

## 5 學術誌等掲載論文

山階鳥類学雑誌, 55: 1-12, 2023

繁殖期に魚類を利用するイヌワシ

前田 琢

A Pair of Golden Eagles Suppling Fish as Food for Nestlings

Taku Maeda

Abstract

Video monitoring on breeding nests of a pair of Golden Eagles *Aquila chrysaetos japonica* in the Kitakami Mountains, Iwate, Northern Honshu, Japan, revealed that fish were supplied as food for nestlings. Fishes, totalling 57 individuals, were presented to chicks in four of the seven years analyzed, and comprised 19% of the 306 total prey animals detected. All fish were brought by the male, and most of them were removed from the nest by the female soon after, except for 10 fishes consumed by the female and nestlings. Fish supply during the nestling period was either continuous or temporary, depending on the year. Fish species could not be identified, but the variations in size, coloration, pattern and shape of fishes indicated various species were used. The method of obtaining fish was unknown, but the variation of fish species and irregular supply patterns suggested the possibility that Golden Eagles snatched prey from fish-eating birds, such as Western Ospreys *Pandion haliaetus*.

原著論文 Original Article

## 生息域外保全を目的にした絶滅危惧種チョウセンキバナアツモリソウの発芽法の開発と保存種子を用いた苗生産

小山田智彰<sup>1</sup>・鞍懸重和<sup>1</sup>・千崎則正<sup>1</sup>

岩手県環境保健研究センター<sup>1</sup>

Tomoaki OYAMADA, Shigekazu KURAKAKE, Norimasa SENZAKI:  
Development of a Method for Germinating Endangered *Cypripedium guttatum* and Efficacy of Seedling Production from Preserved Seeds Aimed at Ex Situ Conservation

要旨：国内において、チョウセンキバナアツモリソウ (*Cypripedium guttatum*) の自生地は1カ所であり、保護措置による個体数の増加傾向が見られないことなどから、生息域内保全を補完するための生息域外保全が急務になっている。本研究は、環境省東北地方環境事務所から要請を受け、環境省新宿御苑管理事務所に保存されてきた自生地由来となるチョウセンキバナアツモリソウの種子の発芽に取り組むことにより、生息域外保全の推進のために必要とされる保存種子の活用の有効性について検証した。研究は、保存種子の発芽に取り組む前に、予備試験として栽培地で維持されてきた栽培株より結実した種子900粒を材料にして、アツモリソウ属植物用に開発した小山田培地を使用し発芽試験を実施した。発芽した312個体を育苗培地に継代し、成長が認められた99個体をフラスコから取り出し、用土を充填したトロ箱に植え付けて2年間の育苗を行った。育苗によって生存した苗から30個体を栽培試験地に定植した結果、定植の翌年に初開花を確認した。この結果を参考にし、本試験として新宿御苑管理事務所に保管されていた2014年に採種された種子290粒を培地に播種したところ6個の発芽個体が得られた。これを育成用培地に継代して培養し、最終的に2個体の苗が得られた。また、2015年に採種された保存種子7265粒を材料に発芽試験に取り組んだ。その結果、培養開始50日から800日までの期間に断続的に108個体の発芽が確認された。発芽率は1.5%となり、予備試験として実施した栽培地由来の種子の発芽率34.7%と比較して有意に低くなることが分かった。発芽した108個体から器官分化が見られた101個体を育成培地に継代して育苗し、成長した57個体を培養フラスコから取り出して、栽培用土を充填したトロ箱に移植して2年間の育苗を行い、生存した22個体を環境省に提出した。本研究は、環境省が進めている絶滅危惧植物の保護増殖事業の中で取り組んだものであり、特にチョウセンキバナアツモリソウ野生株から採種された保存種子の発芽は国内外において初報告になり、生息域外保全に用いる保存種子の苗生産に成功した。

**Abstract :** *Cypripedium guttatum* (common name: Spotted Lady's Slipper) grows naturally in one location in Japan. However, the failure of conservation measures to increase the *C. guttatum* population has created an urgent need to supplement in situ conservation with ex situ conservation. We were contracted by the Ministry of the Environment (MOE) Tohoku Regional Environment Office to develop a method to germinate *C. guttatum* seeds collected from *C. guttatum*'s natural habitat and stored by the MOE Shinjuku Gyoen National Garden Management Office and to evaluate the efficacy

<sup>1</sup> 〒020-0857 岩手県盛岡市北飯岡1-11-16; Iwate Prefectural Research Institute for Environmental Sciences and Public Health (I-RIEP), 1-11-16 Kitaiioka, Morioka, Iwate, 020-0857, Japan

of using preserved seeds, which is essential for promoting ex situ conservation. Prior to attempting to germinate preserved seeds, we conducted a pilot study using 900 seeds collected from *C. guttatum* maintained in cultivation plots and Oyamada medium developed especially for *Cypripedium* spp. The 312 seedlings that germinated were subcultured in seedling growth medium; of these, the 99 seedlings that showed growth were removed from their flasks and transplanted to planting trays filled with horticultural soil and grown for 2 years. The 30 surviving seedlings were planted out in experimental plots and the first bloom was observed the following year. Based on these results, we sowed 290 seeds collected in 2014 and stored by the Shinjiku Gyoen National Park Management Office, which resulted in the germination of six seedlings. These seedlings were subcultured in seedling growth medium and ultimately yielded two plants. Based on the results of this pilot study, we conducted a germination experiment using 7,265 preserved seeds collected in 2015. Intermittent germination of 108 seedlings was observed between day 50 and day 800 after sowing. The resulting germination rate (1.5%) was significantly lower than the germination rate of seeds collected from cultivated plants (34.7%) conducted as a control experiment. Of these 108 seedlings, the 101 that exhibited differentiation were subcultured in seedling growth medium; of these, the 57 that showed growth were removed from their flasks and transplanted to planting trays filled with horticultural soil and grown for 2 years. The 22 surviving seedlings were submitted to MOE. This research, conducted as part of MOE's effort to conserve and increase populations of endangered plants, has yielded the first report both inside and outside Japan of germination of *C. guttatum* from preserved seeds collected from wild *C. guttatum* and demonstrates the efficacy of using preserved seeds for ex situ conservation.

キーワード: チョウセンキバナアツモリソウ, 自生地, 生息域外保全, 保存種子, 発芽

**Keywords:** *Cypripedium guttatum*, natural habitat, ex situ conservation, preserved seed, germination

## I. はじめに

チョウセンキバナアツモリソウ (*Cypripedium guttatum*) は、2002年に「絶滅のおそれのある野生動植物の種の保存に関する法律」の国内希少野生動植物種に指定されている（環境省、2018）。2004年には農林水産省・環境省共管で「保護増殖事業計画」を策定し、生育状況のモニタリング、盗掘防止の巡視、生育環境の改善のための植生管理等を実施している。しかし、国内の自生地は1カ所であり、2017年10月、2018年6月と7月に実施した自生地での現地調査では、開花まで到達した植物体はわずかで、開花から結実に至る個体はほとんど見られなかった（小山田ほか、2019；表1）。また、開花個体においても、地上部が黒色化し、その後に地上部が溶け落ちる現象が確認された。これは、栽培管理の中で健全に開花したチョウセンキバナアツモリソウには見られないことから、自生地のチョウセンキバナアツモリソウ群落は衰退傾向にあることが推測され、生息域内保全を補完す

るための生息域外保全が急務になっている。本種の生息域外保全の推進のために環境省東北地方環境事務所の要請を受けて環境省新宿御苑管

表1. チョウセンキバナアツモリソウ自生地の開花および結実状況

**Table 1.** Blooming and fruit bearing of *Cypripedium guttatum* in their natural habitat.

調査年	開花数	結実数	さく果回収
2004	9	0	0
2005	11	4	0
2006	21	2	0
2007	17	2	0
2008	17	2	0
2009	1	1	0
2010	8	3	0
2011	1	1	0
2012	5	1	0
2013	2	1	0
2014	6	1	1 <sup>z</sup>
2015	13	5	4 <sup>z</sup>
2016	10	1	1
2017	5	0	0
2018	6	1	1

<sup>z</sup> 2014年と2015年の種子を本研究に使用

理事務所（以下、「新宿御苑」）に長期保存されていた種子（以下、「保存種子」）の発芽に取り組んだ。

国外において、本種の発芽試験に取り組んだ報告は数例あるが、幼苗の育苗を進めて栽培・開花に至った報告例は確認できず（Bae et al.,2009；Ishii et al.,2005；佐々木,2000），これについて環境省に聞き取りを行ったが同様の回答であった。

そこで本研究では過去に取り組んだアツモリソウに用いた発芽法（小山田,2003;小山田ほか,2008；小山田ほか,2011）を参考に、予備試験として岩手県内に設置したアツモリソウ属植物の栽培試験地（以下、「栽培地」）のチョウセンキバナアツモリソウの種子（以下、「栽培地由来種子」）を用い、発芽・育成技術の確立に取り組んだ。初めに、今後の発芽・育成試験の材料となる種子数を把握、予測するため、栽培地から採取したさく果の外部形態と種子数の関係を明らかにし、その後栽培地由来種子を用いた発芽・育成試験を行った。また、保存種子の発芽・育成技術を確認する事前調査として、発芽、育成培地および培養土の組成を検討するため、自生地の土壌分析を行った。これらの結果を基に本試験として保存種子の発芽、育成試験に取り組む、生息域外保全における保存種子活用の有効性を検討した。

## II. 材料と方法

### 1. 予備試験

チョウセンキバナアツモリソウの種子および発芽の状況を取りまとめた既往の報告等がないため、貴重な保存種子を取扱うにあたり、著者が保有する「種の保存法」規制以前となる1993年に園芸店から購入した2株から増殖した株を使用して予備試験を行った。

#### (1) 栽培地由来のチョウセンキバナアツモリソウのさく果の外部形態と種子数の関係（試験1-1）

2010年に栽培地の19株からさく果を採取し、結実したさく果のサイズを「大」、「中」、「小」、「不

良（虫食い）」に目視分類した後、長さ、径をノギスで、重量を電子天秤（LIBROR EB-28, 島津製作所）で計測した。計測後、さく果内の種子をシャーレーに取り出し、実体顕微鏡（SZX-ILLB100, OLYMPUS）を用いて種子数をカウントした。計測後、さく果のサイズ中のものから900粒を発芽試験に用いた。

#### (2) 栽培地由来種子の発芽（試験1-2）

栽培地由来種子の発芽試験を2010年に行った。種子は、開花・結実から90日が経過した完熟種子を材料に用いた。

種子殺菌は、試験1-1で採取した種子を、次亜塩素酸ナトリウム水溶液（有効塩素0.6%）を入れた遠沈管に移し、振動殺菌を30分間行った。

播種は、殺菌した種子をクリーンベンチに搬入後、小山田培養液（小山田ほか,2011；小山田ほか,2021；表2）に30分間種子を沈殿させ、pH5.7に調整した小山田培地（小山田ほか,

表2. 小山田培養液と培地の組成

**Table2.** Composition of Oyamada culture solution and growth medium.

小山田培養液の組成	添加量	小山田固形培地の組成	添加量
塩酸チアミン	5 mg・L <sup>-1</sup>	Hyponex (6.5 - 6.0 - 19.0)	1 g・L <sup>-1</sup>
ニコチン酸	2.5 mg・L <sup>-1</sup>	ベプトン	1 g・L <sup>-1</sup>
塩酸ピリドキシン	5 mg・L <sup>-1</sup>	スクロース	20 g・L <sup>-1</sup>
ミオイノシトール	10 mg・L <sup>-1</sup>	小山田培養液	1 mL・L <sup>-1</sup>
植物活性剤	0.1 mL・L <sup>-1</sup>	ポテトキューブ <sup>2</sup>	5 mm角
木酢液	0.1 mL・L <sup>-1</sup>	活性炭 <sup>3</sup>	1 g・L <sup>-1</sup>
pH	5.8	ゲランガム	3 g・L <sup>-1</sup>
		pH	5.7-5.8

<sup>2</sup> 培養容器に1個加えた後に、オートクレーブで滅菌する。

<sup>3</sup> 活性炭は発芽培地に用いない。

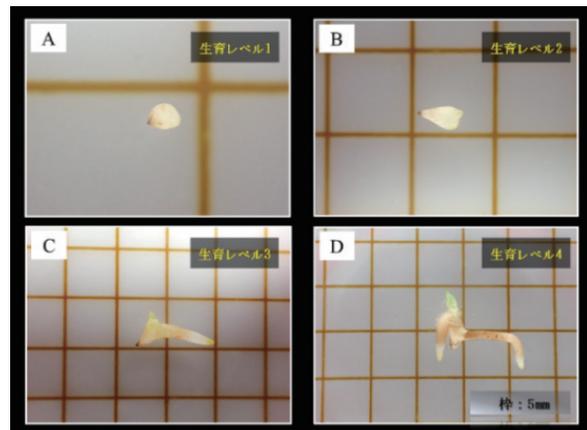


図1. 小山田培地で発芽した苗の生育レベル（4区分）A: 発芽直後のプロトコーム, B: プロトコームが変形肥大した形態, C: 器官分化を開始した形態, D: 芽と根が伸長した苗

**Fig.1.** Development stages of seedlings germinated in Oyamada medium (four stages).

2011；小山田ほか，2021；表2）を100 mL 充填した広口ガラス容器（450 mL，石塚硝子社）へ，1容器あたり300粒を目安に900粒播種した。その後，培養容器に小山田培養液を5 mL 添加した。

培養は，20℃の暗黒条件に設定したバイオマルチンキュベータ（LH-30-8CT,NK システム）に，播種した培養容器を180日間を目安に静置した。発芽の判断は目視により未分化で極小の白い球体形をなすプロコトームを確認した時とし，経過観察から発芽開始日を記録し，培養終了後に発芽個体をカウントした。

### (3) 栽培地由来の育成培養（試験1-3）

継代は，試験1-2で発芽した312個体を小山田培養液に浸透させた後，活性炭入りの小山田培地（pH 5.7に調整）を100 mL 充填した広口ガラスの培養容器（450 mL，石塚硝子社）へ生育レベル別に，1容器当たり1～3個体移植した。この時，生育レベル2で継代したものを生育レベル2区，3で継代したものを生育レベル3区とした。培養は試験1-2と同条件で行った。尚，育成培養の期間は180日を目安に行った。

育成培養完了時に，生存個体をカウントし，その内30個体については草丈，根長をノギスで，重量を電子天秤で計測し，根数，越冬芽数をカウントした。また，生育レベル2区と3区で生存した各15個体についても，重量を除いた項目について同様の方法で調査し，生育レベル間で比較した。

### (4) 栽培地由来の順化・鉢上げ（試験1-4）

栽培地由来種子の培養苗99個体の順化・鉢上げを2011年に行った。順化・鉢上げは，過去に実施した方法（小山田，2003）に従った。つまり，栽培容器（発泡スチロール製トロボ箱縦50 cm × 横30 cm × 高さ20 cm）に硬質鹿沼土（小粒），硬質赤玉土（小粒），砂，もみ殻くん炭を比率6:2:1:1で混合した用土を深さ15 cm まで充填し，その上に約1 cm 深でスギ・ヒノキの樹皮を薄く敷き詰め，そこに栽培地由来種子の培養苗を栽培容器あたり20個体を目安に植栽した。植栽した栽培容器は屋外の棚に設置し，自然環境下

で管理した。順化・鉢上げは，野外の最高気温が20℃以下の日が連続するようになったのを確認してから実施した。

苗の管理は，4月から11月まで月1回の割合で小山田培養液（表2）を各栽培容器に1000 mL 散布した（小山田ほか，2021）。また，夏期の間は寒冷紗を設置して遮光を行ない，日中の照度が10000 lx 以下になるよう調節した。降雪が確認された時点で苗の管理を終了し冬越した。

順化・鉢上げから2年後に生存個体をカウントした。また，生存した45個体については，縦230 cm × 横100 cmに設定した栽培地へ30個体を定植，15個体を栽培容器のまま管理し，その後の生育，開花状況を観察して本試験で用いた発芽法の実用性について評価した。

## 2. 自生地の土壌分析（試験2）

自生地の土壌は，2017年，2018年に自生地の上の群落（「以下，U区」）（開花あり；写真1）および下の群落（「以下，L区」）（開花なし）の植物体付近から，1サンプル60 g，3サンプルずつ採取した。この時，開花期に自生地に入り，土壌の採取は植物体の発生地点から根系の伸長範囲となる深度5 cm を目安とし，植物体に損傷を与えないようにした。

土壌分析は，採取した土壌のアンモニア態窒素，硝酸態窒素，可給態鉄，交換性マンガン，可給態リン酸，交換性カリウム，交換性カルシウムおよび交換性マグネシウムを土壌養分検定器（Dr. ソイル，富士平工業）で分析した。pH と EC は測定器（H198129，HANNA 社）を使用して測定した。比較のために行ったアツモリソウとクマガイソウ自生地の土壌分析も同じ方法で分析した。分析11項目から，特にアツモリソウ属植物の土壌比較に用いられている窒素，交換性カルシウム，pH と EC に着目し（小山田ほか，2020；小山田ほか，2021），開花数および個体数の多いU区と，開花のないL区の土壌を評価し，培地調整に活用した。

## 3. 本試験

本試験については，予備試験および自生地の土壌試験結果を基に本試験を実施した。

### (1) 保存種子の発芽 (試験 3-1)

試験の種子には、新宿御苑の種子保存施設内で種子調整および乾燥作業を行った後に $-20^{\circ}\text{C}$ で長期保存されていた保存種子(環境省自然環境局, 2009; 環境省新宿御苑管理事務所, 2010)を用いた。保存後2年経過した2016年と2017年に受け取った保存種子は、それぞれ2014年開花1株の種子(290粒)と、2015年開花4株の種子(7265粒)で、どちらも完熟種子である(表1)。これらを材料に発芽試験に取り組んだ。

播種は、種子を小山田培養液に2時間沈殿させたこと、また、小山田培地のpHを5.8に調整したことを除き、試験1-2と同様の方法で行った。

培養は、試験1-2と同条件で行ったが、培養容器の静置期間は、発芽開始日数が個体によって50日から800日と大きく異なったため、培養容器を定期的に確認し、発芽後生育レベル2で取り出した。

計測は、試験1-2と同様の項目を同様の方法で行い、栽培地由来種子と比較した。

### (2) 保存種子の育成培養 (試験 3-2)

継代および培養は、試験3-1で発芽後生存した2014年採取種子由来の5個体および2015年採取種子由来の101個体について、試験1-3と同様の方法で行った。尚、発芽に要する日数に差が生じたため、育成培養の期間は180日を目安に実施した。計測は、育成培養完了時に試験1-3の生

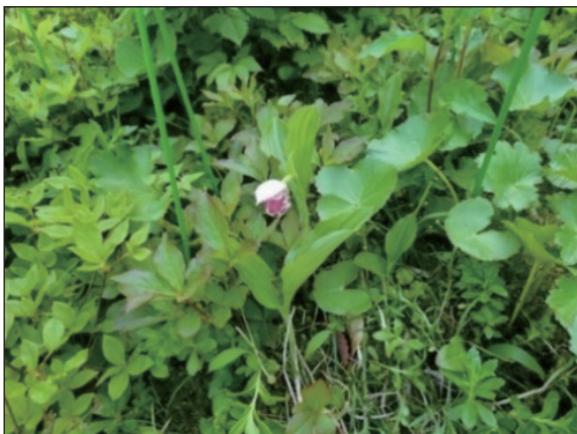


写真1. 土壌採取地点のチョウセンキバナアツモリソウ野生株(U区)

Photo1. Wild *Cypripedium guttatum* growing at a soil sampling location (upper community).

