

G614D 変異の検知による SARS-CoV-2 変異株の

研究所流出可能性の検証

神崎達也・二平泰知・三浦健人

アドバイザー教員：掛谷英紀

1. 背景と目的

Covid-19（新型コロナウイルス感染症）は2019年12月に中華人民共和国湖北省武漢市において確認された。当初は、ヒトからヒトへの感染は否定されていたが、2020年1月に入ってさらに感染が拡大し、中国国内だけにとどまらず世界中で多数の感染者が確認されることになった。そのような状況から、2020年3月11日にWHO（世界保健機関）はCovid-19をパンデミックとみなせると表明した。

パンデミックは世界中に大きな影響を与えたが、パンデミックを引き起こしたSARS-CoV-2ウイルス（新型コロナウイルス）がどのように出現したかについては明確に解明されていない。拡散したウイルスの起源として主に、動物が持つウイルスが人に感染した可能性と、研究所で保管されていたウイルスが何らかの要因によって研究所の外に流出し、人に感染したという可能性が指摘されている。

研究所からウイルスが流出した事例は実際に確認されている。台湾では研究所で保管されていたSARS-CoV-2のデルタ株が研究所の外に流出し、100名以上に感染が広がった。[1]この流出が起きた当時の台湾では厳しい水際対策がとられており、市中感染は広がっていなかった。そのような状態

で、研究所に勤務する女性やその周囲の人物が相次いでCovid-19に罹患したため、研究所からの流出が発覚した。仮に、同様の流出事故が他の国や地域で起こっていたとしても、市中感染が拡大している中では見落とされている可能性がある。

実際に、SARS-CoV-2の変異株については研究所から流出した可能性を示唆する研究がある。

Tanakaら[2]は、オミクロン株を中心にSARS-CoV-2変異株の異常を指摘している。

オミクロン株に分類される変異株には多くの変異が存在するが、そのほとんどが非同義変異（アミノ酸を変化させる塩基の変異）である。自然環境では、同義変異（アミノ酸を変化させない塩基の変異）が一定の割合で混じることから、非同義変異への異常な偏りは、人為的な遺伝子改変の可能性を示唆する。

また、オミクロン株において、約30あるスパイクタンパクのほぼ全ての変異で、もとの武漢株に先祖返りする復帰変異が見られている点も不自然であると指摘されている。復帰変異自体は、既存の株との相同組み換えで起きやすいことが知られているが、全ての変異で復帰変異が見られること、またその一部は変異部位が非常に近く相同組み換えでは起こりえないことなど、起こる可能性は自然界では極めて低い。

Takeya ら[3]は、SARS-CoV-2 において 2 つ以上の復帰変異がある株の復帰変異位置のパターンについて解析を行っている。相同組換えであれば、復帰変異位置は隣接するのが普通であるが、孤立した位置で複数の復帰変異が見られる株も多く検出されており、その不自然さを指摘している。これらの復帰変異株は、個々の変異の影響を調べるため、復帰変異を意図的に挿入した実験で生じ、それが事故で研究所から漏れた可能性も考えられる。

Takeya らは D614G 変異の復帰変異にも注目して解析を行っている。D614G 変異とは、SARS-CoV-2 のスパイクタンパクの 614 番目のアミノ酸が、アスパラギン酸 (略号: D) からグリシン (略号: G) に置換される変異のことである。

一般的に、ヒトの体内では配列の 614 番目のアミノ酸が D である場合は不安定で、614 番目が G である方が安定であることが知られている。そのため、ヒトからヒトへの感染によって伝播する過程では、D から G の変異は起きうる一方で、G から D への復帰変異は極めて起こりにくい。よって、ヒトかヒトへの感染で伝播するウイルスの配列の 614 番目は G であるはずである。一方で、実験室などで培養される環境では、614 番目 D でも安定した状態を保つことが知られている。

このような性質から、D614G の復帰変異である G614D 変異がまとまって検出された場合、その検出日時の直前、検出場所の近くで、研究所から流出した可能性が示唆される。

