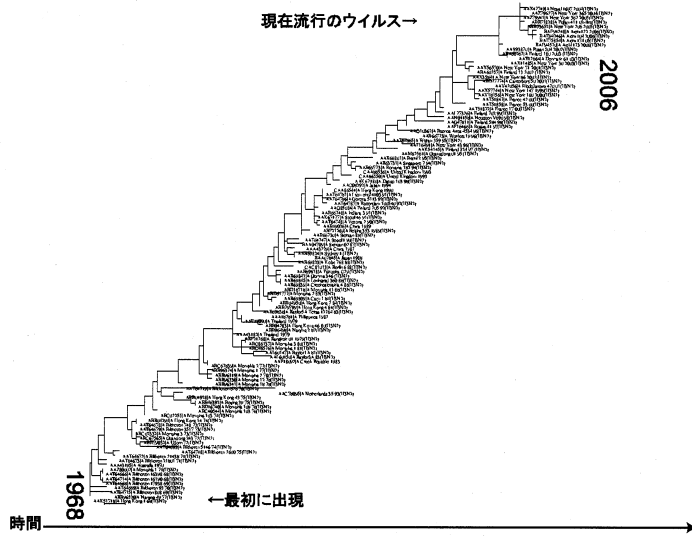


図 2: インフルエンザウイルス (H3N2 亜型) の進化系統樹

ノ酸が異なっている場合に、どちらのアミノ酸が祖先型であるかが明らかになる。進化系統解析を応用することにより、進化過程における分岐や時間的關係を考慮した共変異の解析が実現できると考えられる。

過去 40 年間にヒトから分離された H3N2 亜型のインフルエンザウイルスの HA タンパク質アミノ酸配列に対して、相互情報量を用いて共変異の検出を行った [4]。その結果、検出された共変異における残基位置間の関係は一様でないことが判った。ウイルスタンパク質の三次元立体構造の複雑な相互作用のために、あるときには共変異していた残基位置同士がその後も同様な共変異を示すとは限らず、今まで全く共変異をしたことがなかった残基位置同士が共変異を示している可能性がある。

また、インフルエンザウイルスのそれぞれの残基位置において、アミノ酸は一様に置換するわけではない。アミノ酸が頻繁に置換される位置があれば、ほとんど置換せず同じアミノ酸である残基位置もある。これらを考慮すると、他の残基位置間の共変異に影響を与えるような残基位置が存在する可能性がある。そこで本研究では、共変異の変化を検出する手法を提案する。

2 進化系統樹を利用して共変異を検出する手法

本節では進化系統解析を利用し、アミノ酸残基位置での共変異を検出する手法を提案する。図 2 に H3N2 亜型インフルエンザウイルスの進化系統樹を示す。

一般のタンパク質の進化過程において、二つ以上の異なる位置のアミノ酸が共に変異する場合は知られている。二つ以上のアミノ酸が共に置換する進化過程を、アミノ酸残基の共変異と呼ぶ。文献によっては、アミノ酸残基の共変異を、“correlated-mutation”、“co-mutation”、“co-evolution”、“co-variation”、“co-ordinated amino-acid substitution”等の用語で呼ぶことがある。アミノ酸残基の共変異は、タンパク質の構造や機能の保持に重要な突然変異であると考えられている。一つの自然な解釈として、ある位置でのアミノ酸変異がタンパク質の構造上の歪みを引き起こし、別の位置のアミノ酸変異がこの歪みを修正する現象であることが挙げられる。

以下に、アミノ酸残基の共変異の例を挙げる。 n 番目の残基位置における、アミノ酸 x から y への置換を、 xny と記述する。

Example 1 下記の5つのアミノ酸配列 $seq_{1968}, \dots, seq_{2000}$ を考える。添え字は、ウイルス株が分離された年を表す。

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|--------------|---|---|---|---|---|---|---|---|---|
| seq_{1968} | Q | D | L | P | G | N | R | N | S |
| seq_{1976} | Q | D | I | P | G | A | R | N | S |
| seq_{1984} | Q | K | L | P | G | T | D | N | S |
| seq_{1992} | Q | K | I | P | G | Q | D | N | S |
| seq_{2000} | Q | K | L | P | G | S | D | N | T |

2番目のアミノ酸残基位置に注目すると、1976年から1984年の間に、アミノ酸がD(アスパラギン酸)からK(リシン)に置き換わっている。上述の記法に従いこの変異をD2Kと記述する。