育レベル間の生育比較を除いた項目を、同様の方法で計測した。この時、生育調査は57個体を対象とし、栽培地由来の培養苗と比較した。

### (3) 保存種子の順化・鉢上げ (試験 3-3)

保存種子の培養苗57個体の順化・鉢上げは2019年に行った。順化・鉢上げおよび苗の管理の方法は試験1-4と同様の方法で行った。ラン科植物は土中の菌と共生関係を結んでいること(松本ほか, 1988; 清水ほか2001)や、発芽と実生時の栄養摂取に菌根菌を必要とし、自生分布は、菌根の特異性によって制限されるという報告もある(Shefferson et al., 2005)ため、土壌分析に使用したU区の土壌を苗の根系に接するように充填した。順化・鉢上げ2年後に生存個体をカウントし、生存率を栽培地種子由来の培養苗の生存率と比較した。

## 4. 統計解析

さく果の外部形態と種子数の解析は、さく果内の種子数を目的変数、さく果の長さ、径および重量を説明変数とし、単回帰分析を行った。また、それぞれの相関係数について無相関の検定を行った。

栽培地由来種子の育成培養における継代時の生育レベル間での培養苗の生育比較では、生育レベル2区と生育レベル3区の培養苗の草丈および根長はウェルチのt検定で、根数と越冬芽数はマンホイットニーのU検定で、生存率はフィッシャーの正確確率検定で比較した。

2015年採取の保存種子と栽培地由来種子の発芽開始日数および発芽率の比較では、発芽開始日数はマンホイットニーのU検定で、発芽率はフィッシャーの正確確率検定で比較した。

育成培養完了時の保存種子および栽培地由来種子の培養苗の生育および生存率の比較では、培養苗の草丈、根長および重量はウェルチのt検定で、根数と越冬芽数はマンホイットニーのU検定で、生存率はフィッシャーの正確確率検定で比較した。

順化・鉢上げ2年後の生存率の比較では、保存種子および栽培地由来種子の培養苗の生存率をフィッシャーの正確確率検定で比較した。

### Ⅲ. 結果

#### 1. 予備試験結果

##### (1) 栽培地由来のチョウセンキバナアツモリソウのさく果の外部形態と種子数の関係 (試験 1-1)

目視分類により、さく果は大5個、中5個、小5個、不良(虫食い)4個であった。また、サイズ別の種子数の平均値およびレンジは、大8337粒(レンジ:7271-10118粒)、中5816粒(レンジ:4651-6670粒)、小2910粒(レンジ:100-5229粒)、不良1772粒(レンジ:217-4469粒)であった(表3)。大のさく果と比較して、中、小および不良のさく果の平均種子数は少なく、小および不良のさく果に含まれる種子数にはばらつきが見られた。

図2にさく果の長さ、径、重量と種子数の関係を示す。その結果、さく果の長さ、径および重量と種子数で正の相関が見られ、特に、長さよりも径と重量に強い相関が見られた。また、径が10 mm以上であること、重量が0.4 g以上に達したさく果を回収することで、平均的な種子数と言える5000粒の種子を確保できることが示唆された。

##### (2) 栽培地由来種子を用いた発芽 (試験 1-2)

播種後80日目からプロトコームが発生し、発芽が確認された。平均発芽開始日数は $126.3 \pm 22.7$ 日(平均 $\pm$ 標準偏差)、発芽率は34.7%であった。

##### (3) 栽培地由来の育成培養 (試験 1-3)

育成培養完了時の培養苗の草丈、根長、根数、越冬芽数、重量及び生存率は、それぞれ $8.9 \pm 0.3$

表3. 栽培地で回収したさく果の種子数調査

**Table3.** Numbers of capsules and seeds collected from cultivated *Cypripedium guttatum*.

さく果の状況 (目視判定)	さく果数 (個)	平均種子数 (最大-最小値)
大	5	8337 (10118-7271)
中	5	5816 (6670-4651)
小	5	2910 (5229-100)
不良	4	1772 (4469-217)
参考: 人工交配	2	6143 (6908-5377)

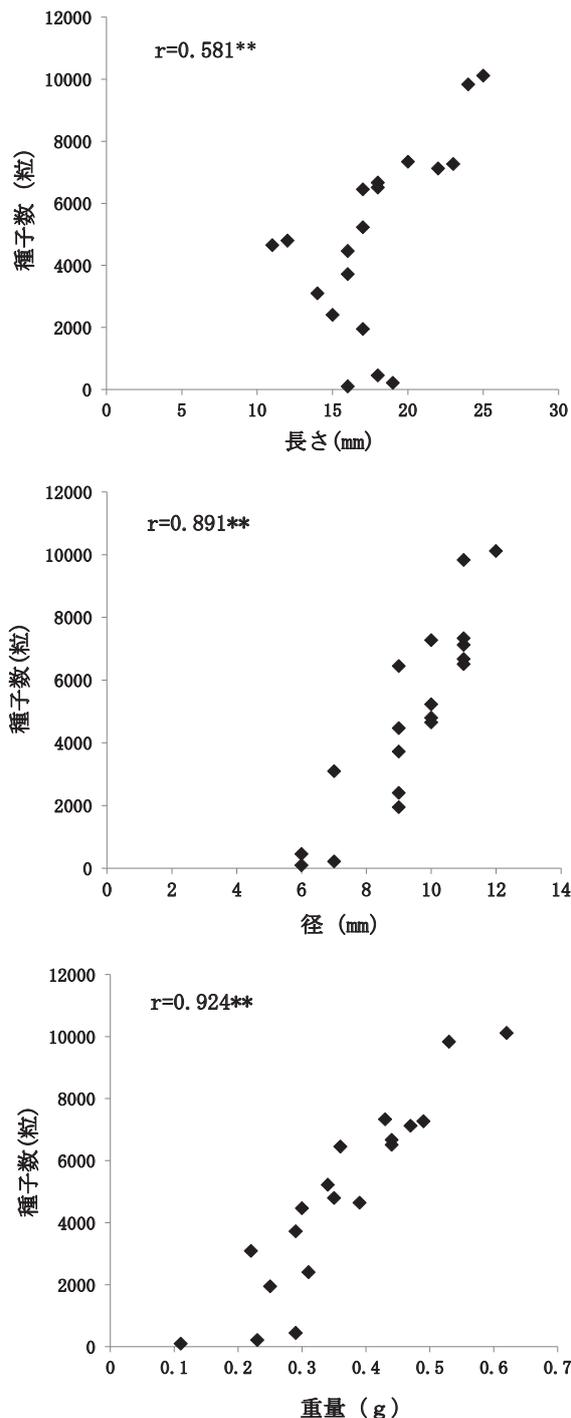


図2. さく果の長さ・径・重量と種子数の関係

図中rはピアソンの相関係数を示す。\*\*は無相関の検定により、1%水準で有意差があることを示す。

**Fig.2.** Relationship between capsule length, diameter, weight, and seed number.

mm(平均 $\pm$ 標準誤差),  $46.7 \pm 2.2$  mm,  $4.7 \pm 0.3$  本,  $0.7 \pm 0.1$  本,  $0.23 \pm 0.11$  gおよび31.7%であった。

表4に生育レベル2区と3区における培養苗の草丈、根長、根数、越冬芽数および生存率を

示す。その結果、根長と生存率に有意な差が見られ、生育レベル2の根長が51.4 ± 2.3 mm（平均 ± 標準誤差）で、3の41.9 ± 2.4 mmより有意に長かった。また、生育レベル2区の生存率が40.8%で、生育レベル3区の23.0%より有意に高かった。

これらのことから、生育レベル2で育成培養することで、培養苗の根長および生存率が高まることが示唆された。これを受け保存種子の育成試験は、生育レベル2で育成培地に継代した。

(4) 栽培地由来の順化・鉢上げ（試験1-4）

栽培地由来種子の培養苗の順化・鉢上げ2年後の生存率は45.5%であった（図3；2013年5月19日）。

露地の栽培試験地に定植した30個体（図3；2013年6月18日）は翌年に初開花が確認され（図3；2014年5月18日）、2021年には、開花数304株まで開花が認められた（図3；2021年5月16日）。

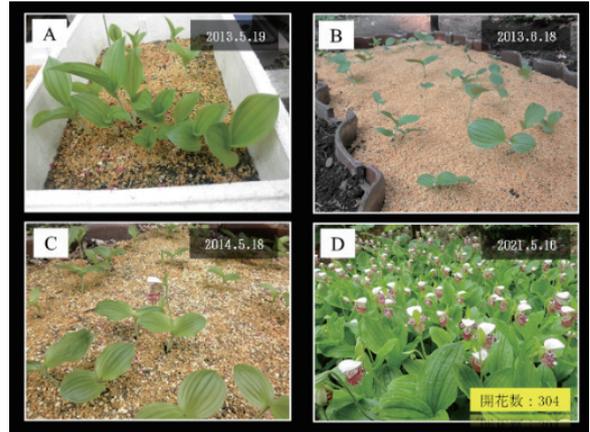


図3. 栽培地に定植した苗の初開花と群落化の様子（2013年定植：30個体）

A: 順化・鉢上げ2年後の栽培地由来種子の培養苗, B: 露地の栽培試験地に定植後の個体, C: 初開花個体, D: 露地栽培8年後の開花株群

Fig.3. First bloom and clustering of *Cypripedium guttatum* seedlings planted out to a cultivation plot (30 plants planted out in 2013).

表4. 生育レベル2区と生育レベル3区の培養完了時の生育と生存率

Table4. Growth of seedlings subcultured at development stages 2 and 3 at completion of subculturing (acclimatization, potting).

試験区	草丈 (cm)	根長 (mm)	根数 (本)	越冬芽数 (本)	生存率 (%)
生育レベル2の継代区	9.3 ± 0.4 <sup>z</sup> (15)	51.4 ± 3.3 (15)	4.5 ± 0.3 (15)	0.6 ± 0.2 (15)	40.8 (179)
生育レベル3の継代区	8.6 ± 0.5 (15)	41.9 ± 2.4 (15)	4.9 ± 0.6 (15)	0.7 ± 0.2 (15)	23.0 (113)
有意性 <sup>y</sup>	n. s.	*	n. s.	n. s.	**

<sup>z</sup> 平均 ± 標準誤差 (供試数)

<sup>y</sup> 草丈および根長の\*はウェルチのt検定により、根数および越冬芽数の\*はマンホイットニーのU検定により、生存率の\*\*はフィッシャーの正確確率検定により、5%水準で有意差があることを示す。n. s.は有意差がないことを示す。

表5. チョウセンキバナアツモリソウ自生地の土壌分析結果

Table5. Analysis of soil from the *Cypripedium guttatum* natural habitat.

分析項目	上の群落 (開花あり)			平均	標準偏差	下の群落 (開花なし)			平均	標準偏差
	サンプル①	サンプル②	サンプル③			サンプル④	サンプル⑤	サンプル⑥		
アンモニア態窒素 (mg/100g)	1.6	1.7	1.2	1.5	0.3	1.8	1.8	1.0	1.5	0.5
硝酸態窒素 (mg/100g)	1.0	1.2	1.5	1.2	0.3	0.9	1.3	0.7	1.0	0.3
可給態リン酸 (mg/100g)	1.0	2.1	2.0	1.7	0.6	3.0	2.2	1.3	2.2	0.9
交換性カリウム (mg/100g)	44.0	66.0	30.0	46.7	18.1	40.0	35.0	38.0	37.7	2.5
交換性カルシウム (mg/100g)	160.0	112.0	157.0	143.0	26.9	103.0	88.0	94.0	95.0	7.5
交換性マグネシウム (mg/100g)	40.0	17.0	82.0	46.3	33.0	43.0	98.0	9.0	50.0	44.9
可給態鉄 (PPm)	4.0	60.0	53.0	39.0	30.5	6.0	42.0	16.0	21.3	18.6
交換性マンガン (PPm)	6.0	16.0	11.0	11.0	5.0	6.0	4.0	9.0	6.3	2.5
塩分 (%)	0.006	0.006	0.03	0.0	0.0	0.008	0.059	0.07	0.0	0.0
pH	6.05	5.55	5.71	5.8	0.3	5.87	5.8	5.91	5.9	0.1
EC (μS/cm)	4	5	4	4.3	0.6	10	5	11	8.7	3.2

表6. アツモリソウ自生地とクマガイソウ自生地の土壌分析結果

**Table6.** Analysis of soil from the *Cypripedium macranthos* var. *speciosum* and *Cypripedium japonicum* Thunb. natural habitat.

試験項目	アツモリソウ自生地 (1株)		クマガイソウ自生地 (群生)	
	平均	標準誤差	平均	標準誤差
アンモニア態窒素 (mg/100g)	0.3 ± 0.1		15.0 ± 1.7	
硝酸態窒素 (mg/100g)	2.2 ± 1.8		11.3 ± 4.4	
可給態リン酸 (mg/100g)	1.5 ± 0.4		3.7 ± 0.9	
交換性カリウム (mg/100g)	42.5 ± 10.0		96.0 ± 11.6	
交換性カルシウム (mg/100g)	383.4 ± 40.6		41.3 ± 8.2	
交換性マグネシウム (mg/100g)	6.4 ± 3.4		36.0 ± 7.4	
可給態鉄 (ppm)	2.1 ± 0.6		22.7 ± 13.2	
交換性マンガン (ppm)	3.1 ± 1.6		9.7 ± 1.5	
塩分 (%)	0.003 ± 0.002		0.010 ± 0.001	
pH	7.6 ± 0.3		4.7 ± 0.1	
EC ( $\mu$ S/cm)	37.8 ± 23.3		48.3 ± 5.8	

以上の結果から、チョウセンキバナアツモリソウの種子を利用した無菌播種による発芽個体の作出と、発芽した苗の順化・鉢上げによる野外での育苗を経て、露地試験地に定植することによって開花および開花数の増加が確認できた。

## 2. 自生地の土壌分析 (試験 2)

表5にチョウセンキバナアツモリソウ自生地の土壌分析結果を示す。また、比較のためにアツモリソウとクマガイソウ自生地の土壌分析結果を表6に示す。

チョウセンキバナアツモリソウ自生地のU区とL区のアンモニア態窒素の平均値は、それぞれ1.5 mg/100 g, 1.5 mg/100 gであった。硝酸態窒素では1.2 mg/100 g, 1.0 mg/100 gとなり、U区とL区で大きな差は見られなかった。アンモニア態窒素と硝酸態窒素の合計値をアツモリソウ自生地と比較した場合、両者に大きな差は見られなかったが、クマガイソウ自生地と比較すると、チョウセンキバナアツモリソウとアツモリソウ自生地はクマガイソウ自生地の1/10程度になり低かった。

交換性カルシウムでは、143.0 mg/100 g, 95.0 mg/100 gとなり、アツモリソウ自生地よりも低くなったが、クマガイソウ自生地よりも高い値を示した。

U区のpHは平均5.8, L区は平均5.9となり、いずれも弱酸性土壌を示した。アツモリソウ自

生地の7.6よりも低くなったが、クマガイソウ自生地の4.7よりも高い値を示した。

本試験に使用する培地は、群落内の株数と開花数が多いU区に合わせて調整した。つまり、アンモニア態窒素と硝酸態窒素の合計値やpHに注目して培地作成時の窒素量とpHの調整を行った。

## 3. 本試験

### (1) 保存種子の発芽試験 (試験 3-1)

2014年採取種子の発芽率は2.4%, 2015年採取種子の発芽率は1.5%であり、それぞれ6個体と108個体の発芽個体が得られた。2015年採取種子と栽培地由来種子の発芽率を比較したところ、栽培地由来種子の方が有意に高かった(表7)。また、2015年採取種子の平均発芽開始日数は、203.4 ± 150.4日(平均 ± 標準偏差)となり、栽培地由来種子の126.3 ± 22.7日と比べて有意に長く、また、標準偏差も大きく、保存種子は、発芽率が低いとともに、発芽の開始も遅く、発芽開始日数のばらつきが大きかった(表7, 図4)。

### (2) 保存種子の育成培養 (試験 3-2)

表8に育成培養完了時の保存種子の培養苗(写真2)と栽培地由来種子の培養苗の草丈、根長、根数、越冬芽数および重量を示す。その結果、根数に有意な差が見られ、保存種子の培養苗の根数は3.5 ± 0.2本(平均 ± 標準誤差)と、栽培

表7. 小山田培地に播種した保存種子と栽培地由来種子の発芽開始日と発芽率の比較

**Table7.** Germination start dates and germination rates of preserved seeds vs. seeds from cultivated plants sown in Oyamada medium.

	発芽開始日数 (日)	発芽率 (%)
保存種子	203.4 ± 150.4 (108) <sup>2</sup>	1.5 (7265)
栽培地由来種子 (保存なし)	126.3 ± 22.7 (312)	34.7 (900)
有意性 <sup>3</sup>	**	**

<sup>2</sup> 平均±標準偏差 (供試数)

<sup>3</sup> 発芽開始日数の\*\*はマンホイットニーのU検定により、発芽率の\*\*はフィッシャーの正確確率検定により、1%水準で有意差があることを示す。

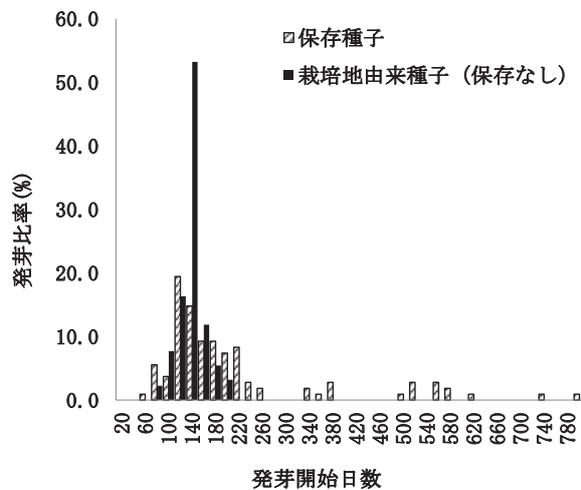


図4. 小山田培地に播種した保存種子と栽培地由来種子の発芽開始日と発芽率

**Fig.4.** Germination start dates and germination rates of preserved seeds vs. seeds from cultivated plants sown in Oyamada medium.

地由来種子の培養苗の  $4.7 \pm 0.3$  本より有意に少なかった。

育成培養完了時の保存種子の培養苗と栽培地由来種子の培養苗の生存率は、それぞれ 52.8% と 31.7% で保存種子の培養苗が有意に高かった。これらのことから、保存種子の培養苗は出根数が少ないが、生存率が高い苗であることが示唆された。

### (3) 保存種子の順化・鉢上げ (試験 3-3)

保存種子由来の培養苗の順化・鉢上げ 2 年後の生存率は 38.6% で、栽培地種子由来の培養苗の生存率と比較し、有意な差は見られなかった。

## IV. 考察

### 1. さく果と種子数の調査

本種の種子形成に関する情報を確認することができなかったことから、さく果の長さ、径、重量と種子数の関係を調査したところ、長さよりも径と重量で強い正の相関が見られた。特に、径が 10 mm 以上あるものや、重量が 0.4 g 以上に達しているものを選ぶことで、平均的な種子数を持ったさく果を選択できることが明らかになった。

これらの数値結果を参考にすることで、自生地等で結実が確認された際のさく果の選択が可能になると思われる。

### 2. 自生地の土壌分析

本種の自生地点を土壌分析した前例はなく、本研究の分析が国内初となる (小山田ほか, 2020)。群落の株数および開花数が多い U 区と、少ない L 区について、土壌分析結果の窒素、交換性カルシウムおよび pH の 3 項目を評価するとともに、アツモリソウとクマガイソウの自生地の分析結果 (小山田ほか, 2020) と比較して培地調整の参考とした (表 6)。

窒素成分は、U 区と L 区の両群落で少なく、交換性カルシウムは U 区でわずかに高い傾向にあり、U 区はクマガイソウよりもアツモリソウに近かった (小山田, 2020)。一般的な作物



写真2. 保存種子由来の苗の順化・鉢上げ時の様子  
**Photo2.** *Cypripedium guttatum* seedling germinated from a preserved seed just prior to acclimatization.

表 8. 保存種子由来の苗の順化・鉢上げ時の生育

Table 8. Growth of seedlings germinated from preserved seeds just prior to acclimatization.

	供試数	草丈 (mm)	根長 (mm)	根数 (本)	越冬芽数 (本)	重量 (g)
保存種子由来の苗	57	8.5 ± 0.3 <sup>2</sup>	51.1 ± 3.9	3.5 ± 0.2	0.8 ± 0.1	0.22 ± 0.10
栽培地由来種子の苗	30	8.9 ± 0.3	46.7 ± 2.2	4.7 ± 0.3	0.7 ± 0.1	0.23 ± 0.11
有意性 <sup>2</sup>		n. s.	n. s.	*	n. s.	n. s.