そこで本研究では、遺伝子データベースに登録された SARS-CoV-2 の配列を解析し、G614D 変異を含む配列を抽出し、それらの採取場所と採取日時の分布がもつ地理的、時系列的特徴を調べる。

2. 手法

本研究では SARS-CoV-2 のデータベースとして、NCBI (アメリカ国立生物工学情報センター) の GenBank [4]を用いる。収集するデータは、Takeya らの先行研究で調べられていない変異株のうち、登録データ数が 1000 以上のものとした。

Takeya らは挿入や脱落のある配列は解析対象外としており、また全く同じ配列をもつものはダブルカウントしない集計方法をとっている。本研究では、挿入や脱落のある株も対象とし、全く同じ配列パターンであっても G614D 復帰変異を含むものは加算してカウントする手法をとる。

具体的には、614 番目のアミノ酸配列の前 2 つと後ろ 2 つのアミノ酸が変化しないで、614 番目のアミノ酸が D に復帰しているときに現れる YQDVN の配列を探し、その配列の Accession 番号を抽出する。

集計方法が異なるため、Takeya らの論文で G614D 復帰変異が多く見られた B.1.617.2 と BA.2 についても再度集計を行う。そして、G614D が多く検出された変異株について、Accession 番号から採取日時と採取場所を特定し、それを集計・分析する。

3. 結果

本研究で調査した変異株の名称、G614D 復帰変異の数、全データ数、G614D 復帰変異を含む配列の割合を表 1 に示す。

その結果、先行研究で G614D 復帰変異が多いことが分かっている B.1.617.2 と BA.2 に加え、AY.103、AY.44、B.1.15 で G614D 復帰変異が多いことが分かった。そこで、これら 5 つの変異株について発生位置と発生時期について分析を行った。

表 1 G614D 変異の発生数と割合

変異	AY.103	AY.44	B.1.1.519	B.1.427
G614D	163	134	5	4
全体数	242444	208550	12534	13062
百分率	0.06723	0.06425	0.03989	0.03062
B.1.429	B.1.526	B.1.637	BA.1.15	B.1.619.2
7	10	8	57	403
29260	33588	11185	84850	45699
0.02392	0.02977	0.07152	0.06718	0.88186
BA.2	BA.5.2.1	BA.5.5	BE.1	BE.3
189	5	6	1	0
84062	82004	34641	6004	5185
0.22483	0.00610	0.01732	0.01666	0.00000
CH.1.1	D.2	XBB.1.16	XBB.1	
0	0	0	2	
1311	11423	3388	3186	
0.00000	0.00000	0.00000	0.06277	

3-1. B.1.617.2(デルタ株)

最初に、Kakeya らの研究でも G614D 復帰変異が多く見つかっている B.1.617.2 について、採取時期と採取場所の情報を集計した結果を図 1 から図 3 に示す。

Kakeya らの研究でも明らかになってお

り、G614D 復帰変異の発生時期は、同変異株の採取のピークからかなり遅れており、流行後期での G614D 復帰変異の上昇は、不安定な変異株が流行後期に現れている点において、極めて不自然である。