また、この期間には、アミノ酸置換D2Kと同時にR7Dが共起している。したがって、2番目と7番目のアミノ酸残基が共変異していると見なすことが出来る。なお、2番目と7番目の両アミノ酸残基位置においては、この期間以外にアミノ酸置換は見られない。

2.1 進化系統樹からのアミノ酸置換履歴の抽出

インフルエンザウイルス進化の過程における共変異を検出するために、HAタンパク質アミノ酸配列から、アミノ酸置換履歴を抽出する。

1. アミノ酸配列の多重配列アライメントを行い、進化系統樹を推定する。
2. 進化系統樹からアミノ酸置換の集合を抽出する。

上記1では、ヒトから分離されたH3N2亜型インフルエンザウイルスのHAおよびNAのアミノ酸配列を、公共データベースより大量にダウンロードし、進化系統解析を行いアミノ酸配列の祖先関係を推定する。一般に、進化系統樹の作成には、近隣結合法が多く用いられる。本研究では、仮想的な祖先(進化系統樹の葉以外のノード)のアミノ酸配列を推定するために、進化系統解析には最節約法を用いる。進化系統樹の葉以外のノードを内点と呼ぶ。仮想的な祖先は内点となる。

進化系統樹のデータ構造は、木構造である。最節約法および最尤法を用いて作成される進化系統解析では、内点のアミノ酸配列が推定されるため、木の全てのノードにはアミノ酸配列が対応付けられる。つまり、各ノードはウイルス株および仮想的な祖先ウイルスのアミノ酸配列に対応する。根は、新型インフルエンザとして出現した株である。幹はインフルエンザウイルスの過去の抗原変異の履歴を表す。幹の末端は、現在流行しているウイルス株である。枝は、突然変異によるアミノ酸置換に対応する。各ノードにはアミノ酸配列が対応付けられているので、枝には具体的なアミノ酸置換の集合を対応付けることが可能である。

上で議論したように、最節約法あるいは最尤法で作成された進化系統樹を用いると、進化系統樹の各枝に対応する、具体的なアミノ酸置換の集合を抽出する事ができる。

図3に進化系統樹からアミノ酸置換履歴を抽出する手法の概要を示す。図3左の進化系統樹には、根・葉・内点を含む各ノードに、アミノ酸配列が割り当てられている。根ノードと、根ノードの左の子ノードに割り当てられたアミノ酸は、それぞれQDNRQS...とQGNRIS...である。これらのアミノ酸配列の差を求めることにより、これらのノードを結ぶ枝には、アミノ酸置換の集合{D2G, Q5I}を割り当てる。進化系統樹上の各枝に割り当てられた全てのアミノ酸置換の集合は、過去の同時期に起こったアミノ酸置換の集合と解釈することが出来る。

2.2 アミノ酸置換履歴からの共変異の検出

前副節では、最節約法によって作成した進化系統樹から同時期に起こったアミノ酸置換の集合を抽出できることを示した。本副節では、これらのアミノ酸置換の集合から、共変異を検出する

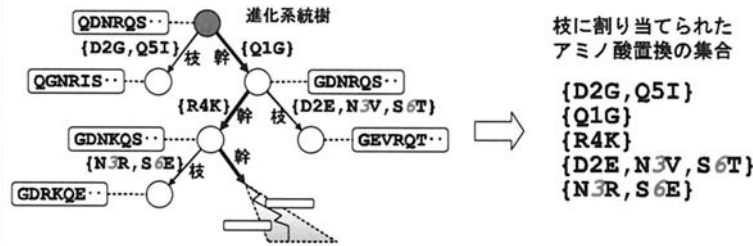


図 3: 進化系統樹からのアミノ酸置換履歴の抽出

手法について議論する。

アミノ酸の集合 A を $\{A, R, N, D, C, Q, E, G, H, I, L, K, M, F, P, S, T, W, Y, V\}$ で表す。進化系統樹上のある枝 e に割り当てられたアミノ酸置換の集合 T_e を

$$T_e = \{x_1 n_1 y_1, x_2 n_2 y_2, \dots, x_k n_k y_k\},$$

とする。ここで、 $x_1, \dots, x_k, y_1, \dots, y_k$ はアミノ酸であり、 n_1, \dots, n_k は、アミノ酸残基位置である。このとき、枝 e においてアミノ酸置換が起こったアミノ酸残基位置の集合 S_e を、

$$S_e = \{n_1, n_2, \dots, n_k\},$$

とする。

進化系統樹に現れる各枝 e 対し、 S_e に含まれる残基位置の集合は、同時期にアミノ酸置換が起こった残基位置と見なすことができる。進化系統樹の全ての枝に対してそれぞれ S_e を求め、集合 S_e の集合に頻出するアミノ酸残基位置の集合 n_1, \dots, n_k を求める。求まる集合は、同時期にアミノ酸置換が起こる傾向にあるアミノ酸残基位置の集合である。すなわち、集合 S_e の集合に頻出するアミノ酸残基位置の抽出は、共変異の検出に他ならない。