<sup>2</sup> 平均±標準誤差

<sup>2</sup> 草丈、根長および重量の\*はウェルチのt検定により、根数および越冬芽数の\*はマンホイットニーのU検定により、5%水準で有意差があることを示す。n. s. は有意差がないことを示す。

の畑土では、土壌中の好適な濃度範囲はアンモニア態窒素が5～15 mg/100 g、硝酸態窒素が5～10 mg/100 gとされており（富士平工業, 1988）、チョウセンキバナアツモリソウとアツモリソウの硝酸態窒素は5 mg/100 g未満と貧栄養土壌に自生すること、また、交換性カルシウムが両種ともに100 mg/100 g以上と、カルシウム成分の多い土壌に自生する傾向が見られた。

pHは、両群落とも弱酸性土壌であった。アツモリソウでは平均7.6で弱アルカリ性土壌を示し、クマガイソウでは平均4.7で酸性土壌を示したことから、チョウセンキバナアツモリソウは、その中間的なpHの土壌に自生していた。

硝酸態窒素との間に正の相関関係を示すECは（藤原ほか, 2005）、両群落とも低く、窒素成分の傾向と一致した。分析数は少ないが本種はアツモリソウに近く、クマガイソウよりも水溶性塩類を必要としない植物である可能性が認められた。

以上の分析結果を参考にして、本試験に用いる小山田培地（表2）の組成は、開花株の見られるU区の窒素量からアンモニア態窒素と硝酸態窒素の平均値の合計からアツモリソウに用いる培地と同量にした。また、pHの数値を5.8に決定した。

### 3. 発芽試験

本研究では、生息域外保全の推進のために必要となる自生地由来の保存種子を発芽させることを第一の目的にしている。保存種子の発芽に取り組む前に、予備試験として栽培株より採種した種子を用いて小山田培地に播種した。その

結果、発芽率は34.7%となり、過去に同じ方法で実施したアツモリソウの68%よりは低い発芽率となった（小山田ほか, 2011）。

本種の無菌播種による発芽報告は、国内外で3例がある（Bae et al., 2009; Ishii et al., 2005; 佐々木, 2000）。Baeは、チョウセンキバナアツモリソウの完熟種子を材料に、1%NaOClを用いた種子への刺激によって発芽が進むことを確認しており、発芽したプロトコームからの発根には、Ga 3（ジベレリン酸）を添加している（Bae et al., 2009）。Ishiiらは、1 μMのBAP（6-ベンジルアミノプリン）を含むMS培地で、未成熟種子のプロトコーム誘導を介して発芽を確認している（Ishii et al., 2005）。佐々木は、受粉から47日後の未熟種子を使用して、ホルモンを含まない1/3MS寒天培地で発芽を確認しており、その後、0.2 mg/LのBAPを添加した場合において複数のシュートを増殖させ、根の伸長にはホルモンを含まない培地が好ましいことを確認している（佐々木, 2000）。これら3者が行った無菌播種法では、フラスコ内での発芽およびシュートの増殖や発根についてまとめているものの、作出した苗の栽培段階での成長や開花は取り組まれていない。

Harvaisは、アツモリソウ属植物のレギナエ（*C. reginae*）の完熟種子を用いて、培地にサイトカイニンを添加すると実生苗作出に生育・生存率が高まることを示している（Harvais, 1982）。また、城らは、アツモリソウ（*C. macranthos* var. *speciosum*）、ホテイアツモリソウ（*C. macranthos* var. *hoteiatsumorianum*）、レブンアツモリソウ（*C. macranthos* var. *rebunense*）

のアツモリソウ属3種について、未熟種子を材料にBAPを0.2 mg/Lを添加した培地で発芽させている(城ほか, 1996)。以上のように、チョウセンキバナアツモリソウを含むアツモリソウ属植物の無菌播種による発芽では、植物成長調節物質(松本ほか, 1998; 日本植物培養学会・日本植物組織培養学会, 1993)を積極的に用いて発芽を進めている例が多いが、植物の組織培養で一般的に活用される植物成長調節物質の利用は、体細胞突然変異による培養植物の突然変異の発生を助長することが知られており(日本植物培養学会・日本植物組織培養学会, 1993)、絶滅危惧植物の野生復帰を目的にした利用には課題がある。

本研究に用いた保存種子は自生地から採種された種子を用いており、発芽した苗はクローン苗ではない。また、植物成長調節物質を含まない小山田培地を用いており、絶滅危惧植物の野生復帰を進める場合において、有効性が高いと言える。

2014年および2015年に採種された保存種子を材料に発芽試験に取り組んだ結果、培養開始50日から800日までの期間に断続的に発芽が確認され、発芽率は1.5%となり、栽培地由来の種子の発芽率34.7%と比較して有意に低くなった。両者はいずれも完熟種子を用いているが、栽培地の種子は乾燥処理を行わず採種から直ちに播種試験を行っており、保存種子は自生地で完熟種子を採取した後に保存処理を行っている。また、自生地のチョウセンキバナアツモリソウ群落は衰退傾向にあり、近交弱性や自生地で確認された植物体の生育不良による種子の結実不良の影響も懸念される。これらのことから発芽率に影響を及ぼしたものと推察されたが、今回の結果からはどのような原因によるものかについては究明できず、今後の課題である。また、Zengらはアツモリソウ属の未熟な種子は、成熟した種子よりも発芽が良いことが、多くの研究で示されていると報告しており(Zeng et al., 2014)、過去に未熟種子を用いたアツモリソウの発芽試験では、発芽率が68%であった(小山田ほか, 2011)が、本研究では材料が限られていたため、同培地による未熟種子の発芽試験

は行わなかった。今後、種子の熟度によるチョウセンキバナアツモリソウ発芽の知見を得ることも本種の生息域外保全に向けた一つの検討課題となる。

保存種子を試験開始前に顕微鏡観察したところ、微細な種子が縮れて変形しているものが多数見られた。また、種皮の中に観察できる未分化の胚は、どれも褐色をなしていた。本研究で行った発芽試験では、種子殺菌を行った後に、小山田培養液に液浸処理を行い、その後、小山田培地に種子を播種した。この時、培養液に液浸することで乾燥した種子の湿潤が種子の状態に応じて促進され、発芽日数のばらつきを生じさせたものと推察された。この培養液は、本来アツモリソウ専用として開発したものであるが(小山田ほか, 2011)、チョウセンキバナアツモリソウ自生地の分析結果を参考にpHの調整を行うことで、発芽や発芽後の育成培養、さらに順化・鉢上げ後の野外栽培において活用することができた。ラン科植物は、根の表皮部分に水滴が付着したり、湿った状態の所に接触すると、その部分の表皮が伸長して吸水能を高めることが知られている(農文協, 2003)ことから、発芽から野外での栽培まで培養液を継続的に与えたことで、アツモリソウと同様に成長に効果が認められたものと推察された。

#### 4. 育成培養

チョウセンキバナアツモリソウの発芽から苗の成長過程の変化を報告した例がないことから、本研究では、栽培地由来種子から発芽させた苗を観察し、苗の成長過程を4段階の生育レベルに区別した(図1)。このことで、苗の成長進度や成長区分を把握することが可能になり、その結果、保存種子から発芽させた苗の評価や、育成培地への継代時期の判断、さらに、順化・鉢上げの野外栽培へ移行する際の判断が容易になった。

苗の成長を進める育成培地への継代適期を確認するために、生育レベル2と生育レベル3の苗を育成培地に継代した結果、根長および生存率が生育レベル2区で高くなった。ラン科植物の培養では、培地中に活性炭を添加することに

より、成長阻害物質である老廃物を吸着することが知られており（富山，2000）、チョウセンキバナアツモリソウの育成培養においても培地に活性炭を添加し、生育レベル2の早い段階で育成培地に継代することでその効果が発揮されたものと推察された。

この結果を受け、環境省から依頼を受けた保存種子の発芽試験では、発芽した全ての苗を生育レベル2の段階で育成培地に継代し、栽培地由来種子の苗よりも高い生存率が得られた。一方で、根数は栽培地由来種子より少なくなり、この要因について不明であるが、チョウセンキバナアツモリソウの自生地では開花株が極めて少なく、自家受粉による結実の可能性が高い。以上のことより、種子の持つ多様性の差が影響しているものと推察した。

本研究によって、チョウセンキバナアツモリソウの発芽による苗の成長を進める場合、生育レベル2で育成培地へ継代することで、より高い生存率が得られることを明らかにした。

## 5. 順化・鉢上げ

現在、チョウセンキバナアツモリソウの発芽に関する報告は数例あるが、その中で順化・鉢上げについて具体的に報告したものはない。そこで、本研究では、我々が過去に実施したアツモリソウの順化法（小山田，2003；小山田ほか，2011）に沿って行った。この方法により、栽培地由来の苗について2年間の栽培管理を行ったところ、最終的に定植した30個体から7年後の2021年に304株の開花が確認された。これは本栽培管理によって地下茎（清水ほか，2001）の形成が進み、群生化することができたことを示す（図3）。このことからチョウセンキバナアツモリソウにおいても、アツモリソウの順化法が活用できることを確認した。

保存種子由来の苗は、栽培地由来の苗と同じ方法で順化・鉢上げを行ったが、栽培地由来の苗と大きく異なる生存率であり、本手法が保存種子由来の苗においても有効であることが示された。

## 6. チョウセンキバナアツモリソウ保護増殖事業

### の展望および今後の課題

チョウセンキバナアツモリソウは、絶滅の危険性が高い植物であり、種の保存対策が急務となっている（環境省自然環境局野生生物課，2018）。特に国内で確認できる自生地は1ヵ所のみで、このことがより緊急性を高めている（森，2002；高橋，2009）。

アツモリソウ属植物の場合は、「生息域外保全」を進める必要があるが、その前段階として、「生息域内保全」（環境省，2009）を進めることが急務である（自然環境研究センター，2011；小山田ほか，2021）。

そのためには、これまでに取り組んだ保護措置の効果と課題を科学的に検証し、より効果的な保全策を進める必要がある。「生息域外保全」については、新宿御苑において、種子保存に合わせて発芽確認を積極的に進めており、今後は希少植物の自生地復元や再導入を可能にするための種子保存方法の改良をめざす研究も進むと思われる。

本研究では、栽培地由来のチョウセンキバナアツモリソウの種子を用いてアツモリソウ属植物用に開発した小山田培地を用いて発芽を試みた。その結果、発芽個体を得ることに成功し、その苗を育苗して開花まで導いた。続いて環境省から依頼された保存種子を材料に発芽に取り組み、自生地の土壌分析結果を参考に調整した培地を用いて発芽させ、苗を環境省へ提出することができた（環境省，2020；写真3）。

保存種子7265粒から最終的に22個体の育成を導いたことになり、得苗率にすると0.3%となる。これは、栽培地由来種子900粒から45個体を導いた得苗率約5.0%と比較すると、「保存種子の活用」は「非保存種子の活用」よりも結果で劣っており、長期間にわたり保存された種子は、保存処理を行っていない種子と比較して得苗率で劣るという結果となった。これは種子の保存処理によって種子内の胚がダメージを受けているためと推察したが、さらに、現地個体の近交弱性や植物体の生育不良による種子の不全によって発芽・生育が低下していることも考えられる。

本研究では、乾燥冷凍保存したチョウセンキ

バナアツモリソウの種子からも苗の生産が可能であることを証明できた。作出した苗については、クローン増殖によるものではなく、環境省が「生息域外保全の実施に係る基本的な事項」として示す、「生息域外において保存される個体は可能な限り野生復帰させることが期待されるため、野生復帰させ得る資質を保つことが原則となる」という方針（環境省，2009）に適合しており、生息域外保全の苗として利用できるものと思われる。今後、保存種子からの苗が開花に至れば、将来的には自生地で衰退した野生株に代わって種子繁殖による遺伝的多様性の復元（自然環境研究センター，2011）にも活用が期待できる可能性があり、そうなる本研究で開発した発芽技術と本論文が重要な前例として役立つことができる。チョウセンキバナアツモリソウを含め国内に自生するアツモリソウ属植物の保護は緊急性が高いと考えられることから、本研究の取り組みが今後の種の保存対策に役立てば幸いである。

### 謝辞

環境省新宿御苑管理事務所温室第一科長・保護増殖専門官の関勝雄氏より、種子の保存方法について説明をいただいた。環境省東北地方環境事務所野生生物課生息地保護連携専門官の松

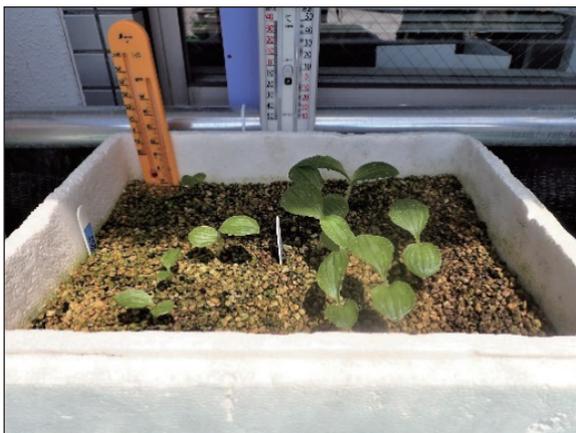


写真3. 環境省に提出したチョウセンキバナアツモリソウ保存種子由来の苗（2021年7月2日：秋田県立農業科学館）

**Photo3.** *Cypripedium guttatum* seedlings germinated from preserved seeds submitted to the Ministry of the Environment (July 2, 2021: Akita Prefectural Museum of Agricultural Science).

原豊氏と環境省秋田自然保護官事務所自然保護官の小笠原孝記氏より、研究のサポートや他機関との連絡調整に尽力をいただいた。本研究で作出したチョウセンキバナアツモリソウ苗の野外管理下の生育状況について、秋田県立農業科学館学芸班学芸主事の照井梓氏と、北海道大学北方生物圏フィールド科学センター植物園技術専門職員の永谷工氏より報告を受けた。ここに記して感謝を申し上げる。

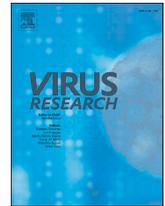
### 引用文献

- Bae, K. H., Kwon, H. K. & Choi, Y. E. (2009) In vitro Germination and Plantlet Conversion from the Culture of Fully Mature Seeds of *Cypripedium guttatum* swartz. *Propagation of Ornamental Plants* 9 (3) : 160-165.
- 富士平工業株式会社研究所 (1988) 土壌診断と施肥. 10-22. 富士平工業株式会社, 東京.
- 藤原俊六郎・安西徹郎・小川吉雄・加藤哲郎 (2005) 土壌肥料用語辞典. 78pp. 農文協, 東京.
- Harvais, G. (1982) An improved medium for growing the orchid *Cypripedium reginae* axenically, *Can.T. Bot.*, 60 : 2547-2555.
- Ishii, K., Maruyama, E., Hosoi, Y., Kanetani, S. (2005) IN VITRO PROPAGATION OF THREE ENDANGERED SPECIES IN JAPANESE FORESTS. *Propagation of Ornamental Plants* 5 (4) : 173-178.
- 城真一郎・秋田奈美樹・馬場徹代・筒井澄 (1996) アツモリソウ (*C.macranthum*) の無菌発芽における種子親による発芽率の変動, *園芸学会雑誌*, 65 別 2:632-633.
- 環境省 (2009) 絶滅のおそれのある野生動植物種の生息域外保全に関する基本方針.
- 環境省 (2020) 岩手県環境保健研究センターにおいてチョウセンキバナアツモリソウ自生地種子の発芽試験及び栽培を実施. <http://www.env.go.jp/nature/kisho/hogozoushoku/chosenkibanaatsumoriso.html> (参照：2020年3月17日).
- 環境省新宿御苑管理事務所 (2010) Seed Conservation Project リーフレット. 2, 51, 55pp.
- 環境省自然環境局 (2009) 絶滅危惧植物種子の

- 収集・保存等に関するマニュアル. 51, 54pp.  
環境省自然環境局野生生物課 (2018) 絶滅のおそれのある野生生物 - レッドデータブック -. 233pp.  
環境省自然環境局野生生物課 (2018) 絶滅のおそれのある野生動植物の種の保存に関する法律改訂法の施行リーフレット. 松本正雄・大垣智昭・大川清 (1998) 園芸辞典. 178pp. 朝倉書店, 東京.  
森和夫 (2002) 雲南の植物. 51. トンボ出版, 大阪.  
日本植物培養学会・日本植物組織培養学会 (1993) 組織培養辞典. 158, 190pp. 学会出版センター, 東京.  
農文協 (編) (2003) 花卉園芸大百科 15 ラン. 23-26. 農山漁村文化協会, 東京.  
小山田智彰 (2003) 地域希少植物 (アツモリソウ) の無菌播種. 植物バイオテックの実際. 131-135. 農山漁村文化協会, 東京.  
小山田智彰・平塚明・間山秀信 (2008) アツモリソウの種子発芽による苗の育成に関する研究. 自然環境復元研究 4, 43-50.  
小山田智彰・平塚明・鞍懸重和 (2011) ロールペーパーとバーミキュライトを培地支持材量に用いた絶滅危惧植物アツモリソウの苗生産に関する研究. 園芸学研究 10 (3) : 315-320.  
小山田智彰・鞍懸重和・千葉文也・佐藤香菜・長谷川啓一・古澤輝雄 (2019) 生息域外保全を目的としたチョウセンキバナアツモリソウの苗生産. 自然環境復元学会全国大会研究発表・講演要旨集, 25-28.  
小山田智彰・鞍懸重和 (2020) チョウセンキバナアツモリソウの自生地と栽培地の土壌分析. 自然環境復元学会全国大会研究発表・講演要旨集, 33-36.  
小山田智彰・鞍懸重和・高柳茂暢・吉田馨 (2021) 生息域内保全を目的にしたアツモリソウ野生株の移植と保全措置の有効性. 自然環境復元研究 12, 3-16.  
自然環境研究センター (編) (2011) 絶滅する前にできること. <http://www.env.go.jp/content/900491893.pdf> (参照: 2023年9月21日).  
佐々木揚 (2000) 日射量積算値による播種適期判定を利用したチョウセンキバナアツモリソウ (*Cypripedium guttatum*) の未熟種子培養. 植物工場学会誌 12 (4) : 268-274.  
Shefferson, R. P., Weiß, M. Kull, T. & Taylor, D. L. (2005) High specificity generally characterizes mycorrhizal association in rare lady's slipper orchids, genus *Cypripedium*. *Molecular Ecology* 14 : 613-626.  
清水建美 (2001) 植物用語辞典. 267pp. 八坂書房, 東京.  
高橋英樹 (2009) 日本産アツモリソウ属 *Cypripedium* (ラン科) の地理分布パターン. 分類 9 (2) : 143-157.  
富山昌克 (2000) ラン科植物のクローン増殖. 345pp. トンボ出版, 大阪.  
Zeng S., Zhang Y., Teixeira da Silva J.A., Wu K., Zhang J., Duan J. (2014) Seed biology and in vitro seed germination of *Cypripedium*. *Critical Reviews in Biotechnology*, 34 (4) : 358-371.

受付日: 2023年 2月 13日

受理日: 2023年 12月 2日



## Molecular evolutionary analyses of the fusion protein gene in human respirovirus 1

Tomoko Takahashi<sup>a</sup>, Mao Akagawa<sup>b</sup>, Ryusuke Kimura<sup>c,d</sup>, Mitsuru Sada<sup>b,c</sup>, Tatsuya Shirai<sup>c</sup>, Kaori Okayama<sup>b</sup>, Yuriko Hayashi<sup>b</sup>, Mayumi Kondo<sup>e</sup>, Makoto Takeda<sup>f</sup>, Akihide Ryo<sup>g</sup>, Hirokazu Kimura<sup>b,c,\*</sup>

<sup>a</sup> Iwate Prefectural Research Institute for Environmental Science and Public Health, Morioka-shi, Iwate 020-0857, Japan

<sup>b</sup> Department of Health Science, Gunma Paz University Graduate School of Health Sciences, Takasaki-shi, Gunma 370-0006, Japan

<sup>c</sup> Advanced Medical Science Research Center, Gunma Paz University Research Institute, Shibukawa-shi, Gunma 377-0008, Japan

<sup>d</sup> Department of Bacteriology, Gunma University Graduate School of Medicine, Maebashi-shi, Gunma 371-8514, Japan

<sup>e</sup> Department of Clinical Engineering, Faculty of Medical Technology, Gunma Paz University, Takasaki-shi, Gunma 370-0006, Japan

<sup>f</sup> Department of Microbiology, Graduate School of Medicine and Faculty of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

<sup>g</sup> Department of Microbiology, Yokohama City University School of Medicine, Yokohama-shi, Kanagawa 236-0004, Japan

### ARTICLE INFO

#### Keywords:

Human respirovirus 1  
Molecular evolutionary analyses  
Fusion protein gene  
B cell conformational epitope

### ABSTRACT

Few evolutionary studies of the human respiratory virus (HRV) have been conducted, but most of them have focused on HRV3. In this study, the full-length *fusion* (*F*) genes in HRV1 strains collected from various countries were subjected to time-scaled phylogenetic, genome population size, and selective pressure analyses. Antigenicity analysis was performed on the F protein. The time-scaled phylogenetic tree using the Bayesian Markov Chain Monte Carlo method estimated that the common ancestor of the HRV1 *F* gene diverged in 1957 and eventually formed three lineages. Phylodynamic analyses showed that the genome population size of the *F* gene has doubled over approximately 80 years. Phylogenetic distances between the strains were short (< 0.02). No positive selection sites were detected for the F protein, whereas many negative selection sites were identified. Almost all conformational epitopes of the F protein, except one in each monomer, did not correspond to the neutralising antibody (NT-Ab) binding sites. These results suggest that the HRV1 *F* gene has constantly evolved over many years, infecting humans, while the gene may be relatively conserved. Mismatches between computationally predicted epitopes and NT-Ab binding sites may be partially responsible for HRV1 reinfection and other viruses such as HRV3 and respiratory syncytial virus.

### 1. Introduction

Human respirovirus 1 (formerly called human parainfluenza virus 1, HRV1) is an RNA virus that belongs to the genus Respirovirus of the family Paramyxoviridae. HRV1 is a causative agent of acute respiratory diseases, such as common colds, acute laryngotracheobronchitis (croup), bronchiolitis, and pneumonia, and is distributed world-wide as the most prevalent type of the former human parainfluenza virus as well

as HRV3 (Henrickson, 2003; Karron, 2007). Epidemiological studies showed that HRV is a causative agent for croup in children under five years of age, among which approximately 26–74% experience HRV1 infection (Denny et al., 1983). HRV1 reinfection and HRV3 may occur throughout life; however, the reinfection mechanisms are not exactly known (Henrickson, 2003).