G614D 復帰変異の採取位置は、同変異株全体の分布と比較すると、特にアメリカ北東部のミシガン州、イリノイ州にて検出が多く見られる。

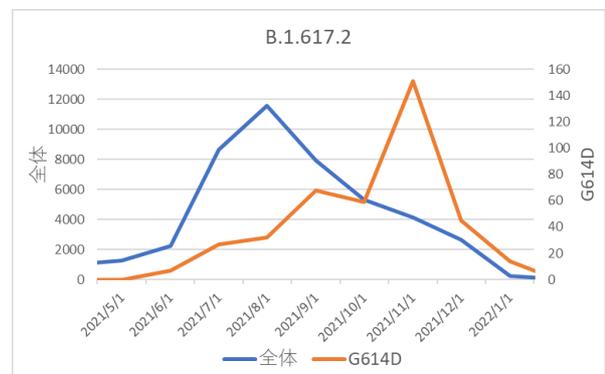


図 1 B.1.617.2 の発生時期

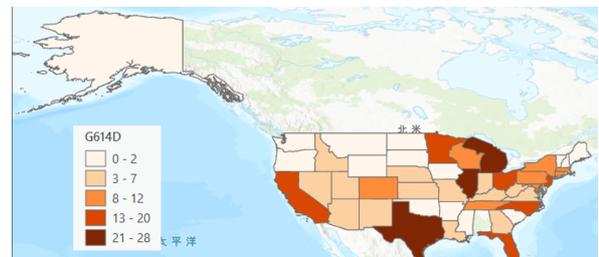


図 2 B.1.617.2 の G614D 変異の発生位置

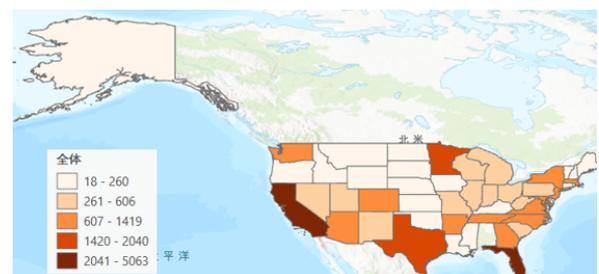


図 3 B.1.617.2 の株全体の発生位置

3-2. AY.103

次にデルタ株の亜種である AY.103 について採取時期と採取場所の情報を集計した結果を図4から図6に示す。

図4が示す通り、G614D 変異のピークは全体のピークから少し遅れており、B.1.617.2 の G614D 復帰変異のピーク（2021年11月）と一致している。