アイテム集合の列中に、同時に現れる傾向にあるアイテムの集合を抽出する問題は、頻出アイテム集合マイニングと呼ばれ、近年データマイニングの研究分野で盛んに研究されている [1]。LCM [12] などの、高速なアルゴリズムも開発されている。次の副節では、インフルエンザウイルスの進化系統樹から抽出された、アミノ酸置換の集合における頻出アイテム集合を求め、アミノ酸残基の共変異を検出する。

2.3 インフルエンザウイルスの進化系統樹における共変異の検出

香港風邪の病原体ウイルスである H3N2 亜型インフルエンザウイルスの HA タンパク質における共変異を、頻出アイテム集合マイニングを用いて検出する実験を行った。

HA タンパク質のアミノ酸配列を、NCBI Influenza Virus Resource [7] からダウンロードし、多重配列アライメントによって、長さの異なるアミノ酸配列を整列した。H3N2 亜型ウイルスの HA タンパクのアミノ酸配列の長さは約 550 であり、約 330 の長さの HA1、約 220 の長さの HA2 の二つの領域をもつ。抗原変異は、1) 主に HA1 領域で起こること、2) HA1 領域のみデータベースに登録されている配列が多数あることから、本実験では、多重配列アライメントから切り出した HA1 領域のアミノ酸配列 2183 本を解析対象とした。

HA1 領域のアミノ酸配列 2183 本から、最節約法を用いて進化系統樹を作成した。表 1 に作成された進化系統樹の概要を示す。

表 1 の進化系統樹を用いて、2183 株のウイルス株それぞれに対し、2 年以上離れた祖先株を割り出した。その上で、祖先株から、各子孫株へのアミノ酸置換が起きている残基位置の集合を求めた。配列アミノ酸配列が全く等しい子孫株を除外し、同時期アミノ酸置換が起こった残基位置

表 1: 作成された進化系統樹の概要

| 葉の数 | 枝の数 | アミノ酸置換の総数 | 幹の長さ | 幹上の置換の総数 |
|------|------|-----------|------|----------|
| 2183 | 4364 | 2182 | 286 | 242 |

表 2: 検出された共変異の個数

| 共変異残基位置の数 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 計 |
|---------------------|----|----|---|---|---|---|---|---|----|----|----|----|
| 頻出アイテム (残基位置) 集合の個数 | 27 | 19 | 6 | 6 | 4 | 2 | 0 | 0 | 0 | 0 | 1 | 65 |

を含む集合 1062 個を分離した。これらの集合から、頻出アイテム集合を求めるアプリアリアルゴリズム [1] により、アミノ酸の共変異が起こったアミノ酸残基位置の集合を計算した。

インフルエンザウイルスの HA のアミノ酸配列において、25 回以上共にアミノ酸置換が起きた残基位置の集合 60 種類が検出された。表 2 に検出された共変異の概要を示す。

検出された共変異の一例として、{172, 193}, {144, 276}, {62, 133, 158}, {135, 190, 226}, {121, 124, 275, 278}, {25, 50, 75, 83, 106, 131, 155, 156, 202, 222, 225, 261} がある。現在、これらのアミノ酸残基位置に関して、ウイルス学的な意味の考察を行っている。

最小サポート値を 25 回とし、アプリアリアルゴリズムを用いて計算したところ計 4768 種類の頻出アイテム集合が見つかった。頻出アイテム集合の性質から、計算結果には包含関係にある集合が重複されて数えられている。たとえば、アミノ酸残基位置の集合 {62, 133, 158} が頻出アイテム集合である場合には、この部分集合である {62, 133}, {62, 158}, {133, 158} も同様に頻出アイテム集合である。表 2 に示した共変異では、このような自明なアイテム集合を除外して数えた。

3 共変異の差異の検出

2 節において進化系統樹を利用し共変異を検出した。この節ではその抽出された共変異を解析する手法として共変異の変化を検出するための手法について議論する。

3.1 準備

この副節では、アイテムおよびアイテム集合の定義を行う。

本論文においては、インフルエンザウイルスの HA1 領域のアミノ酸配列を扱うため、アミノ酸残基の位置の集合 $I = \{i_1, i_2, \dots, i_{328}\}$ をアイテムの集合とする。このとき、 $S \subseteq I$ をアイテム集合と呼ぶ。