The HRV1 genome encodes six genes that are translated into seven proteins (Karron, 2007). Among these, fusion protein (F protein) and

**Abbreviations:** HRV, Human respirovirus; F protein, fusion protein; HN, haemagglutinin-neuraminidase; NT-Ab, neutralising antibodies; RSV, respiratory syncytial virus; *F* gene, *fusion* gene; BMCMC, Bayesian Markov chain Monte Carlo; ESS, effective sample sizes; HPDs, highest posterior densities; ML, marginal likelihood; BSPs, Bayesian skyline plots; *dN*, non-synonymous substitution rates; *dS*, synonymous substitution rates; SLAC, single-likelihood ancestor counting; FEL, fixed effects likelihood; IFEL, internal fixed-effects likelihood; FUBAR, fast unconstrained Bayesian approximation; MEME, mixed-effects model of evolution; 3D, three-dimensional; BRV, bovine respiratory virus.

\* Corresponding author at: Department of Health Science, Gunma Paz University Graduate School of Health Sciences, Takasaki-shi.

E-mail address: [kimhiro@nih.go.jp](mailto:kimhiro@nih.go.jp) (H. Kimura).

<https://doi.org/10.1016/j.virusres.2023.199142>

Received 22 January 2023; Received in revised form 26 April 2023; Accepted 31 May 2023

0168-1702/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

haemagglutinin-neuraminidase (HN) proteins are the major viral antigens (Karron, 2007). In particular, the F protein consists of a homotrimer and may be associated with infection of airway epithelial cells in the host (Karron, 2007). Moreover, the existence of two conformations of the F protein, prefusion and postfusion, have been confirmed (Yin et al., 2006). However, detailed F protein structure is not well understood (Shao et al., 2021).

Antibody responses are central to acquired immunity against viral infections. Epitopes are classified into two categories: conformational and linear epitopes (Van Regenmortel, 2001). Linear epitopes are continuous amino acid sequences of the primary antigen structure. Conformational epitopes are composed of discontinuous residues that are in proximity on the protein three-dimensional (3D) structure. Both epitopes are recognised by the immune system, triggering the production of antibodies (Sharon et al., 2014; Collins and Karron, 2013; Van Regenmortel, 2001). A previous report showed that over 90% of B cell epitopes are conformational, and only a few are linear (Van Regenmortel, 2001). In contrast, due to an explicit distinction between antigenicity and immunogenicity, these epitopes in antigenic proteins may not be adequately recognised by a neutralising antibody (NT-Ab) (Sharon et al., 2014; Collins and Karron, 2013). Our previous data suggested that the computationally predicted conformational epitopes in the respiratory syncytial virus (RSV) and HRV F proteins do not correspond to the NT-Ab binding sites of these proteins (Aso et al., 2020; Saito et al., 2021). However, to the best of our knowledge, the relationship between conformational and linear epitopes and the NT-Ab binding sites of the HRV1 F protein is not known. Moreover, the detailed phylogeny of the viral *fusion* (F) gene is unknown. Therefore, detailed molecular evolutionary analyses of the HRV1 F gene were performed in strains collected globally, using bioinformatic technologies.

## 2. Materials and methods

### 2.1. Strains used in this study

To better understand the molecular evolution of the HRV1 F gene, nucleotide sequences, including the full-length coding region of the gene (positions 5060–6727; 1668 nt for the hPIV1/USA/ATCC VR-94/1957 strain; GenBank accession No. JQ901971) was retrieved from GenBank on 11 June 2019. Among these, sequences from strains with confirmed information on the detected/isolated years and regions were selected. In addition, data of strains that displayed ambiguity with undermined sequences (e.g., N, Y, R, and V) were omitted from the dataset, providing data from 71 strains for the analysis. Homologous sequences were identified using Clustal Omega (Sievers and Higgins, 2021). When three or more similar sequences were present, only two among them were randomly retained, which reduced the final sequence set to those from 66 strains.

Temporal signal analysis of the sequences from 66 HRV1 strains was performed to determine whether the dataset was suitable for molecular clock analysis. Maximum likelihood method was used to generate a phylogenetic tree using molecular evolutionary genetics analysis version 7.0 (MEGA 7; for bigger datasets) software. The data were analysed using TempEst software (version 1.5.3) (Rambaut et al., 2016).

All data are presented in Supplementary Table S1. These sequences were aligned using MAFFT version 7 (Kato and Standley, 2013) and subsequently trimmed to 1668 nt based on the prototype F gene sequences.

### 2.2. Time-Scaled phylogenetic analysis and phylodynamic analyses using the bayesian markov chain monte carlo method (BMCMC)

To investigate the evolution of HRV1 strains, a time-scaled phylogenetic analysis of full-length sequences of the HRV1 F gene was conducted using the Bayesian Markov chain Monte Carlo (BMCMC) method

in BEAST (version 2.4.8) (Bouckaert et al., 2014). To select a suitable substitution model, jModelTest program (version 2.1.10) was used (Darriba et al., 2012). The path-sampling implemented in the BEAST package was used to determine the best of four clock models (strict clock, exponential relaxed clock, log-normal relaxed clock, and random local clock) and three prior tree models (coalescent constant population, coalescent exponential population, and coalescent Bayesian skyline). The TrN + I substitution model, log-normal of the relaxed clock model, and coalescent exponential from the tree prior model were used for BMCMC analysis of all strains. An BMCMC tree was constructed using BEAST software with the obtained strains and selected models. We analysed the MCMC chains for the 100,000,000 steps with sampling performed after every 2000 steps. To confirm convergence, the effective sample sizes (ESS) were evaluated using Tracer (version 1.6), and values above 200 were considered acceptable. After burn-in of the first 10% of the trees, a maximum clade credibility tree was generated using TreeAnnotator (version 2.4.8) in the BEAST package. The BMCMC phylogenetic tree was visualised using FigTree (version 1.4.03), and the 95% highest posterior densities (HPDs) of all internal nodes were computed. Moreover, strain clustering in the constructed phylogenetic tree of the HRV1 F gene followed the illustrated tree topology. Simultaneously, the evolutionary rates of the 66 HRV1 strains and strains of each lineage determined by the BMCMC phylogenetic tree were estimated using the BMCMC method, and the values were confirmed using the Tracer software. The marginal likelihood (ML) values for model selection and the detailed parameters of the BMCMC analyses are shown in Supplemental Tables S2 and S3. The statistics calculated by Tracer for each dataset are listed in Supplementary Tables S2–S6. Statistical analysis for comparing evolutionary rates between the lineages was performed using the Kruskal–Wallis test with the EZR software (Kanda, 2013). The evolutionary rates sampled every 2000 steps from the MCMC chains after discarding the 10% burn-in (45,001 samples) were used for statistical analysis. Statistical significance was defined as  $p < 0.05$ . Past genome population dynamics of the HRV1 F gene were examined using Bayesian skyline plots (BSPs) in BEAST. A coalescent Bayesian skyline was selected as the prior tree model.

### 2.3. Phylogenetic distance calculation

The phylogenetic distances among all HRV1 strains were analysed to estimate F gene diversity. A phylogenetic tree of all HRV1 strains was constructed using the ML method with MEGA7 software (Kumar et al., 2016), and branch reliability was supported by 1000 bootstrap replications. The jModelTest program was used to select the best substitution model for the ML method. Subsequently, the phylogenetic distance of the ML tree was calculated using Patristic (Fourment and Gibbs, 2006).

### 2.4. Selective pressure analyses

The selective pressure sites for the F protein of HRV1 were analysed by calculating the non-synonymous ( $dN$ ) and synonymous ( $dS$ ) substitution rates at each amino acid site using the Datamonkey web server (<https://www.datamonkey.org/>) (Weaver et al., 2018). Single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), internal fixed-effects likelihood (IFEL), fast unconstrained Bayesian approximation (FUBAR) (Murrell et al., 2013), and the mixed-effects model of evolution (MEME) (Murrell et al., 2012) were used to estimate positive selection sites, whereas, SLAC, FEL, IFEL, and FUBAR were used to predict negative selection sites. The positive ( $dN/dS > 1$ ) and negative ( $dN/dS < 1$ ) selection was determined based on the  $p$  values ( $p < 0.05$ ) for SLAC, FEL, IFEL, and MEME and on the posterior probability values ( $> 0.9$ ) for FUBAR.

### 2.5. Modelling of three-dimensional structure of the HRV1 F protein

Experimentally validated 3D structure of the HRV1 F protein is not

available. Hence, we employed a homology modelling method to construct trimeric structural models of the prefusion F protein of HRV1 for representative strains from each group, determined using the BMCMC phylogenetic tree (prototype, ATCC VR-94/USA/1957 strain, JQ901971; lineage 1, HPIV1/WI/629-008/1997 strain, JQ901978; lineage 2, HPIV1/WI/629-007/1997 strain, JQ901979; and lineage 3, HPIV1/USA/629-D02161/2009 strain, KF687308). The cryo-electron microscopy structure of HRV3 F protein (Protein Data Bank accession ID: 6MJZ) was selected as the template based on the results from BLAST web server (Shao et al., 2021). The amino acid sequences of each strain and template were aligned using MAFFT. The percentage sequence identity of each strain to the template was calculated using Clustal Omega. Based on the template sequence, 3D structures were constructed using Modeller software (version 10.2). The generated models were assessed by Ramachandran plot analyses using WinCoot implemented in the CCP4 package, and the models with the best scores were selected. Energy minimisation of the generated structures was performed using GROMOS96, which was implemented in Swiss PDB Viewer (version 4.1.0) (Guex and Peitsch, 1997).

## 2.6. Analyses of conformational and linear epitopes and amino acid substitution sites

To accurately analyse the pressure of human immune defence against the natural state of the HRV1 F protein, epitopes in the trimeric prefusion state were predicted. The conformational epitopes of the constructed models were analysed using Disco-Topo (version 2.0) (Kringelum et al., 2012), ElliPro (Ponomarenko et al., 2008), SEMA (Shashkova et al., 2022) and SEPPA (version 3.0) (Zhou et al., 2019) with cut-off values of  $-3.7$ ,  $0.5$ ,  $0.76$ , and  $0.064$ , respectively. The accuracy of the analyses was also supported by the consensus sites predicted by more than three of the four methods, and regions with residues close to two of the sites on the trimeric structure models were determined as conformational epitopes. Subsequently, the linear epitopes were analysed using LBtope (Singh et al., 2013) and BECEPS (Ras Carmona et al., 2021), BepiPred (version 2.0) (Galgonek et al., 2017) and ABCpred (Saha and Raghava, 2006). Cut-off values were set as 80%

(LBtope),  $0.5$  (BECEPS),  $0.5$  (BepiPred 2.0), and  $0.51$  (ABCpred), respectively.

Regions that had more than 10 continuous amino acids and were estimated in common by at least three of the four methods were regarded as linear epitopes. Finally, the predicted and previously identified epitopes were mapped onto the constructed, pre-fusion F protein models using PyMOL (version 2.3) (WL, 2002.).

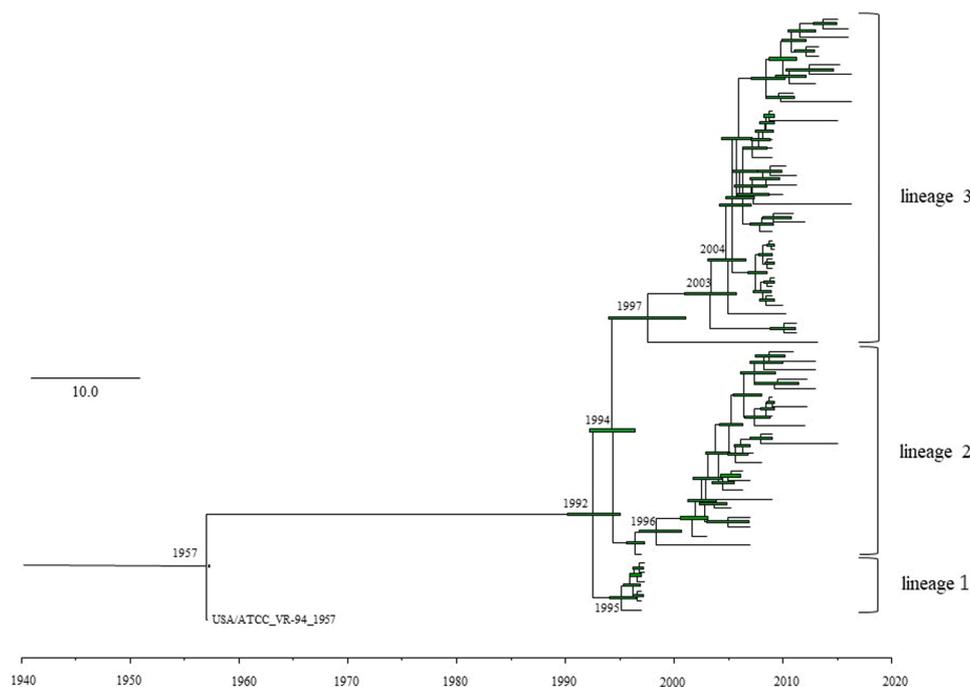
## 3. Results

### 3.1. Time-Scaled phylogenetic analysis and phylodynamic analysis of the HRV 1 F gene using the BMCMC method

To estimate the time-scaled evolution of the HRV1 F gene, a phylogenetic tree was constructed using the BMCMC method. In this study, we used only the sequences from HRV1 strains (66 strains) that were detected in humans, because sequence data from bovine respiratory virus (BRV) type 1, which may be a common ancestor of both BRV and HRV, were not available. Before constructing the BMCMC tree, the temporal signal of the dataset was estimated using TempEst (version 1.5.3). The plots of root-to-tip genetic distance against sampling time exhibited a positive correlation between genetic divergence and sampling time, and the R square value was calculated as  $0.87$  (Figure S1). These results suggest that the dataset of the 66 HRV1 strains was adequate for molecular clock analysis. Hence, we used this dataset to carry out the BMCMC method.

As shown in Fig. 1, a common ancestor of the HRV1 prototype strain (hPIV1/USA/ATCC\_VR-94\_1957; GenBank accession No. JQ901971) and other existing HRV1 strains diverged in 1957 (95% HPD, 1956–1957), resulting ultimately in the formation of three major lineages 1–3. After the first divergence in 1957, strains belonging to lineage 1 further diverged from a common ancestor of strains belonging to the three lineages in 1992 (95% HPD, 1989–1994), and the opposite side of lineage 1 diverged into lineages 2 and 3 in 1994 (95% HPD, 1991–1996). Currently, strains belonging to lineage 3 are widespread and form several clusters.

Next, the evolutionary rate of HRV1 F gene was estimated (Table 1).



**Fig. 1.** Time-scaled evolutionary tree of the full length Human respirovirus 1 (HRV1) fusion gene constructed by the Bayesian Markov chain Monte Carlo (BMCMC) method. The scale bar represents time (years). Green bars indicate the 95% highest posterior density (HPD) for each branch year.

**Table 1**  
Evolutionary rates of all HRV1 strains and each lineage.

	Evolutionary rates (95% HPD) (substitutions/site/year)	Effective sample size
All strains (66 strains)	$8.504 \times 10^{-4}$ ( $7.003 \times 10^{-4}$ to $1.0008 \times 10^{-3}$ )	220
Lineage 1 (6 strains)	–	–
Lineage 2 (23 strains)	$6.580 \times 10^{-4}$ ( $4.784 \times 10^{-4}$ to $8.4595 \times 10^{-4}$ )	4053
Lineage 3 (36 strains)	$1.205 \times 10^{-3}$ ( $7.159 \times 10^{-4}$ to $1.6866 \times 10^{-3}$ )	954

The evolutionary rate of all strains was estimated to be  $8.504 \times 10^{-4}$  substitutions/site/year (s/s/y) (95% HPD,  $7.003 \times 10^{-4}$  to  $1.0008 \times 10^{-3}$  s/s/y). The calculations for each of the above lineages showed that the evolutionary rate of strains belonging to lineage 2 was  $6.580 \times 10^{-4}$  s/s/y (95% HPD,  $4.784 \times 10^{-4}$  to  $8.4595 \times 10^{-4}$  s/s/y), and that of strains belonging to lineage 3 was  $1.205 \times 10^{-3}$  s/s/y (95% HPD,  $7.159 \times 10^{-4}$  to  $1.6866 \times 10^{-3}$  s/s/y). The evolutionary rate of the strains in lineage 1 with the same detection year (1997) was not calculated. The evolutionary rate of strains belonging to lineage 3 was significantly higher than that of strains belonging to lineage 2 ( $p < 2^{-16}$ ), which may indicate that strains belonging to lineage 3 are more adapted to humans, although the mechanisms underlying the significance of these values are not known.

### 3.2. Phylodynamics of the HRV1 F gene using the bayesian skyline plot (BSP) analysis method

As shown in Fig. 2, the phylodynamics of the F gene in HRV1 strains were analysed using the BSP analysis method to detect fluctuations in effective population size. The genome population size of all the strains doubled between 1995 and 2008 (Fig. 2A). Similarly, strains belonging to lineage 2 showed a two-fold increase in genome population around 2003 and 2008 (Fig. 2B). In contrast, in lineage 3, a steep increase in genome population size was observed, even though it occurred only once around 2008 (Fig. 2C). Because the detection year of the strains belonging to lineage 1 was the same (1997), we could not calculate the genome population size of this lineage. To summarise the results of these BSP analyses, the rapid increase in genome population size around 2008 for all strains was speculated to be mainly due to an increase in the genome population size of lineage 3.

### 3.3. Phylogenetic distances calculation of the HRV1 F gene

The phylogenetic distance and distribution between strains were evaluated based on their nucleotide sequences. A histogram of the distances between the sequence pairs of all the strains revealed a bimodal distribution (Fig. 3). Furthermore, histograms of lineages 1 and 2 showed bimodal distributions, although the apparent phylogenetic

distances of lineage 1 may not represent the phylogenetic distances of the actual lineage, owing to the small sample size (Fig. 3B and 3C). In contrast, the histogram of lineage 3 showed a unimodal pattern (Fig. 3D). The mean distance ( $\pm$  SD) between each pair of F gene sequence in the 66 HRV1 strains examined in this study was  $0.018575 \pm 0.01227$ . The results of the study showed that lineages 1, 2 and 3 had phylogenetic distances of  $0.0022 \pm 0.0012$ ,  $0.0073 \pm 0.0030$ , and  $0.0092 \pm 0.0058$ , respectively. Thus, the phylogenetic distance for each of the lineage was less than 0.02, suggesting conservation of the F gene sequence.

### 3.4. Homology modelling

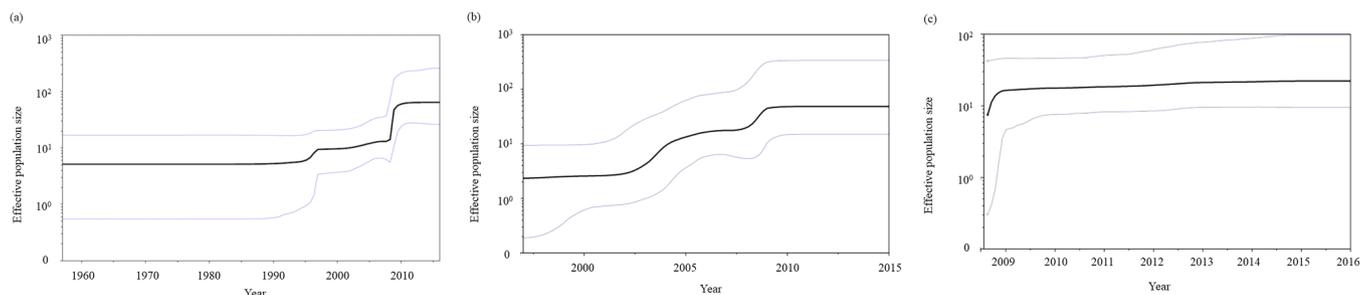
To visualise the relationship between NT-Ab binding sites and predicted B-cell conformational epitopes, we constructed the homotrimer (chains A, B, and C) of the HRV1 pre-fusion protein structure (Fig. 4). The amino acid sequence of the template covered amino acids 24–98 and 126–550 in each strain (Fig. 5). In this range, the amino acid residues on the protein surface were the same between representative strains of lineages 1–3. Moreover, from homologous analysis using Clustal Omega, the percentage sequence identity value of the prototype strain against the representative strains was high (96.8%). Hence, we presented the prototype structural model alone and showed the sites where amino acid substitutions occurred in representative strains from lineages 1 to 3 (Figs. 4 and 5). Both the prototype and lineage 1 representative strains had the same sequence identity (44.3%) as the template.