採取場所についても、B.1.617.2 同様、G614D 復帰変異はアメリカ北東部で多く検出されている。

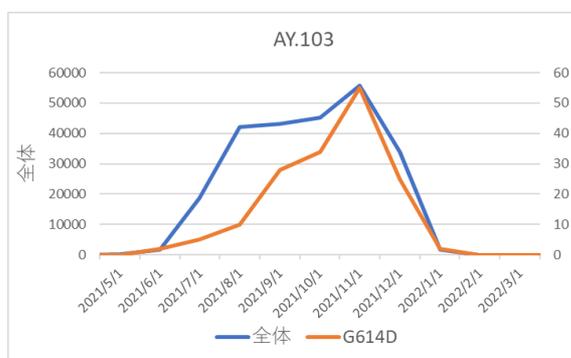


図4 AY.103 発生時期

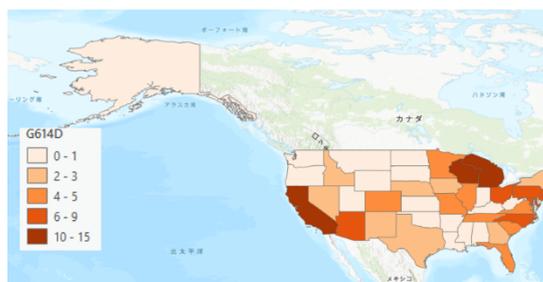


図5 AY.103 の G614D 変異発生位置

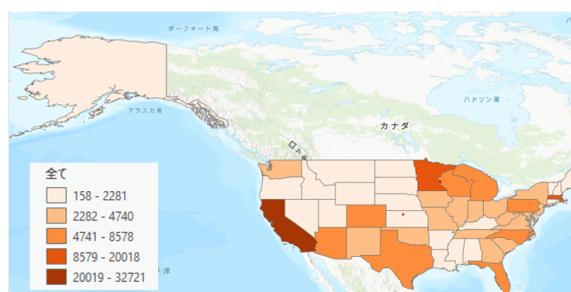


図6 AY.103 の株全体の発生位置

3-3. AY.44

次に同じくデルタ株の亜種である AY.44 について採取時期と採取場所の情報を集計した結果を図7から図9に示す。

図7が示す通り、G614D 復帰変異のピークは全体のピークより遅れている。

G614D 復帰変異検出位置は、他のデルタ株およびその亜種と同様北東部で若干多い。

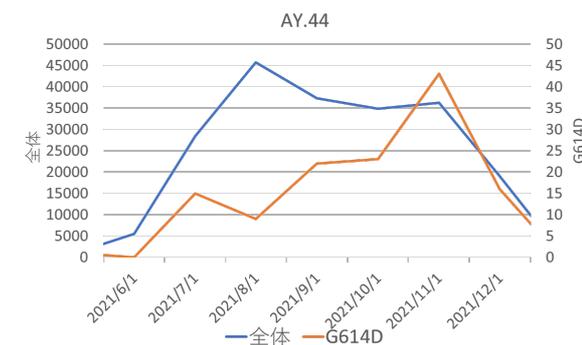


図7 AY.44 発生時期



図8 AY.44 の G614D 発生位置

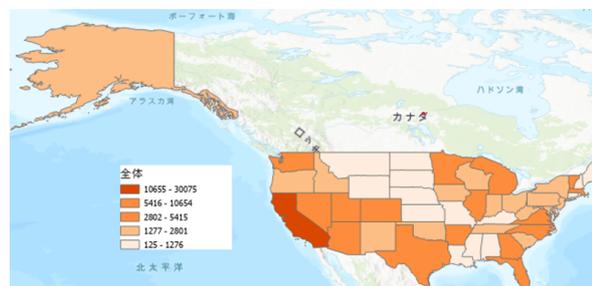


図9 AY.44 の株全体の発生位置

3-4. BA.1.15

次にオミクロン株 BA.2 の採取時期と採取場所の情報を集計した結果を図 10 から図 12 に示す。

図 12 が示す通り、G614D 復帰変異のピークは全体のピークとほぼ一致している。

また、採取場所はいずれも、アメリカ南部のテキサス州やカリフォルニア州が多く、傾向が似ている。

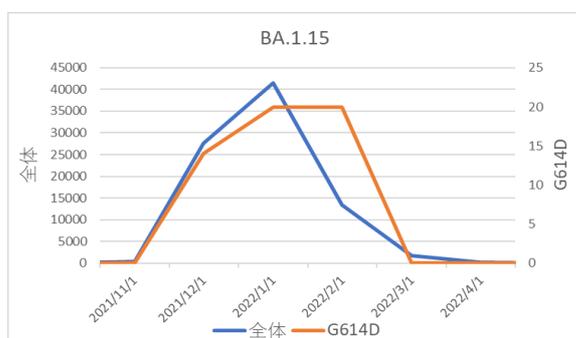


図 10 BA.1.15 変異の発生時期

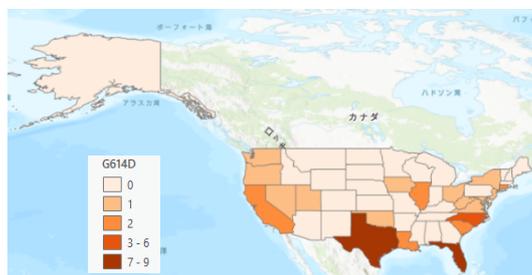


図 11 BA.1.15 の G614D 発生位置

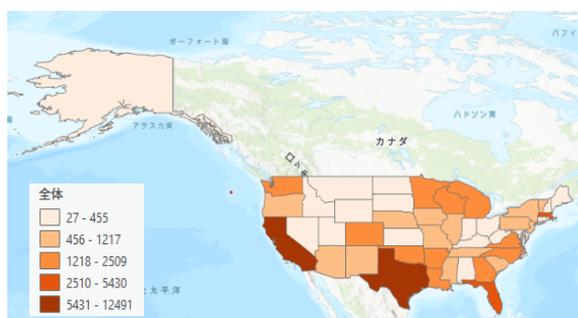


図 12 BA.1.15 の株全体の発生位置

3-5. BA.2

最後にオミクロン株 BA.2 の採取時期と採取場所の情報を集計した結果を図 13 から図 15 に示す。

採取時期については、全体のピークが 2022 年の 4 月なのに対し、G614D 復帰変異のピークが同年 5 月と少し遅れている。

G614D 復帰変異は全データの採取場所の分布に似ている者の、特にニューヨーク州で多く検出されていることが分かる。

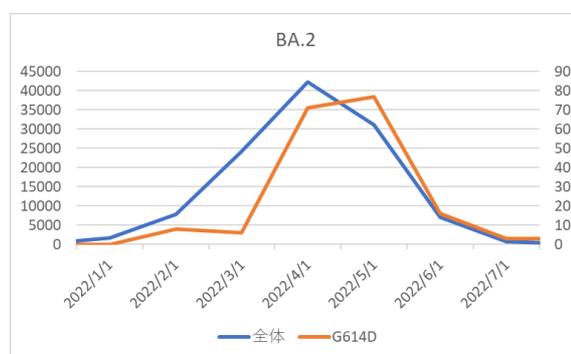


図 13 BA.2 の発生時期



図 14 BA.2 の G614D の発生位置

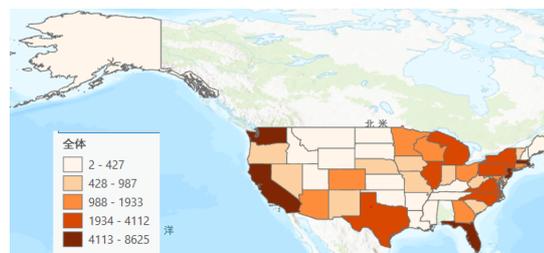


図 15 BA.2 の株全体の発生位置

4. 考察