本研究では、タイムスタンプ t とアイテム集合 S の組 (t, S) をトランザクションとして持つ以下のようなデータベースを解析の対象とする。

$$D = \{(t_1, S_1), (t_2, S_2), \dots, (t_n, S_n)\}.$$

ここでデータベース D に対するアイテム集合 S の出現 $O(S, D)$ を次のように定義する。

$$O(S, D) = \{d \mid S \subseteq d \wedge d \in D\}.$$

そのときに、データベース D に対するアイテム集合 S の出現の確率 Pr を

$$Pr(O(S, D)) = |O(S, D)| / |D|$$

と定義する。以後の議論において、 $Pr(O(S, D))$ は簡単に $P(S)$ と記述する。

また、本研究ではある期間における共変異も評価の対象とするため、以下のようなローカルデータベースを定義する。

$$D_t = \{d \mid d \text{ のタイムスタンプ } t \text{ が, } t_{\text{before}} \leq t \leq t_{\text{after}} \wedge d \in D\}.$$

ここで, $t_{\text{before}}, t_{\text{after}}$ はユーザによって与えられているとする. D と同様に, データベース D_t に対するアイテム集合 S の出現 $O(S, D_t)$ は

$$O(S, D_t) = \{d \mid S \subseteq d \wedge d \in D_t\}$$

と定義する. さらに, データベース D_t に対するアイテム集合 S の出現の確率 Pr_t を

$$Pr_t(O(S, D_t)) = |O(S, D_t)| / |D_t|$$

とし, 簡単に $P_t(S)$ と記述する.

3.2 共変異の変化の検出

この副節では, 共変異の変化を検出するための手法について議論する. 共変異は複数のアミノ酸残基位置が共に置換する現象なので, いくつかの株において共起するアイテム (アミノ酸置換) の集合とみなすことができる. したがって共変異の度合いが高い場合, 共起の割合はそれぞれ単独の置換に対して高いだろうし, 度合いが低い場合, 共起の割合は低いであろう. ここで, 異なる残基位置間の共変異を評価するため, \log をとらない自己相互情報量である以下の式を用いる.

$$\begin{aligned} Cor(S_n, S_m) &= Pr(O(S_n, D) \cap O(S_m, D)) / Pr(O(S_n, D)) Pr(O(S_m, D)) \\ &= Pr(O(S_n \cup S_m, D)) / Pr(O(S_n, D)) Pr(O(S_m, D)) \\ &= P(S_n \cup S_m) / P(S_n) P(S_m). \end{aligned}$$

S_n と S_m の共変異が独立に起こっている場合には, $P(S_n \cup S_m) = P(S_n)P(S_m)$ となる. よって, $Cor(S_n, S_m) > 1$ のときには S_n と S_m の共変異は正の相関であると考えることができる. 一方, $Cor(S_n, S_m) < 1$ の場合, S_n と S_m の共変異は負の相関である.

同様に, あるタイムスタンプがついた株もしくは, 進化系統樹の内点における共変異の度合いは以下の式で評価することができる.

$$\begin{aligned} Cor(S_n, S_m; t) &= Pr_t(O(S_n, D_t) \cap O(S_m, D_t)) / Pr_t(O(S_n, D_t)) Pr_t(O(S_m, D_t)) \\ &= P_t(S_n \cup S_m) / P_t(S_n) P_t(S_m). \end{aligned}$$

本研究における目的は, 共変異の変化をとらえることである. そのための一つの方法として共変異の度合いの変化を評価することを考える. 上記で定義した式の比をとって共変異の変化を評価するための評価関数とする.

$$Change(S_n, S_m; t) = Cor(S_n, S_m; t) / Cor(S_n, S_m).$$

$Change(S_n, S_m; t) > 1$ であるとき, ある期間 t において残基位置の集合 S_n と S_m の共変異の度合いが高くなっていることを意味し, $Change(S_n, S_m; t) < 1$ である場合は, 共変異の度合いが低くなっていることを意味する. また値が 1 から離れているほど, その変化は顕著である. 本研究では, パラメータ $\rho_{upper} (> 1)$, $\rho_{lower} (< 1)$ を導入し, $Change(S_n, S_m; t) > \rho_{upper}$ もしくは $Change(S_n, S_m; t) < \rho_{lower}$ となるようなアイテム集合 S_n , S_m を発見することにより, 共変異の変化が顕著である残基位置間の関係を検出することが可能であると考えられる. 本稿の第一著者らは相関が変化するアイテム集合を効率的に求めるアルゴリズムを提案しており [11], 本副節における解析にも利用できる. 現在, そのアルゴリズムを利用し, 本副節において議論した解析を行っている.