### 3.5. Selective pressure analyses of HRV1 F protein

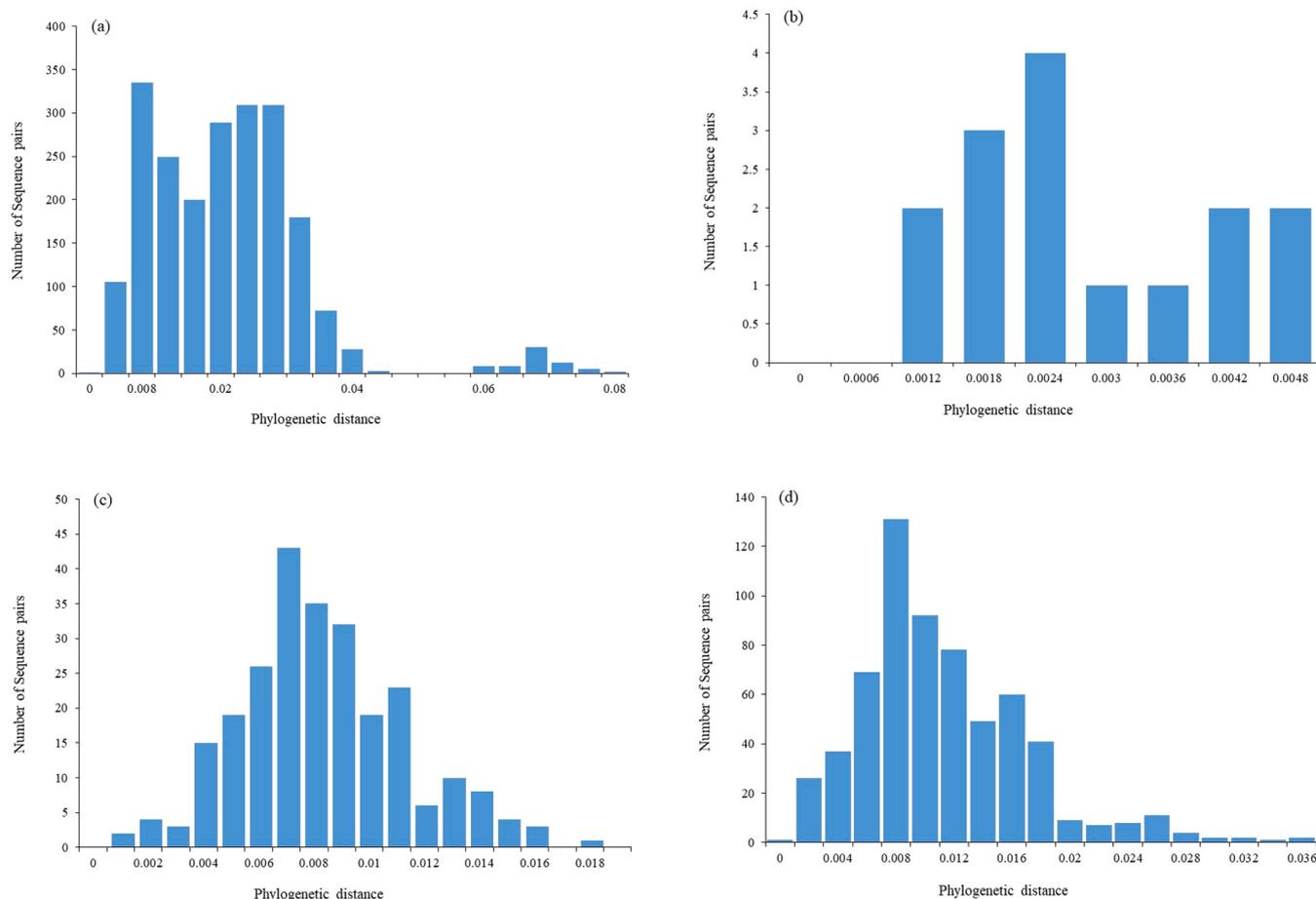
The rates of dS and dN substitutions were estimated using the DataMonkey web server to identify the positive and negative selection sites of F proteins in all 66 strains (Table S7). Only one method (FUBAR) predicted a positive selection site (amino acid residue 5), and the other four methods identified no positive selection sites. Thus, a positive selection site of the F protein is absent. In contrast, many negative selection sites were identified (Table S7). Among them, four negative selection sites (amino acids 150, 382, 460, and 473) were detected using all methods employed (Fig. 5).

### 3.6. Analyses of B-Cell epitopes and amino acid substitution sites

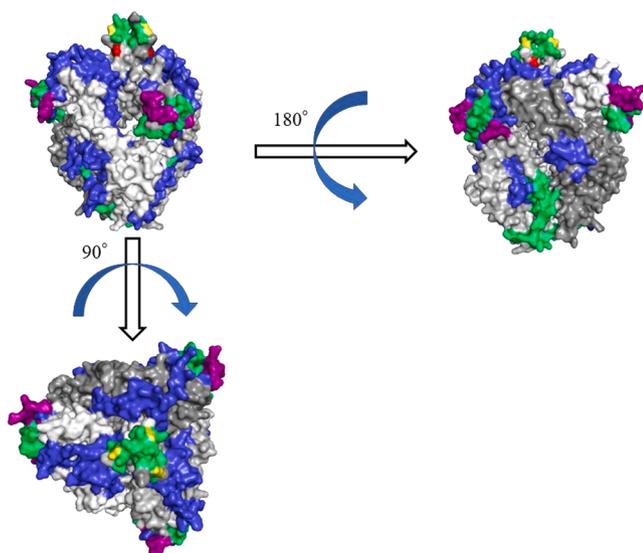
The amino acid substitution sites, NT-Ab binding sites, and predicted epitopes are shown in the HRV1 F chain A amino acid sequences and the trimeric structural model of the prototype (Figs. 4 and 5). First, the amino acid substitution sites in the F protein chain A among the prototype strain and the representative strains of lineages 1, 2, and 3 were compared. Seventeen amino acid substitutions were common in lineages 1, 2, and 3 (Fig. 5 and Table S8). An amino acid substitution unique to lineage 1 is present in Glu5Lys. Moreover, four amino acid residues



**Fig. 2.** Bayesian skyline plot for the Human respirovirus 1 (HRV1) fusion gene. Each panel illustrates the phylodynamics of all 66 strains (a), lineage 2 (b), and lineage 3 (c). Y and x-axes indicate the effective population size and time in years, respectively. Thick black lines show the median values over time; thin blue lines represent the 95% highest posterior density (HPD) intervals.



**Fig. 3.** Distribution of phylogenetic distances between the full-length sequences of the *fusion* gene of all Human respirovirus 1 (HRV1) strains. Each panel illustrates the histogram of all 66 strains (a), lineage 1 (b), lineage 2 (c), and lineage 3 (d). The y-axis and x-axis indicate the number of sequence pairs and phylogenetic distances, respectively.



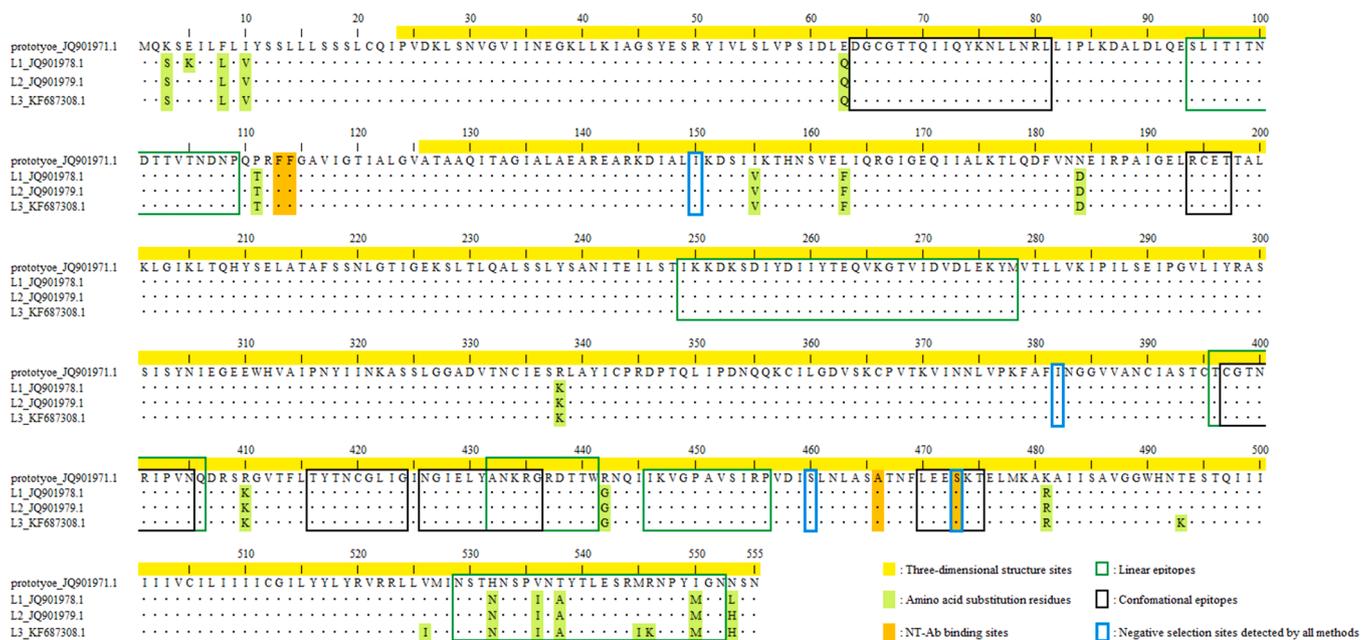
**Fig. 4.** Structural models of the prefusion protein of USA/1957 strain. Chains of the trimeric structures are given in white (chain A), light grey (chain B), and dark grey (chain C). The conformational and linear epitopes have been indicated in green and blue, respectively. Areas where the conformational and linear epitopes overlap are shown in purple. Experimentally identified epitopes (neutralising antibody binding sites) are indicated in red, and those overlapping with the conformational epitope are coloured yellow.

(Thr493Lys, Val526Thr, Met545Ile, and Arg546Lys) showed substitutions that were unique to lineage 3. However, some of the common substitution regions of these amino acid substitution sites are not located on the surface of the 3D structure. Similarly, uncommon substitution residues between lineages 1 and 3 were not present in the 3D structural model. Thus, only seven residues (Glu63Gln, Ile155Val, Leu163Phe, Asn184Asp, Arg338Lys, Arg410Lys, and Arg442Gly) of the common substitution regions in each lineage were located on the surface of the 3D structure model.

Next, conformational and linear B-cell epitopes on HRV1 F protein were analysed. Six conformational epitope sites and seven linear epitope sites were identified for the prototype HRV1 F protein chain A (18 conformational epitopes and 21 linear epitopes in the trimeric structure) (Figs. 4 and 5 and Table S9). No amino acid substitutions were found at these sites of strains in lineages 1, 2, or 3, whereas only one residue substitution (Glu63Gln) was found near one of the predicted conformational epitope sites (Fig. 5). Notably, in the HRV1 F protein chain A, five of the six conformational epitope sites and all the linear epitope sites failed to coincide with the experimentally determined NT-Ab binding sites, whereas only one residue of the conformational epitope (aa 473) coincided (Fig. 5). This mismatch may be a possible mechanism by which HRV1 can reinfect humans.

#### 4. Discussion

Evolutionary studies of HRV have been reported (Bose et al., 2019; Mao et al., 2012; Mizuta et al., 2014), but most of these focus on HRV3 (Mao et al., 2012; Mizuta et al., 2014). A few reports regarding the HRV1



**Fig. 5.** The Human respirovirus 1 (HRV1) fusion (F) protein chain A amino acid sequences investigated in this study. The amino acid residue numbers used for constructing HRV1 F protein three-dimensional (3D) structure are highlighted in yellow. The amino acid residues of neutralising antibodies (NT-Ab) binding sites and mutation are highlighted in orange and light green, respectively. The black and green line boxes in the sequences are the conformational epitope and linear epitope sites, respectively. Blue line boxes in the sequence indicate negative selection sites detected by all four methods (single-likelihood ancestor counting, fixed effects likelihood, internal fixed-effects likelihood, and fast unconstrained Bayesian approximation).

*F* gene have been published involving domestic or partial *F* gene analyses (Ambrose et al., 1995; Aso et al., 2020). To study the detailed molecular evolution of the full-length *F* gene in HRV1 strains from various countries, we performed time-scaled phylogenetic, genome population size, and selective pressure analyses on the gene, as well as antigenicity analysis of the F protein. From the time-scaled phylogenetic tree, constructed using the BMCMC method, it was estimated that the common ancestor of the HRV1 *F* gene diverged in 1957 and that their progenies continuously evolved and formed three lineages (lineages 1–3, Fig. 1). Strains belonging to lineage 3 predominate in various countries. Second, phylodynamic analyses using the BSP method showed that the genome population size of the *F* gene doubled over approximately 80 years. Third, phylogenetic distances among the strains were short ( $< 0.02$ ; Fig. 3). Finally, no positive selection sites were detected in the F protein, whereas many sites were identified as negative selection sites. Moreover, five sites of the six conformational epitopes and all linear epitopes in each chain of the F protein lacked correspondence to the NT-Ab binding sites (Figs. 4 and 5). These results suggest that despite the apparent conservation, the HRV1 *F* gene has evolved over many years. Yet, the conformational and linear epitopes did not correspond to the NT-Ab binding sites in either the pre or postfusion forms of the protein. This mismatch may be partially responsible for HRV1 reinfection and may extend to related viruses, such as HRV3 and RSV.

A phylogenetic analysis of HRV1 *F* gene was performed using the BMCMC method. It revealed that this gene continuously evolved and formed three lineages with many clusters (Fig. 1). Our previous report showed full-length F protein genes in HRV1 among patients with acute respiratory infections in Eastern Japan during 2011–2015 (Tsutsui et al., 2017). Phylogenetic analyses using the ML method showed that HRV1 strains formed three lineages and that the lineage 3 strains were dominant during the investigation period (2011–2015). This finding is consistent with the results of the present study. However, the time-scaled phylogeny of the HRV1 *F* gene was not assessed in the previous study (Tsutsui et al., 2017). Here, time-scaled phylogenetic analyses were done using the BMCMC method. As a result, the

divergence year of a common ancestor and each lineage was estimated (Fig. 1). Although the analysed gene was distinct, our previous report suggested that the HRV1 haemagglutinin-neuraminidase (HN) glycoprotein gene (full-length) isolated from Yamagata prefecture, in northern Japan, was classified into two lineages and formed many clusters using different phylogenetic analysis methods, including neighbour-joining and ML methods (Mizuta et al., 2011, 2014). Thus, to the best of our knowledge, the present study may be the first time-scaled phylogenetic analysis of the HRV1 *F* gene based on full-length sequences from globally collected strains. However, the present study has some limitations, including, the relatively small number of strains used. This is due to the paucity of studies on HRV1 molecular epidemiology. Another limitation is selection bias owing to the limited number of countries studying HRV1 and HRV3 (Aso et al., 2020).

The evolutionary rate of the HRV1 *F* gene has also been estimated. The mean evolutionary rate was  $8.504 \times 10^{-4}$  s/s/y. This is similar to the rates reported for the *F* genes of HRV3 and RSV (Aso et al., 2020). Furthermore, the evolutionary rate of HRV1 strains belonging to lineage 3 was found to be faster than that of lineage 2, whereas a previous evolutionary study on the HRV3 *F* gene did not find this difference (Aso et al., 2020). Moreover, these findings were not found for the RSV *F* gene, a virus belonging to a different genus and species (Saito et al., 2021). Thus, these findings were only observed for HRV1 *F* gene, which to the best of our knowledge, is the first report of lineage differences in evolutionary rate. In addition, the rapid evolutionary rate reflects short generation times and/or strong positive selection, which may generate a phenotype that is more adapted to the host (Collins and Karron, 2013). Together, the strains belonging to lineage 3 were more adaptive to humans and could become dominant strains, although our study did not address the mechanisms underlying the difference in the evolutionary rate.

The mean phylogenetic distance of the *F* gene HRV1 strains was approximately 0.02 (Fig. 3). This agrees with our previous study of Japanese strains, which reported a phylogenetic distance for the HRV1 *F* gene of 0.026 (Tsutsui et al., 2017). Moreover, the distribution of distances in our study was similar to that in previous reports on the HRV3 *F*

gene (Aso et al., 2020) as well as the HRV3 *HN* gene (Mao et al., 2012; Takahashi et al., 2018). These data suggest that the diversity of various viruses carrying the *F* gene may be similar and restricted within each species. With its high conservation and pivotal role in entry, the HRV1 *F* protein can be an attractive target for prophylaxis therapy, as it does for RSV *F* protein (Battles and McLellan, 2019).

Phylogenetic analyses of HRV1 *F* gene (Fig. 2) showed that the genome population of the gene has doubled over approximately 80 years. These fluctuations corresponded to the emergence of strains belonging to lineages 2 and 3 (Figs. 1 and 2). Our previous report showed that the genome population size of HRV3 *F* gene increased only once between 2000 and 2010 (Aso et al., 2020). Thus, genome fluctuations between HRV1 and 3 were different.

The selective pressure analyses of the HRV1 *F* gene did not reveal any positive selection sites for the present strains, whereas many negative selection sites were identified. In general, positive selection sites reflect escape from host defence mechanisms, such as cellular or humoral immunity (Barreiro and Quintana-Murci, 2010). Thus, HRV1 *F* protein may not be affected by such defence mechanisms. Conversely, negative selection sites may act to prevent deterioration of antigenicity (Holmes, 2013; Loewe, 2008). These findings may reflect the essential role the HRV1 *F* protein plays in host cell infection. Similar findings have been reported for the HRV3 *F* protein (Henrickson, 2003; Tsutsui et al., 2017). The negative selection sites of the HRV1 *F* protein were clustered near predicted epitopes, based on the 3D structural modelling. This indicates that these sites play important roles, for example, in the cellular receptor binding domain.

Finally, we analysed the amino acid substitutions and conformational and linear epitopes in the HRV1 *F* protein and evaluated their relationship with the NT-Ab binding sites (Figs. 4 and 5). Notably, amino acid substitutions were not estimated for computationally predicted epitopes or NT-Ab binding sites. Furthermore, almost all experimentally determined NT-Ab binding sites were incompatible with the computationally estimated conformational epitopes and linear epitopes. Antibody responses play a pivotal role in virus neutralisation or elimination, and antigenicity and immunogenicity differ explicitly (Collins and Karron, 2013). Conformational and linear epitopes may stimulate the production of antiviral antibodies (Lo et al., 2021; Sharon et al., 2014). In contrast, predicted epitope sites on antigenic proteins may not be adequately recognised by NT-Abs, which might mean that these predicted epitopes have weak potential for producing NT-Abs (Collins and Karron, 2013; Lo et al., 2021; Sharon et al., 2014). Thus, incompatibilities between NT-Ab binding sites and predicted epitopes may indicate low immunogenicity of the *F* protein and may be partially responsible for HRV1 reinfection, as has been reported for reinfections with HRV3 and RSV (Aso et al., 2020; Saito et al., 2021). However, the paucity of antigenic/antibody complex structures may lead to a mismatch between the NT-Ab binding sites and the predicted epitopes. Hence, the interpretation of the computationally conducted epitope analyses in this study may be limited, although conformational and linear epitopes were investigated using multiple computational methods for increased accuracy.

## 5. Conclusions

In this study, molecular evolutionary analyses of HRV1 *F* gene were performed based on full-length sequences collected globally. The time-scaled phylogenetic tree generated by the BMCME method estimated that the common ancestor of the HRV1 *F* gene diverged in 1957, and that their progenies have continuously evolved and formed three lineages. The phylogenetic analyses using the BSP method showed that the genome population size of the *F* gene doubled over approximately 80 years. The phylogenetic distances among the strains were short with no positive selection sites. Almost all conformational and linear epitopes in the *F* protein did not correspond to NT-Ab binding sites. These results showed that the HRV1 *F* gene has evolved over many years, although the

gene may be relatively conserved. Moreover, incompatibility between predicted epitopes and the NT-Ab binding sites in both the pre and postfusion forms of the protein may be responsible for HRV1 virus reinfection, as well as reinfections with HRV3 and RSV.

## Funding

This work was supported by a commissioned project for Research on Emerging and Reemerging Infectious Diseases from the Japan Agency for Medical Research and Development (AMED; grant number JP23fk0108661).

## Data availability statement

The data sets generated and analysed during the current study are available from the corresponding author upon reasonable request.

## CRediT authorship contribution statement

**Tomoko Takahashi:** Methodology, Data curation, Writing – original draft. **Mao Akagawa:** Methodology, Visualization. **Ryusuke Kimura:** Methodology, Visualization. **Mitsuru Sada:** Methodology, Writing – original draft, Visualization. **Tatsuya Shirai:** Writing – original draft. **Kaori Okayama:** Writing – original draft. **Yuriko Hayashi:** Writing – original draft. **Mayumi Kondo:** Methodology. **Makoto Takeda:** Writing – review & editing. **Akihito Ryo:** Writing – review & editing. **Hirokazu Kimura:** Conceptualization, Writing – original draft, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

We thank Ms. Miki Kawaji for the skilful support in figure preparation.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.virusres.2023.199142](https://doi.org/10.1016/j.virusres.2023.199142).

## References

- Ambrose, M., Hetherington, S., Watson, A., Scroggs, R., Portner, A., 1995. Molecular evolution of the F glycoprotein of human parainfluenza virus type 1. *J. Infect. Diseases* 171 (4), 851–856.
- Aso, J., Kimura, H., Ishii, H., Saraya, T., Kurai, D., Matsushima, Y., Nagasawa, K., Ryo, A., Takizawa, H., 2020. Molecular evolution of the fusion protein (F) gene in human respirovirus 3. *Front. Microbiol.* 10, 3054.
- Barreiro, L.B., Quintana-Murci, L., 2010. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat. Rev. Genet.* 11 (1), 17–30.
- Battles, M.B., McLellan, J.S., 2019. Respiratory syncytial virus entry and how to block it. *Nat. Rev. Microbiol.* 17, 233–245.
- Bose, M.E., Shrivastava, S., He, J., Nelson, M.I., Bera, J., Fedorova, N., Halpin, R., Town, C.D., Lorenzi, H.A., Amedeo, P., Gupta, N., Noyola, D.E., Videla, C., Kok, T., Buys, A., Venter, M., Vabret, A., Cordey, S., Henrickson, K.J., 2019. Sequencing and analysis of globally obtained human parainfluenza viruses 1 and 3 genomes. *PLoS ONE* 14 (7), e0220057.
- Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10 (4), e1003537.
- Collins, P.L., Karron, R.A., 2013. Respiratory Syncytial Virus and Metapneumovirus. In: Knipe, D.M., Howley, P. (Eds.), *Fields Virology*, 6th ed. Lippincott Williams & Wilkins, Philadelphia, pp. 1086–1123.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9 (8), 772–772.