デルタ株で G614D 復帰変異が多く発見されたミシガン州や、イリノイ州の隣にあるミズーリ州には SARS-CoV-2 の配列を扱っている施設があるため、デルタ株 (B.1.617.2、AY.103、AY.44) の G614D 変異はイリノイ州やミシガン州の周辺の SARS-CoV-2 の配列を扱う施設から漏れ出た可能性があると考えられる。

5. まとめ

本研究では遺伝子データベースに登録された SARS-CoV-2 のデータから G614D 復帰変異の抽出を行い、検出数の多かった株について採取日時の時系列と場所の位置の分析を行った。その結果、デルタ株 B.1.617.2 の G614D 復帰変異時期は、同株の流行から後れて発生していること、デルタ株とデルタ株亜種 (AY.103、AY.44) の G614D 復帰変異時期のピークは 2021 年 11 月で一致していること、デルタ株の G614D 復帰変異はミシガン州とイリノイ州で多く見られることが分かった。

これらのデータは研究所流出の可能性を示すものの、研究所流出の直接的な証拠を提供するというものではない。今後は、関係各所の立ち入り調査などをもとに、研究所流出の可能性を直接検証する必要があると考えられる。

6. 参考資料

[1] A lab assistant involved in COVID-19 research in Taiwan exposed 110 people

after becoming infected at work - Bulletin of the Atomic Scientists (thebulletin.org)

[2] Atsushi Tanaka, and Takayuki Miyazawa, “Unnatural evolutionary processes of SARS-CoV-2 variants and possibility of deliberate natural selection”, August 15, 2023

DOI:10.5281/zenodo.8361577

[3] Hideki Kakeya, and Yoshihisa Matsumoto, “Repeated emergence of probabilistically and chronologically anomalous mutations in SARS-CoV-2 during the COVID-19 pandemic”, August 4, 2023

DOI:10.5281/zenodo.8216232

[4] “NCBI Virus”

<https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>

[5] 宮坂昌之, 変異ウイルスとワクチンの作用機序、効果, 公益財団法人 日本記者クラブ 2021 年 5 月 11 日資料