4 おわりに

本研究では、インフルエンザウイルスの抗原変異における規則性を発見するために、複数の残基位置のアミノ酸が共に置換する「共変異」の現象に着目した。共変異とは、タンパク質を構成するアミノ酸残基のうち、複数の位置のアミノ酸が共に置換する現象である。従来からアミノ酸残基の共変異を解析する手法がいくつか提案されていたが、それらの手法では進化の過程における分岐や時間的關係が考慮されていなかった。そこで、これらの問題を解決するために、進化系統解析によって得られる系統樹を利用する手法を提案した。過去40年間のH3N2亜型インフルエンザウイルスのHAタンパク質を対象として解析を行った。その結果、アミノ酸残基位置の共変異60種類を検出した。検出された共変異の多くは、従来手法では検出されない共変異であった。

さらに、共変異の変化を検出する手法について議論した。ウイルスタンパク質の三次元立体構造の複雑な相互作用のために、あるときには共変異していた残基位置同士がその後も同様な共変異を示すとは限らず、今まで全く共変異をしたことがなかった残基位置同士が共変異を示している可能性が大きい。また、他の残基位置間の共変異に影響を与えるような残基位置が存在する可能性もある。共変異の変化に影響を与えるような残基位置の存在を検出する手法については、現在、考察中である。

参考文献

- [1] Agrawal,R. and Srikant,R. (1994) Fast Algorithms for Mining Association Rules in Large Databases, In Proceedings of the 20th International Conference on Very Large Data Bases, Morgan Kaufmann,VLDB'94:487-499.
- [2] Bush, R.M., Bender,C.A., Subbarao,K., Cox,N.J. and Fitch,W.M. (1999) Predicting the evolution of human influenza A, *Science*, Vol.286(5446):1921-1925.
- [3] Ghedin,E., Sengamalay,N.A., Shumway,M., Zaborsky,J., Feldblyum,T., Subbu,V., Spiro,D.J., Sitz,J., Koo,H., Bolotov,P., Dernovoy,D., Tatusova,T., Bao,Y., St George K., Taylor,J., Lipman,D.J., Fraser,C.M., Taubenberger,J.K., Salzberg,S.L. (2005) Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution, *Nature*, 437(7062):1162-1166.
- [4] Ito,K., Igarashi,M., Kida,H. and Takada,A. (2006) Computer Analysis of Structural Changes in H3 Hemagglutinins of Human Influenza Viruses Isolated During 1968 to 2006, ” In Proceedings of the Thirteenth International Conference Negative Strand Viruses 2006: 28.
- [5] Korber,B.T. Farber,R.M., Wolpert,D.H., and Lapedes,A.S. (1993) Covariation of mutations in the V3 loop of human immunodeficiency virus type 1 envelope protein: an information theoretic analysis, In Proceedings of the National Academy of Sciences, Vol 90:7176-7180.
- [6] Martin, L.C., Gloor, G.B., Dunn, S.D., and Wahl, L.M. (2005), Using information theory to search for co-evolving residues in proteins, *Bioinformatics*, Vol.21(22):4116-4124.
- [7] National Center for Biotechnology Information, “Influenza virus resource,” Jun 23, 2005. Available from: <http://www.ncbi.nlm.nih.gov/genomes/FLU/FLU.html>
- [8] 中島捷久, 信沢枝里, 中島節子 (2006) インフルエンザウイルス血球凝集素 (HA) 蛋白質におけるアミノ酸変異の基盤的解析, *ウイルス* 56巻1号: 91-98.
- [9] 中島捷久 (2003) 間断なき流行と新型ウイルス出現の機構, *日本臨牀*, 61巻11号:1897-1903.
- [10] 清水一史 (2000) インフルエンザウイルスの抗原性変異のメカニズム, *日本臨牀*, 58巻11号:2199-2205.
- [11] Taniguchi,T. and Haraguchi,M. (2006) Discovery of Hidden Correlations in a Local Transaction Database based on Differences of Correlations, *Engineering Applications of Artificial Intelligence*, Vol.19(4):419-428
- [12] Uno,T., Asai,T., Uchida,Y., and Arimura,H.,(2004) An Efficient Algorithm for Enumerating Closed Patterns in Transaction Databases, In Proceedings of the 8th International Conference on Discovery Science 2004 (DS-2004): 16-31.