- Denny, F.W., Murphy, T.F., Clyde Jr, W.A., Collier, A.M., Henderson, F.W., Senior, R., Sheaffer, C., Conley III, W., Christian, R., 1983. Croup: an 11-year study in a pediatric practice. *Pediatrics* 71 (6), 871–876.
- Fourment, M., Gibbs, M.J., 2006. PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. *BMC. Evol. Biol.* 6 (1), 1–5.
- Galgonek, J., Vymetal, J., Jakubec, D., Vondrášek, J., 2017. Amino Acid Interaction (INTAA) web server. *Nucl. Acids. Res.* 45 (W1), W388–W392.
- Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* 18 (15), 2714–2723.
- Henrickson, K.J., 2003. Parainfluenza viruses. *Clin. Microbiol. Rev.* 16 (2), 242–264.
- Holmes, E.C., 2013. Virus evolution. In: Knipe, D.M., Howley, P. (Eds.), *Fields Virology*, 6th ed. Lippincott Williams & Wilkins, Philadelphia, pp. 286–313.
- Kanda, Y., 2013. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. *Bone Marrow Transpl.* 48 (3), 452–458.
- Karron, R.A., 2007. Parainfluenza viruses. *Fields Virol.* 1497–1526.
- Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30 (4), 772–780.
- Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS. Comput. Biol.* 8 (12), e1002829.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874.
- Lo, Y.-T., Shih, T.-C., Pai, T.-W., Ho, L.-P., Wu, J.-L., Chou, H.-Y., 2021. Conformational epitope matching and prediction based on protein surface spiral features. *BMC. Genomics* 22 (Suppl.2), 116–116.
- Loewe, L., 2008. Negative Selection. *Nat. Educ.* 1 (1), 59.
- Mao, N., Ji, Y., Xie, Z., Wang, H., Wang, H., An, J., Zhang, X., Zhang, Y., Zhu, Z., Cui, A., Xu, S., Shen, K., Liu, C., Yang, W., Xu, W., 2012. Human parainfluenza virus-associated respiratory tract infection among children and genetic analysis of HPIV-3 strains in Beijing, China. *PLoS. ONE* 7 (8), e43893.
- Mizuta, K., Saitoh, M., Kobayashi, M., Tsukagoshi, H., Aoki, Y., Ikeda, T., Abiko, C., Katsushima, N., Itagaki, T., Noda, M., Kozawa, K., Ahiko, T., Kimura, H., 2011. Detailed genetic analysis of hemagglutinin-neuraminidase glycoprotein gene in human parainfluenza virus type 1 isolates from patients with acute respiratory infection between 2002 and 2009 in Yamagata prefecture, Japan. *Virol. J.* 8, 533–533.
- Mizuta, K., Tsukagoshi, H., Ikeda, T., Aoki, Y., Abiko, C., Itagaki, T., Nagano, M., Noda, M., Kimura, H., 2014. Molecular evolution of the haemagglutinin-neuraminidase gene in human parainfluenza virus type 3 isolates from children with acute respiratory illness in Yamagata prefecture. *Japan. J. Med. Microbiol* 63 (Pt 4), 570–577.
- Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., Scheffler, K., 2013. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol. Biol. Evol.* 30 (5), 1196–1205.
- Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., Kosakovsky Pond, S.L., 2012. Detecting individual sites subject to episodic diversifying selection. *PLoS. Genet.* 8 (7), e1002764.
- Ponomarenko, J., Bui, H.-H., Li, W., Füsseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. ElliPro: a new structure-based tool for the prediction of antibody epitopes. *BMC. Bioinformatics* 9 (1), 1–8.
- Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G., 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus. Evol.* 2 (1) vew007-vew007.
- Ras Carmona, A., Pelaez Prestel, H.F., Lafuente, E.M., Reche, P.A., 2021. BCEPS: a Web Server to Predict Linear B Cell Epitopes with Enhanced Immunogenicity and Cross-Reactivity. *Cells* 10 (10).
- Saha, S., Raghava, G.P.S., 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 65 (1), 40–48.
- Saito, M., Tsukagoshi, H., Sada, M., Sunagawa, S., Shirai, T., Okayama, K., Sugai, T., Tsugawa, T., Hayashi, Y., Ryo, A., 2021. Detailed evolutionary analyses of the F gene in the respiratory syncytial virus subgroup A. *Viruses* 13 (12), 2525.
- Shao, N., Liu, B., Xiao, Y., Wang, X., Ren, L., Dong, J., Sun, L., Zhu, Y., Zhang, T., Yang, F., 2021. Genetic characteristics of human parainfluenza virus types 1–4 from patients with clinical respiratory tract infection in China. *Front. Microbiol.* 12.
- Sharon, J., Rynkiewicz, M.J., Lu, Z., Yang, C.Y., 2014. Discovery of protective B-cell epitopes for development of antimicrobial vaccines and antibody therapeutics. *Immunology* 142 (1), 1–23.
- Shashkova, T.I., Umerenkov, D., Salnikov, M., Strashnov, P.V., Konstantinova, A.V., Lebed, I., Shcherbinin, D.N., Asatryan, M.N., Kardymon, O.L., Ivanisenko, N.V., 2022. SEMA: antigen B-cell conformational epitope prediction using deep transfer learning. *bioRxiv*.
- Sievers, F., Higgins, D.G., 2021. The Clustal Omega Multiple Alignment Package. Vol. 2231. *Humana Press*, pp. 3–16.
- Singh, H., Ansari, H.R., Raghava, G.P.S., 2013. Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PLoS. ONE* 8 (5) e62216-e62216.
- Takahashi, M., Nagasawa, K., Saito, K., Maisawa, S.-i., Fujita, K., Murakami, K., Kuroda, M., Ryo, A., Kimura, H., 2018. Detailed genetic analyses of the HN gene in human respirovirus 3 detected in children with acute respiratory illness in the Iwate Prefecture. *Japan. Infection, Genetics. and Evolution* 59, 155–162.
- Tsutsui, R., Tsukagoshi, H., Nagasawa, K., Takahashi, M., Matsushima, Y., Ryo, A., Kuroda, M., Takami, H., Kimura, H., 2017. Genetic analyses of the fusion protein genes in human parainfluenza virus types 1 and 3 among patients with acute respiratory infections in Eastern Japan from 2011 to 2015. *J. Med. Microbiol* 66 (2), 160–168.
- Van Regenmortel, M.H., 2001. Antigenicity and immunogenicity of synthetic peptides. *Biologicals* 29, 209–213.
- Weaver, S., Shank, S.D., Spielman, S.J., Li, M., Muse, S.V., Kosakovsky Pond, S.L., 2018. Datamonkey 2.0: a modern web application for characterizing selective and other evolutionary processes. *Mol. Biol. Evol.* 35 (3), 773–777.
- WL, D., 2002. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA.
- Yin, H.-S., Wen, X., Paterson, R.G., Lamb, R.A., Jardetzky, T.S., 2006. Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. *Nature* 439 (7072), 38–44.
- Zhou, C., Chen, Z., Zhang, L., Yan, D., Mao, T., Tang, K., Qiu, T., Cao, Z., 2019. SEPPA 3.0—enhanced spatial epitope prediction enabling glycoprotein antigens. *Nucleic. Acids. Res.* 47 (W1), W388–W394.

## Article

# Molecular Evolutionary Analyses of the RNA-Dependent RNA Polymerase (*RdRp*) Region and *VP1* Gene in Human Norovirus Genotypes GII.P6-GII.6 and GII.P7-GII.6

Tomoko Takahashi <sup>1,2</sup>, Ryusuke Kimura <sup>3,4</sup>, Tatsuya Shirai <sup>3,5</sup>, Mitsuru Sada <sup>3,5</sup>, Toshiyuki Sugai <sup>6</sup> , Kosuke Murakami <sup>7</sup>, Kazuhiko Harada <sup>3</sup>, Kazuto Ito <sup>3</sup> , Yuki Matsushima <sup>8</sup>, Fuminori Mizukoshi <sup>9</sup>, Kaori Okayama <sup>1</sup>, Yuriko Hayashi <sup>1</sup>, Mayumi Kondo <sup>10</sup>, Tsutomu Kageyama <sup>11</sup>, Yoshiyuki Suzuki <sup>12</sup>, Haruyuki Ishii <sup>5</sup>, Akihideo Ryo <sup>13</sup>, Kazuhiko Katayama <sup>14</sup> , Kiyotaka Fujita <sup>1</sup> and Hirokazu Kimura <sup>1,3,10,\*</sup>

- <sup>1</sup> Department of Health Science, Graduate School of Health Sciences, Gunma Paz University, Takasaki-shi, Gunma 370-0006, Japan; tomo-takahashi@pref.iwate.jp (T.T.); okayama@paz.ac.jp (K.O.); hayashi@paz.ac.jp (Y.H.); fujita@paz.ac.jp (K.F.)
- <sup>2</sup> Iwate Prefectural Research Institute for Environmental Science and Public Health, Morioka-shi, Iwate 020-0857, Japan
- <sup>3</sup> Advanced Medical Science Research Center, Gunma Paz University Research Institute, Shibukawa-shi, Gunma 377-0008, Japan; m2220015@gunma-u.ac.jp (R.K.); shirait28@gmail.com (T.S.); rainbow\_orchestra716@yahoo.co.jp (M.S.); k\_harada@bishinkai.or.jp (K.H.); kzito@gunma-u.ac.jp (K.I.)
- <sup>4</sup> Department of Bacteriology, Graduate School of Medicine, Gunma University, Maebashi-shi, Gunma 371-8514, Japan
- <sup>5</sup> Department of Respiratory Medicine, School of Medicine, Kyorin University, Mitaka-shi, Tokyo 181-8611, Japan; h141@ks.kyorin-u.ac.jp
- <sup>6</sup> Department of Nursing Science, Graduate School of Health Science, Hiroshima University, Hiroshima-shi, Hiroshima 734-8551, Japan; tsugai@hiroshima-u.ac.jp
- <sup>7</sup> Department of Virology II, National Institute of Infectious Diseases, Musashimurayama-shi, Tokyo 208-0011, Japan; ko-mura@niid.go.jp
- <sup>8</sup> Caliciviruses Section, Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA; yuki4649m780@gmail.com
- <sup>9</sup> Department of Microbiology, Tochigi Prefectural Institute of Public Health and Environmental Science, Utsunomiya-shi, Tochigi 329-1196, Japan; mizukoshif01@pref.tochigi.lg.jp
- <sup>10</sup> Department of Clinical Engineering, Faculty of Medical Technology, Gunma Paz University, Takasaki-shi, Gunma 370-0006, Japan; kondo@paz.ac.jp
- <sup>11</sup> Center for Emergency Preparedness and Response, National Institute of Infectious Diseases, Musashimurayama-shi, Tokyo 208-0011, Japan; tkage@niid.go.jp
- <sup>12</sup> Division of Biological Science, Department of Information and Basic Science, Graduate School of Natural Sciences, Nagoya City University, Nagoya-shi, Aichi 467-8501, Japan; yossuzuk@nsc.nagoya-cu.ac.jp
- <sup>13</sup> Department of Virology III, National Institute of Infectious Diseases, Musashimurayama-shi, Tokyo 208-0011, Japan; aryo@niid.go.jp
- <sup>14</sup> Laboratory of Viral Infection Control, Graduate School of Infection Control Sciences, Ōmura Satoshi Memorial Institute, Kitasato University, Minato-ku, Tokyo 108-8641, Japan; katayama@lisci.kitasato-u.ac.jp
- \* Correspondence: kimhiro@nih.gov.jp; Tel.: +81-27-386-5648; Fax: +81-27-386-8594



**Citation:** Takahashi, T.; Kimura, R.; Shirai, T.; Sada, M.; Sugai, T.; Murakami, K.; Harada, K.; Ito, K.; Matsushima, Y.; Mizukoshi, F.; et al. Molecular Evolutionary Analyses of the RNA-Dependent RNA Polymerase (*RdRp*) Region and *VP1* Gene in Human Norovirus Genotypes GII.P6-GII.6 and GII.P7-GII.6. *Viruses* **2023**, *15*, 1497. <https://doi.org/10.3390/v15071497>

Academic Editor: Volker Lohmann

Received: 2 May 2023

Revised: 24 June 2023

Accepted: 29 June 2023

Published: 1 July 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** To understand the evolution of GII.P6-GII.6 and GII.P7-GII.6 strains, the prevalent human norovirus genotypes, we analysed both the *RdRp* region and *VP1* gene in globally collected strains using authentic bioinformatics technologies. A common ancestor of the P6- and P7-type *RdRp* region emerged approximately 50 years ago and a common ancestor of the P6- and P7-type *VP1* gene emerged approximately 110 years ago. Subsequently, the *RdRp* region and *VP1* gene evolved. Moreover, the evolutionary rates were significantly faster for the P6-type *RdRp* region and *VP1* gene than for the P7-type *RdRp* region and *VP1* genes. Large genetic divergence was observed in the P7-type *RdRp* region and *VP1* gene compared with the P6-type *RdRp* region and *VP1* gene. The phylodynamics of the *RdRp* region and *VP1* gene fluctuated after the year 2000. Positive selection sites in *VP1* proteins were located in the antigenicity-related protruding 2 domain, and these sites overlapped with conformational epitopes. These results suggest that the GII.6 *VP1* gene and *VP1* proteins evolved uniquely due to recombination between the P6- and P7-type *RdRp* regions in the HuNoV GII.P6-GII.6 and GII.P7-GII.6 virus strains.

**Keywords:** human norovirus; RNA-dependent RNA polymerase (*RdRp*) region; *VP1* gene; epitope mapping; molecular evolution

## 1. Introduction

Human norovirus (HuNoV) is a major causative agent of acute gastroenteritis in humans of all ages [1,2]. Previous epidemiological data suggest that HuNoV may be associated with 30–60% of patients with gastroenteritis [3–5]. Moreover, this agent has caused large outbreaks of food poisoning worldwide [6,7]. However, effective vaccines and antiviral agents are not available at present [7]. Therefore, this agent may be a public health concern [8].

The HuNoV genome is a single-stranded plus-sense RNA with an approximately 7.5 kb nucleotide sequence [9]. The genome contains three open reading frames (ORFs): ORF1, ORF2, and ORF3 [9]. ORF1 encodes six nonstructural proteins designated as nonstructural proteins (NS) 1/2–7 [9]. Of these, the NS7 region encodes the RNA-dependent RNA polymerase (RdRp) protein, while ORF2 and ORF3 encode structural proteins, such as viral protein (VP) 1 and VP2, respectively [7,9]. The VP1 protein acts as an antigen and also shows large antigenic variations [6], although it is not exactly known.

Previous genetic and molecular epidemiological studies have suggested that the HuNoV genome shows large genetic divergence [8]. Currently, HuNoV is classified into three genogroups: genogroup I (GI), genogroup II (GII), and genogroup IV (GIV), with many genotypes based on the genetic divergence of *VP1* [10]. Due to the relatively frequent recombination between ORF1 and ORF2 [11,12], dual nomenclatures such as GII.P6 (*RdRp* genotype)-GII.6 (*VP1* genotype) have been used, utilizing both the *RdRp* region and *VP1* gene for genotyping [10,13]. To date, 60 and 49 types of *RdRp* (P-types) and *VP1* genotypes have been identified, respectively [10]. Moreover, genogroups and genotypes may be associated with disease severity [14]. Furthermore, recombination between ORF1 and ORF2 has resulted in many chimeric viruses [12]. However, the role of these chimeric viruses remains unknown.

Molecular epidemiological data on HuNoV infections in humans suggest that certain GI and GII genotypes are prevalent [15]. These reports also show that GII HuNoV is more dominant than GI HuNoV [15]. Of these, some GII genotypes corresponding to *VP1* genotypes, such as GII.2, GII.3, GII.4, GII.6, and GII.17, are prevalent types [16–19]. However, these epidemiological data may not explain the reasons for the HuNoV epidemics.

Recently, authentic bioinformatic technologies have been used in population genetics, including the study of the evolution of various viruses [20]. Indeed, these methods may allow us to estimate the phylogeny, genome population, and antigenicity using three-dimensional antigen structures. Information that reflects viral evolution may contribute to a better response to these questions. We studied the molecular evolution of chimeric HuNoV, such as GII.P17-GII.17, GII.P2-GII.2, and GII.P16-GII.2, which have caused major outbreaks in many countries [21–23]. However, such studies have not been performed on other GII genotypes to better understand GII HuNoV. In norovirus infections, GII.4 is the predominant genotype worldwide. However, recombinant GII.6 strains have been circulating since the 1970s, with outbreaks reported in Japan in 2008–2009 and in the United States and Italy in 2014–2015, with an overall prevalence second only to GII.4 [18,24,25]. Although this genotype has been reported to play an important epidemiological role in norovirus outbreaks, the molecular epidemiological mechanism underlying these outbreaks has not been studied in detail. Moreover, GII.6 almost always displays P6- or P7-type *RdRp* genotypes [26]. Therefore, in this study, we performed a comprehensive molecular analysis of globally collected HuNoV GII.P6-GII.6 and GII.P7-GII.6 strains.

## 2. Materials and Methods

### 2.1. Strains Used in This Study

To analyse the molecular evolution of HuNoV GII.6, the complete genome sequences of HuNoV were downloaded from GenBank (last accessed on 28 December 2022). In total, 11,810 strains were collected. They were classified into genotypes using the Norovirus Typing Tool (Ver.2.0), and GII.6 strains were selected [10]. HuNoV GII.6 data collected from each local government public health institution were added to the dataset because the number of GII.6 strains available for analysis, especially GII.P6-GII.6 strains, was small. Strains with an uncertain sequence or an unclear year of collection or area were excluded. Finally, 141 strains belonging to GII.6 remained and were used to analyse the molecular evolution of VP1. Similarly, 141 strains belonging to HuNoV GII.6 were obtained and used to analyse the molecular evolution of *RdRp* region. Details of the strains used in this study are presented in Supplementary Table S1.

### 2.2. Time-Scaled Phylogenetic Analyses

To evaluate the molecular evolution of the present strains, phylogenetic trees of the HuNoV *RdRp* region (1530 bp, excluding the stop codon) and the *VP1* gene (1641–1650 bp, excluding the stop codon) were constructed using the Bayesian Markov chain Monte Carlo (MCMC) method in the BEAST package (v.2.6.7), as previously described [27,28]. First, the jModelTest2 program was used to determine the suitable substitution models [29]. Second, the path-sampling/stepping-stone sampling marginal likelihood estimation method was used to evaluate the best of the four clock models (strict clock, relaxed clock exponential, relaxed clock log normal, and random local clock) and the two prior tree models (coalescent constant population and coalescent exponential population). Although these were performed independently for *VP1* gene and *RdRp* region analyses, SYM-Γ-I, relaxed clock exponential, and coalescent exponential population were selected for the molecular evolutionary analysis of *VP1* gene. However, SYM-Γ, relaxed clock exponential, and coalescent exponential population were adopted for the molecular evolutionary analysis of *RdRp*. The lengths of the Bayesian MCMC chains and samples are listed in Table S2. Effective sample sizes (ESS) were calculated using Tracer and the convergence of all parameters was confirmed if the ESS was greater than 200. After a 10% burn-in, phylogenetic trees were generated using TreeAnnotator (v.2.6.7) and rendered using FigTree (v.1.4.0). Representative strains of each cluster were selected based on the most recent age of detection within the cluster. In addition, to compare the amino acid sequences of these representative strains, the GII.6 strain (AB039777) prototype was determined based on a previous report [10]. Molecular evolutionary rates were estimated using suitable models selected for each dataset, as described above. Statistical analyses were performed using the Kruskal–Wallis *t*-test for EZR [30].

### 2.3. Phylogenetic Distance Analyses

To calculate the phylogenetic distances among the strains, we used MEGA7 software [31]. The best substitution models were estimated using the jModelTest2 program. The phylogenetic distances between the present GII.6 strains were calculated from the pairwise maximum likelihood (ML) tree of the ML tree using the Patristic program [32].

### 2.4. Phylodynamic Analyses

To assess the phylodynamics of the GII.6 strains, the effective population sizes of the *RdRp* region and *VP1* gene were calculated using Bayesian skyline plot (BSP) analysis implemented in the BEAST package [27]. Similar to the Bayesian MCMC method, the best substitution and clock models were selected. A Bayesian skyline plot and the 95% highest probability density (HPD) were visualized using Tracer [33].

### 2.5. Selective Pressure Analyses

The non-synonymous ( $dN$ ) and synonymous ( $dS$ ) substitution rates at each amino acid site were calculated to identify the selective pressure sites for the *RdRp* region and *VP1* gene using the Datamonkey server (<https://www.datamonkey.org/> accessed on 8 October 2022) [34]. Five algorithms—single likelihood ancestor counting (SLAC), fixed-effects likelihood (FEL), internal fixed-effects likelihood (IFEL), the mixed-effects model of evolution (MEME) method, and the fast, unconstrained Bayesian approximation (FUBAR) method—were used to identify positively selected sites, and all of them except FUBAR were used to detect negatively selected sites. The significance level was set at  $p < 0.05$  for SLAC, FEL, IFEL, and MEME. Evidence of selective pressure for FUBAR was supported by a posterior probability  $> 0.9$ . In the positive selection analysis, sites common to more than four methods were regarded as positive selection sites, whereas in the negative selection analysis, sites common to more than three methods were considered negative selection sites.

### 2.6. Construction of the 3D Structure of *RdRp* and *VP1* Proteins

To compare the *VP1* and *RdRp* protein structures among genotypes, three-dimensional (3D) structural models of *VP1* and *RdRp* proteins were constructed for each genotype using homology modeling. First, 3D structural models of *VP1* in representative strains of each genotype (AB039777, LC122916, MH791993, MK956199, and JX989075) were generated using Protein Data Bank (PDB) ID: 6OTF as a template. Then, five models for each *VP1* genotype were generated using Modellar software (version 9.23) [35]. These models were evaluated by Ramachandran plot analysis using WinCoot implemented in the CCP4 package [36] and the best-scoring models were chosen. Finally, the energy of the selected models for each strain was minimized using GROMOS96 implemented in the Swiss PDB viewer (ver4.1.0) [37]. Using a similar procedure, the models of *RdRp* protein in each representative strain (AB039777, LC122916, LC760173, MK956199, and JX989075) (Table S1) were constructed using the crystal structure of *RdRp* (PDB ID:1SH0) as a template.

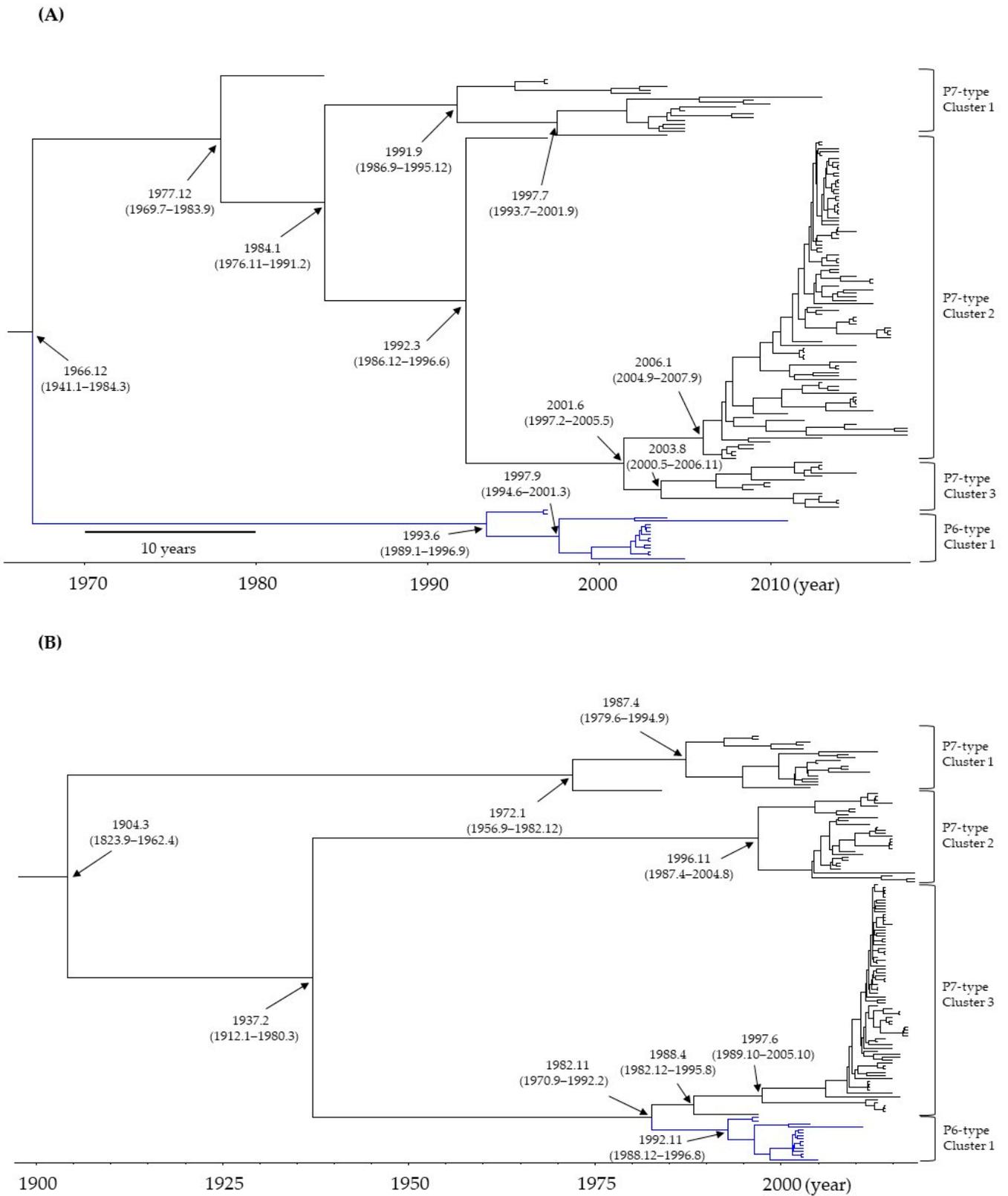
### 2.7. Conformational B-Cell Epitope Prediction

PDB files of the crystal structures of the GII.2 *VP1* protein (PDB ID:6OTF) and FASTA files of their amino acid sequences were downloaded from PDB (<https://pdj.org/?lang=ja> accessed on 30 August 2022) to use as templates in the homology modeling method. To assess the conformational B-cell epitopes of the constructed *VP1* protein models, four methods, DiscoTope 2.0 [38], ElliPro [39], SEMA [40], and SEPPA [41], were used, with cutoff values of  $-3.7$ ,  $0.5$ ,  $0.76$ , and  $0.064$ , respectively. Regions with amino acid sequences predicted by three or more of these methods and those contiguous with three or more residues were regarded as conformational epitopes. Furthermore, conformational epitopes were mapped onto the *VP1* protein models constructed above.

## 3. Results

### 3.1. Time-Scaled Phylogeny of the *RdRp* Region and *VP1* Gene in HuNoV GII.P6-GII.6 and GII.P7-GII.6

Time-scale phylogenetic trees were constructed based on the full-length nucleotides of the *RdRp* region and *VP1* gene using the Bayesian MCMC method. First, as shown in Figure 1A, a common ancestor of the P6- and P7-type *RdRp* regions diverged around December 1966 (mean; 95% HPDs, January 1941–March 1984). Subsequently, the P6- or P7-type *RdRp* regions further diverged and formed clusters 1 and 3, respectively. The main divergence times are shown in Figure 1A. The results suggested that a common ancestor of the P6- and P7-type *RdRp* region diverged approximately 50 years ago and evolved.



**Figure 1.** Time-scaled phylogenetic tree of the (A) *RdRp* region and (B) *VP1* gene in GII.P6-GII.6 (15 strains) and GII.P7-GII.6 strains (126 strains) of the human norovirus (HuNoV) constructed using the Bayesian MCMC method. The divergence times with 95% highest probability densities (HPDs) are indicated on the phylogenetic tree.

Next, as shown in Figure 1B, a common ancestor of the GII.6 VP1 gene diverged around March 1904 (mean; 95% HPDs, September 1823–April 1962). Thereafter, the genes diverged to form four clusters. The main divergence times are shown in Figure 1B. Finally, the GII.6 strains with the P6-type formed only one cluster, while the GII.6 strains with the P7-type formed three independent clusters. Furthermore, this phylogenetic tree estimated that a common ancestor of the P6- and P7-type VP1 genes diverged around November 1982 (mean; 95% HPDs, September 1970–February 1992). Thus, this time may be estimated as a recombination event between the GII.P6-GII.6 and GII.P7-GII.6 genomes in the present strains.

### 3.2. Evolutionary Rates of the RdRp Region and VP1 Gene in HuNoV GII.P6-GII.6 and GII.P7-GII.6

We also calculated the evolutionary rates using the Bayesian MCMC method. As shown in Table 1, the evolutionary rate was higher for GII.6 VP1 than the RdRp region, including P6- and P7-types (141 strains). The evolutionary rate was higher for the P6-type RdRp region than the P7-type RdRp. The evolutionary rate was higher for the P6-type GII.6 VP1 than the P7-type GII.6 VP1. These results suggest that the RdRp region and VP1 gene in the present strains evolved independently, and the evolutionary rates were significantly distinct.

**Table 1.** Evolutionary rates of the present GII.6 strains.

Region/Gene	Evolutionary Rates (95% HPDs) (Substitutions/Site/Year)	Compared Groups and Statistical Values
All RdRp region (141 strains) P6-type 15 strains; P7-type 126 strains	$3.287 \times 10^{-3}$ ( $2.489 \times 10^{-3}$ – $4.098 \times 10^{-3}$ )	All RdRp region vs. All GII.6 VP1 gene $p < 0.001$
All GII.6 VP1 gene (141 strains) P6-type 15 strains; P7-type 126 strains	$3.345 \times 10^{-3}$ ( $2.295 \times 10^{-3}$ – $4.419 \times 10^{-3}$ )	
P6-type RdRp region (15 strains)	$5.063 \times 10^{-3}$ ( $3.525 \times 10^{-3}$ – $6.595 \times 10^{-3}$ )	P6-type RdRp region vs. P7-type RdRp region $p < 0.001$
P7-type RdRp region (126 strains)	$3.022 \times 10^{-3}$ ( $2.268 \times 10^{-3}$ – $3.775 \times 10^{-3}$ )	
GII.P6-GII.6 VP1 gene (15 strains)	$3.725 \times 10^{-3}$ ( $1.843 \times 10^{-3}$ – $5.549 \times 10^{-3}$ )	GII.P6-GII.6 VP1 gene vs. GII. P7-GII.6 VP1 gene $p < 0.001$
GII.P7-GII.6 VP1 gene (126 strains)	$3.482 \times 10^{-3}$ ( $2.419 \times 10^{-3}$ – $4.568 \times 10^{-3}$ )	

### 3.3. Phylogenetic Distances among the Present Strains

To assess the genetic divergence of the RdRp region and VP1 gene in the present strains, we calculated their phylogenetic distances. The phylogenetic distances of the P6- and P7-type RdRp regions and the GII.6 VP1 gene were  $0.112 \pm 0.098$  (mean  $\pm$  1 standard deviation [SD]) and  $0.317 \pm 0.259$  (mean  $\pm$  1 SD). As shown in Figure 2A, the VP1 gene showed statistically greater genetic divergence than the RdRp region (unpaired *t*-test,  $p < 0.001$ ). Moreover, the genetic divergence was greater for the P7-type RdRp region than the P6-type RdRp region (unpaired *t*-test,  $p < 0.001$ ). The detailed statistical data are shown in Table 2.

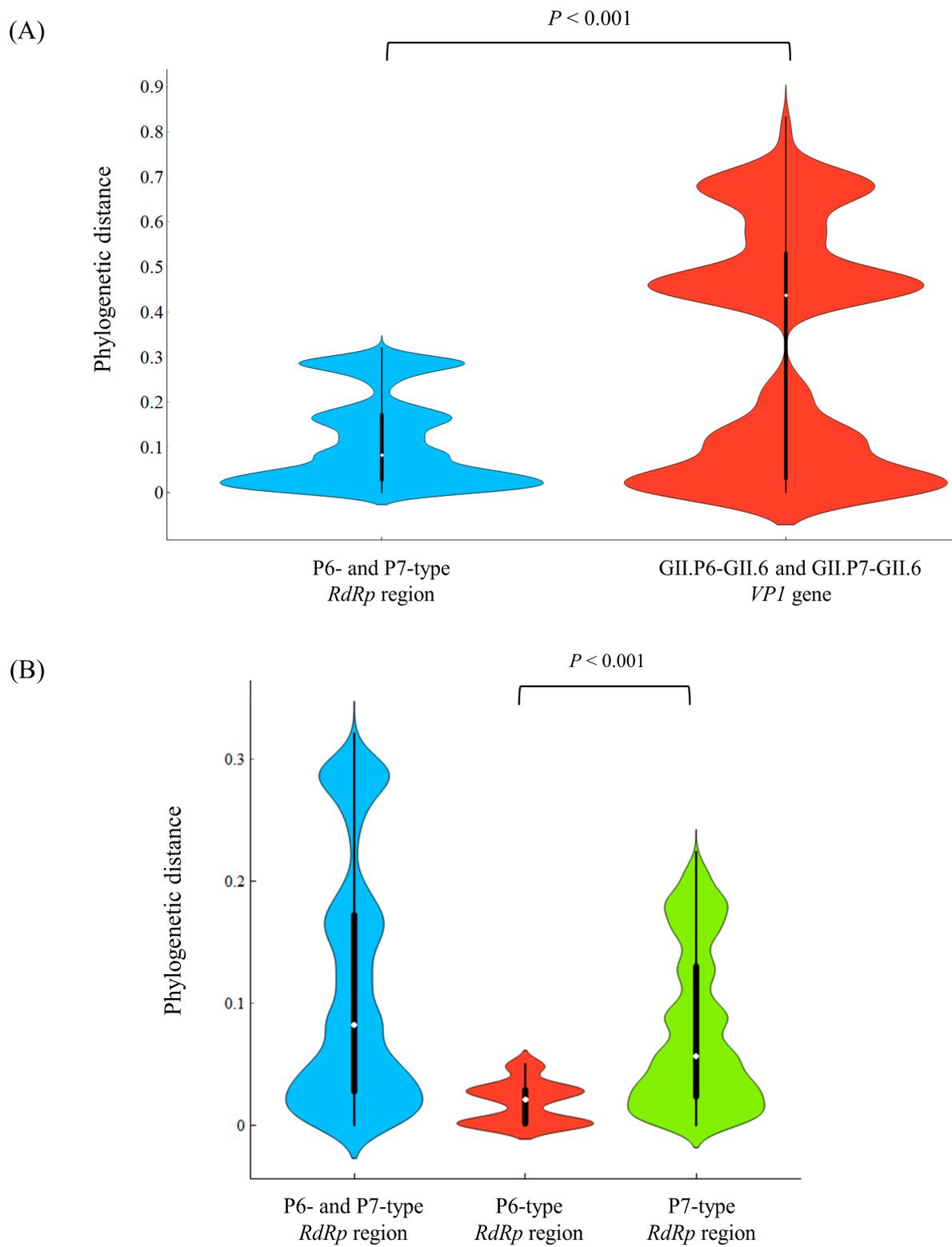
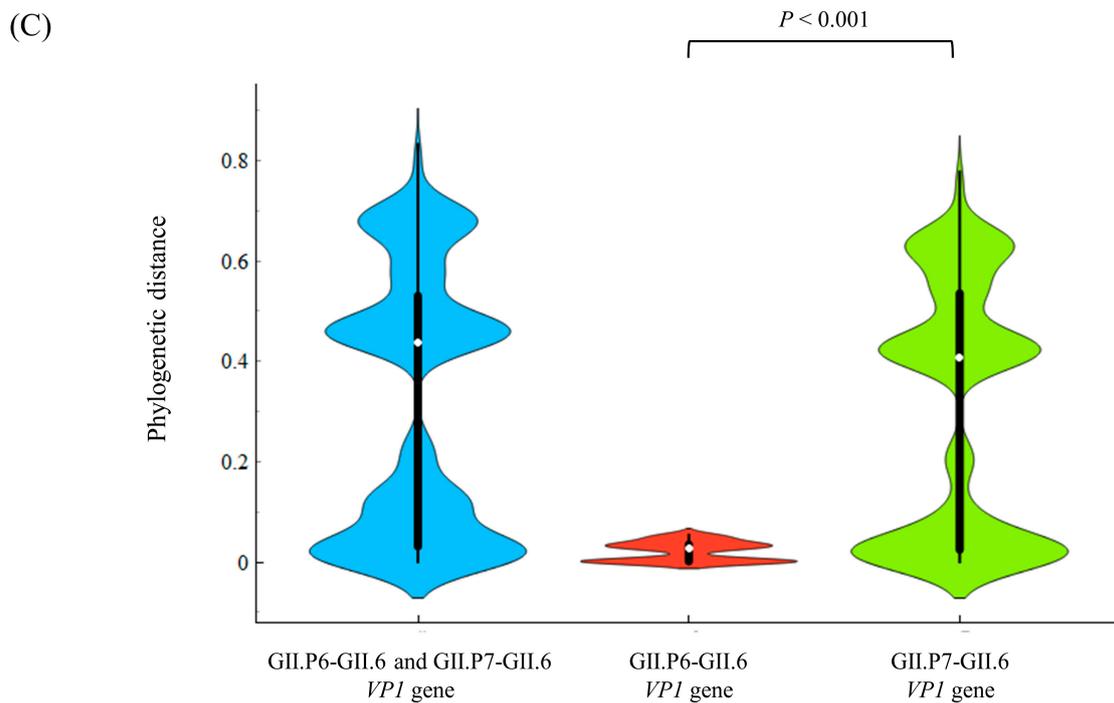


Figure 2. Cont.



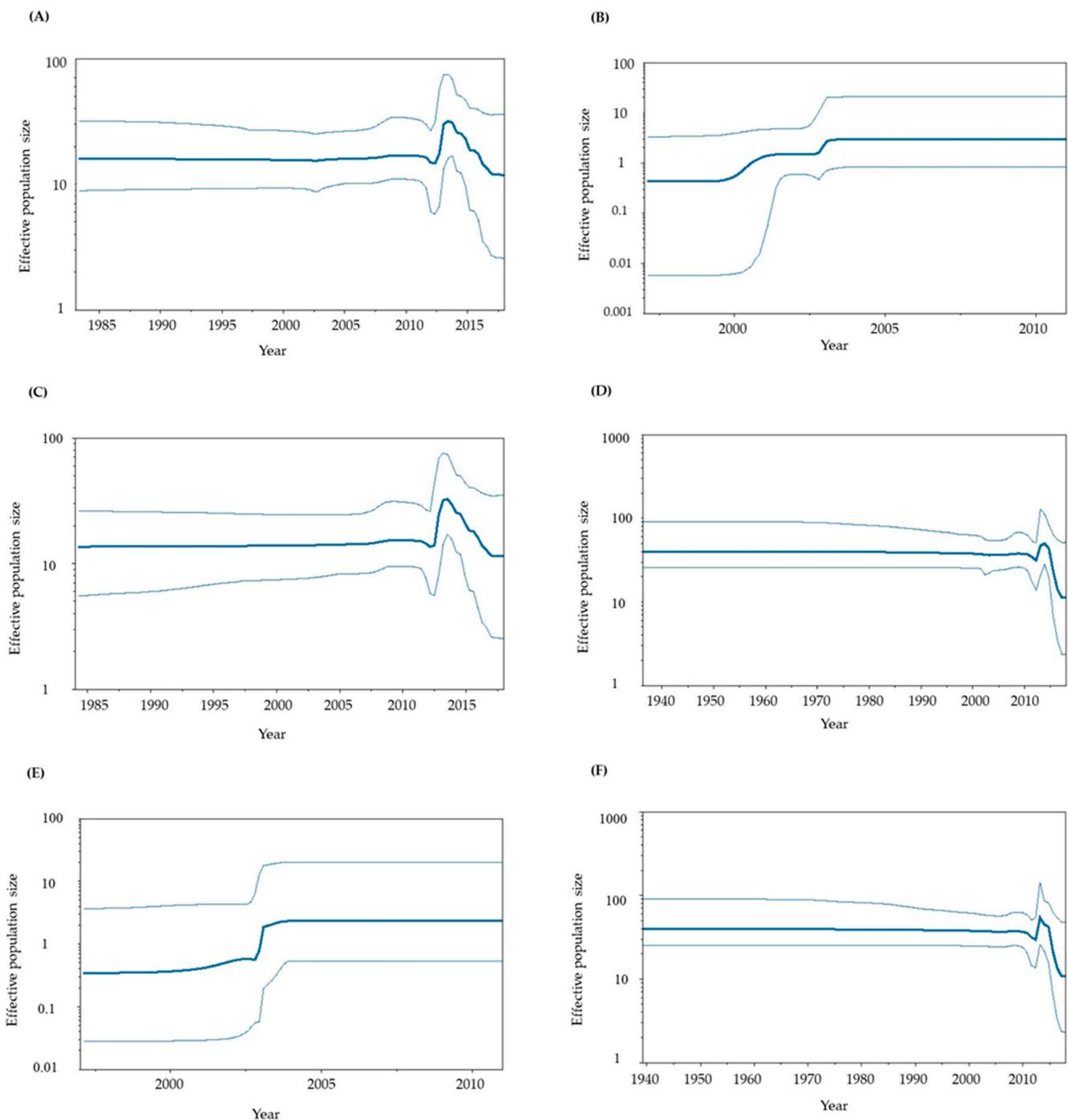
**Figure 2.** Phylogenetic distances of the (A) *RdRp* and *VP1* regions, (B) *RdRp* region, and (C) *VP1* gene in HuNoV GII.P6-GII.6 and GII.P7-GII.6, represented as violin plots. The width of the violin represents kernel density, indicating the distribution shape of the data. The central thick black bar and thin black line show the interquartile range and all data intervals, respectively. The white dot represents the median.

**Table 2.** Phylogenetic distance of the present strains.

Region/Gene	Phylogenetic Distance (Mean $\pm$ SD)	Phylogenetic Distance (Median [IQR])
All <i>RdRp</i> region (141 strains)	0.112 $\pm$ 0.098	0.082 (0.028–0.172)
All GII.6 <i>VP1</i> gene (141 strains)	0.317 $\pm$ 0.259	0.437 (0.031–0.530)
P6-type <i>RdRp</i> region (15 strains)	0.018 $\pm$ 0.017	0.021 (0.001–0.029)
P7-type <i>RdRp</i> region (126 strains)	0.078 $\pm$ 0.064	0.057 (0.023–0.130)
GII.P6-GII.6 <i>VP1</i> gene (15 strains)	0.021 $\pm$ 0.019	0.027 (0.001–0.035)
GII.P7-GII.6 <i>VP1</i> gene (126 strains)	0.305 $\pm$ 0.249	0.408 (0.024–0.535)

### 3.4. Phylodynamics of GII.P6-GII.6 and GII.P7-GII.6

To assess the phylodynamics of the present GII.P6-GII.6 and GII.P7-GII.6 strains, we calculated time-scaled genome population sizes using the BSP method (Figure 3A–F). Until approximately 2010, the genome population sizes of both the *RdRp* region and *VP1* gene remained constant. However, significant fluctuations in genome population sizes were observed at around 2010–2018.



**Figure 3.** Phylodynamics of the present HuNoV GII.P6-GII.6 and GII.P7-GII.6 strains determined using Bayesian skyline plot analysis. (A) Phylodynamics of the P6-type and P7-type *RdRp* regions; (B) P6-type *RdRp* region; (C) P7-type *RdRp* region; (D) P6- and P7-type *VP1* genes; (E) P6-type *VP1* gene alone; (F) P7-type *VP1* gene alone. The *y*-axis shows the effective population size for each distance, and the *x*-axis represents time (years). The thick line in the centre shows the median effective population sizes, and the thin lines at the top and bottom indicate the 95% HPDs.

### 3.5. Positive Selection Sites in the *RdRp* and *VP1* Proteins

We analysed the positive selection sites in the *RdRp* and *VP1* proteins to estimate selective pressure against the host. No positive selection site was detected in the *RdRp* protein. In contrast, a few positively selected sites were identified in *VP1*. Of these, only Lys386His was predicted in the P6-type *VP1* protein, whereas Pro354Thr, Pro354Ser, Pro354Gln,

Asn390Thr, and Asn390Asp were predicted in the P7-type VP1 protein. These sites were located in the protruding 2 (P2) domain of the protein (Table S4). These results suggest that the GII.P7-GII.6 strains may receive stronger selection pressure from the host than the GII.P6-GII.6 strains. Amino acid substitutions were also observed in sites other than the P2 domain, but no positive selection sites were identified.

### 3.6. Negative Selection Sites in RdRp and VP1 Proteins

In general, negative selection sites may prevent the deterioration of protein function. Therefore, we calculated the number of negative-selection sites in these strains. Many negative selection sites were estimated for the P7-type RdRp protein (205 sites) and P7-type VP1 protein (274 sites). However, a small number of negative selection sites were estimated in the P6-type RdRp protein (3 sites) and P6-type VP1 protein (8 sites) (Table 3). A few irregularly positioned negative selection sites in the P6-type RdRp and VP1 proteins were identified. Details of these negative selection sites are shown in Supplementary Table S3A,B. In both the RdRp and VP1 proteins, amino acid variations in the negative selection sites were often located in non-overlapping positions.

### 3.7. 3D Mapping Relationships between Amino Acid Substitutions and Active Sites of the RdRp Dimer Proteins

To better assess the relationships between amino acid substitutions and the active sites of RdRp proteins, we constructed 3D RdRp dimers and mapped them. Within RdRp, there were few amino acid substitutions in both the P6 and P7 types, but none were observed in the RdRp active sites (aa182, aa242, aa243, aa300, aa309, aa343, and aa344) (Figure 4A–E).

#### (A) Prototype: P7-type RdRp Protein (AB039777)

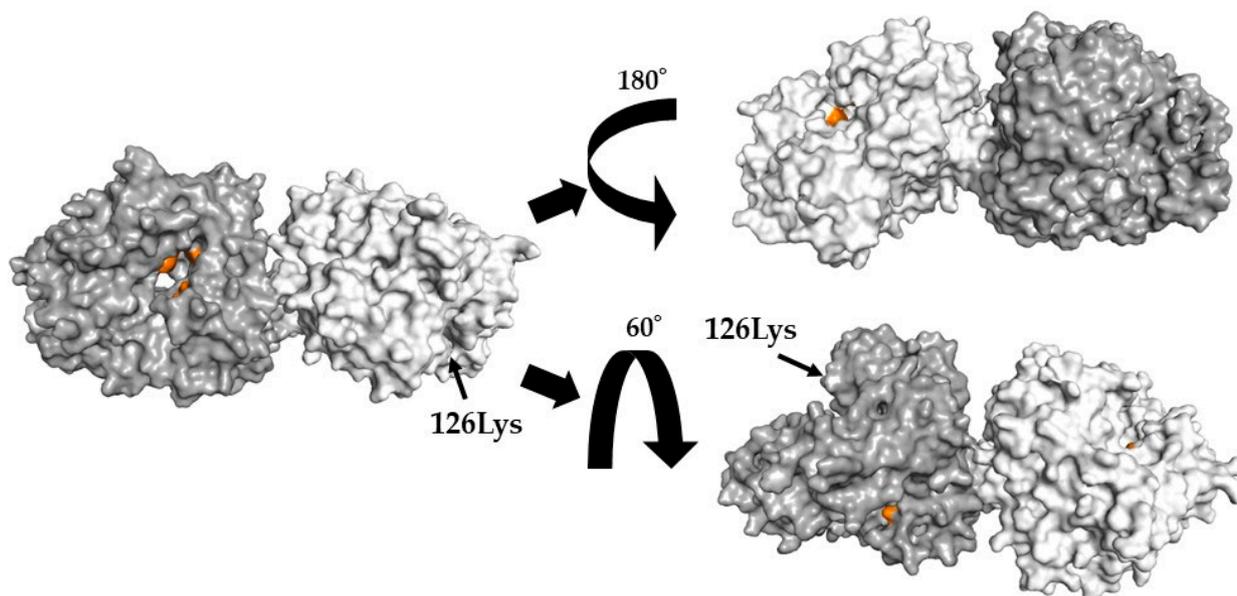
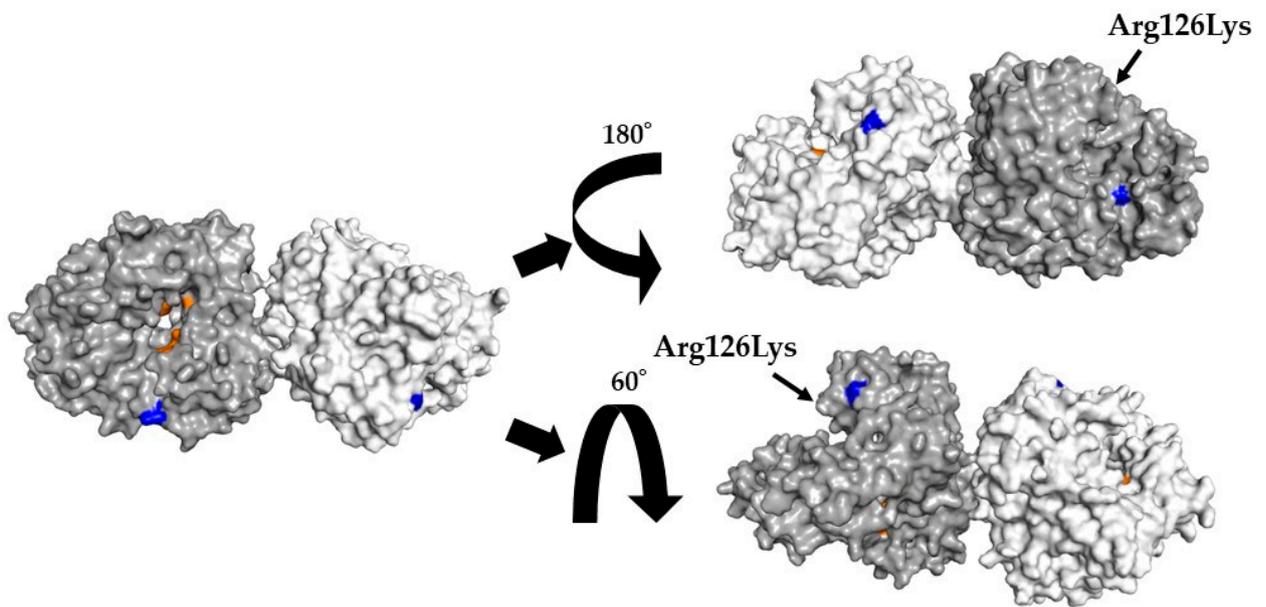
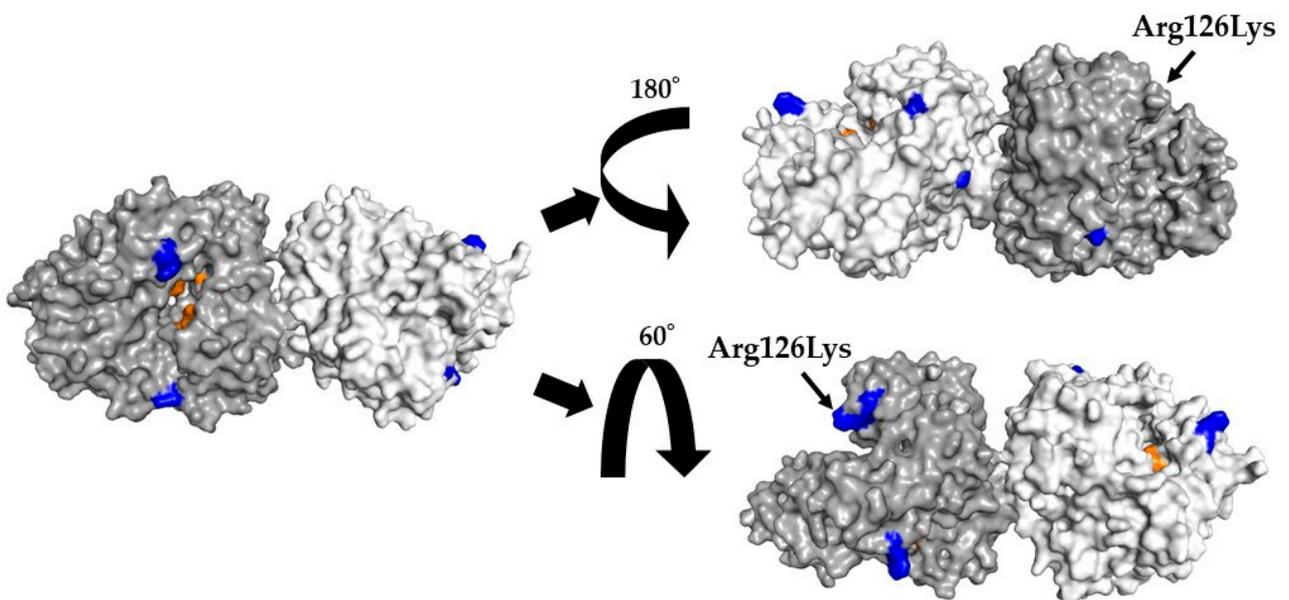
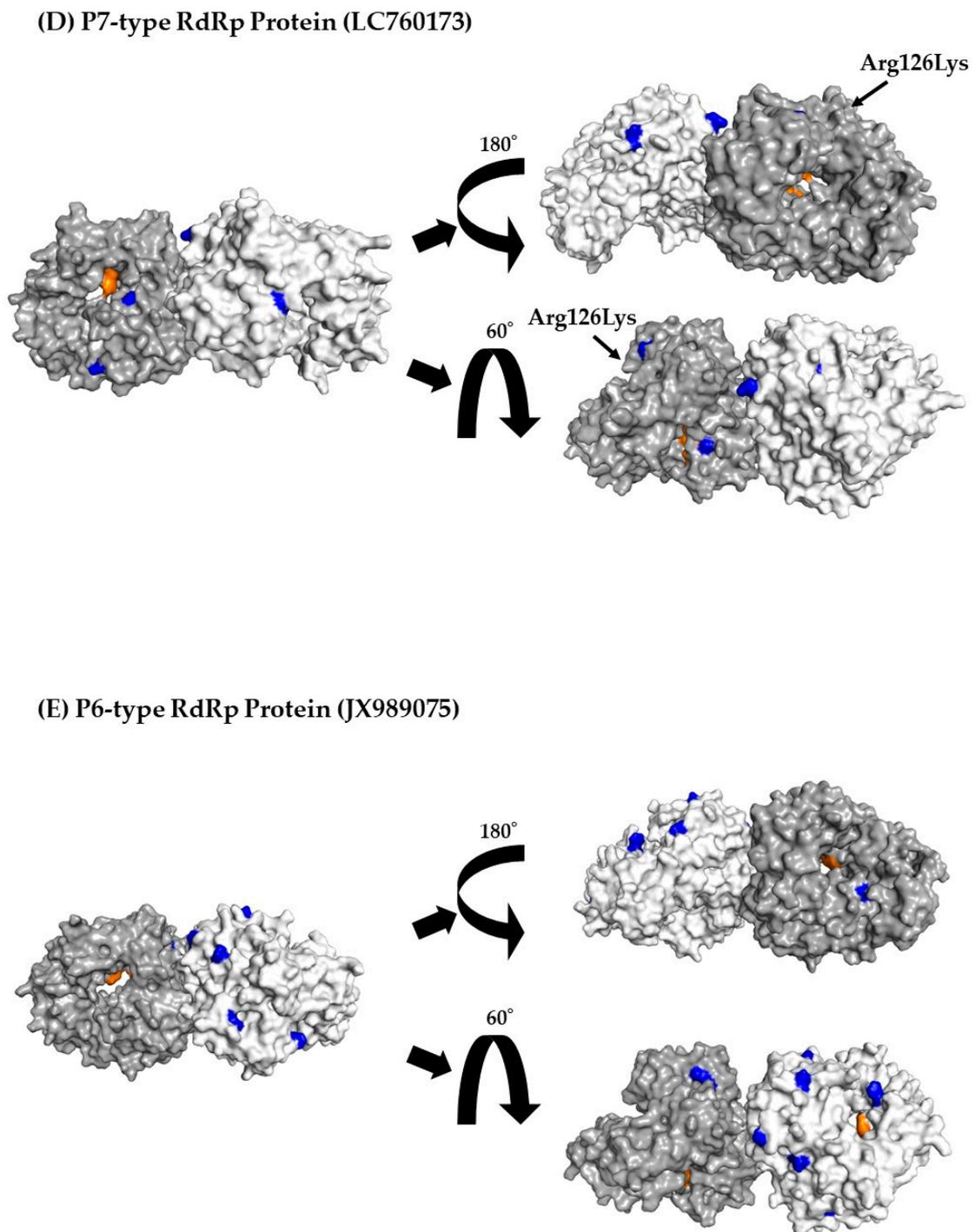


Figure 4. Cont.

**(B) P7-type RdRp Protein (LC122916)****(C) P7-type RdRp Protein (MK956199)**Figure 4. *Cont.*



**Figure 4.** Three-dimensional (3D) RdRp protein (dimer) structure and mapping of amino acid substitutions and active sites. Illustration shows the 3D structures of RdRp protein in the prototype and the most recent strain for each cluster. The strains in each figure are as follows: (A) a P7-type prototype strain (AB039777); (B) a P7-type strain (LC122916) in cluster 1; (C) a P7-type strain (MK956199) in cluster 2; (D) a P7-type strain (LC760173) in cluster 3; (E) a P6-type strain (JX989075) in cluster 1. The chains of the dimer structure are coloured dark grey (chain A) and light grey (chain B). Amino acid substitutions in each variant strain relative to the prototype strain are shown in blue, and the active sites are shown in orange.

### 3.8. 3D Mapping of the Positive Selection Sites and Conformational Epitopes in the VP1 Trimer Proteins

Furthermore, to better evaluate the locations of positive selection sites and conformational epitopes on the VP1 protein, we constructed and mapped 3D VP1 trimer proteins. First, as shown in Figure 5A–E and Table 3, and Section 3.5, positive selection sites for both P6- and P7-type VP1 proteins were located in the P2 domain. Of these, the positive selection sites in the P7-type VP1 proteins (Pro354Thr, Pro354Ser, Pro354Gln, Asn390Thr, and Asn390Asp) overlapped with some conformational epitopes (Tables 3 and S4), whereas a positive selection site (Lys386His) in P6-type VP1 proteins did not. These results suggest that the positive selection sites in P7-type GII.6 VP1 proteins escaped amino acid mutations. In addition, in both representative strains, amino acid mutations overlapped with the conformational epitope at many sites.

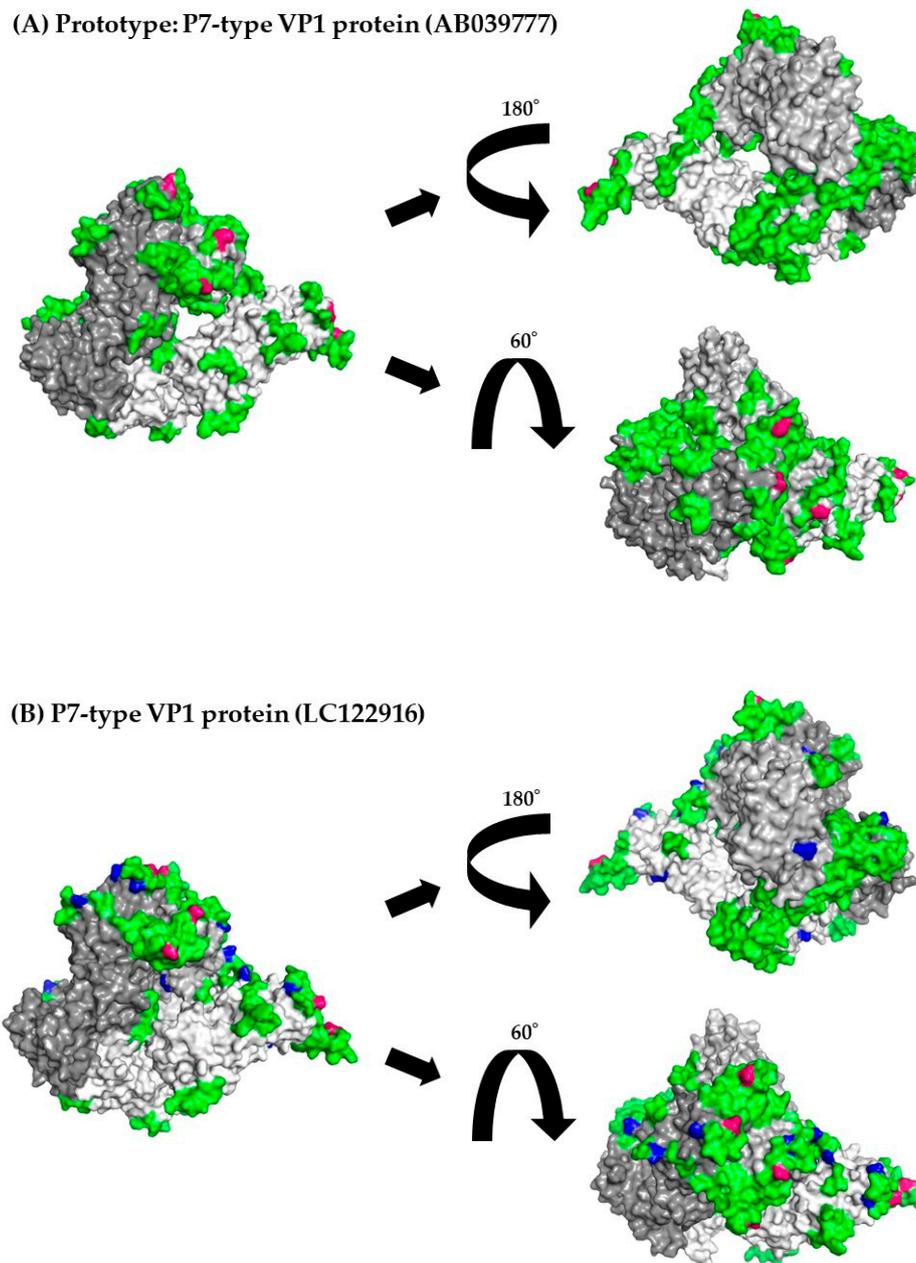


Figure 5. Cont.

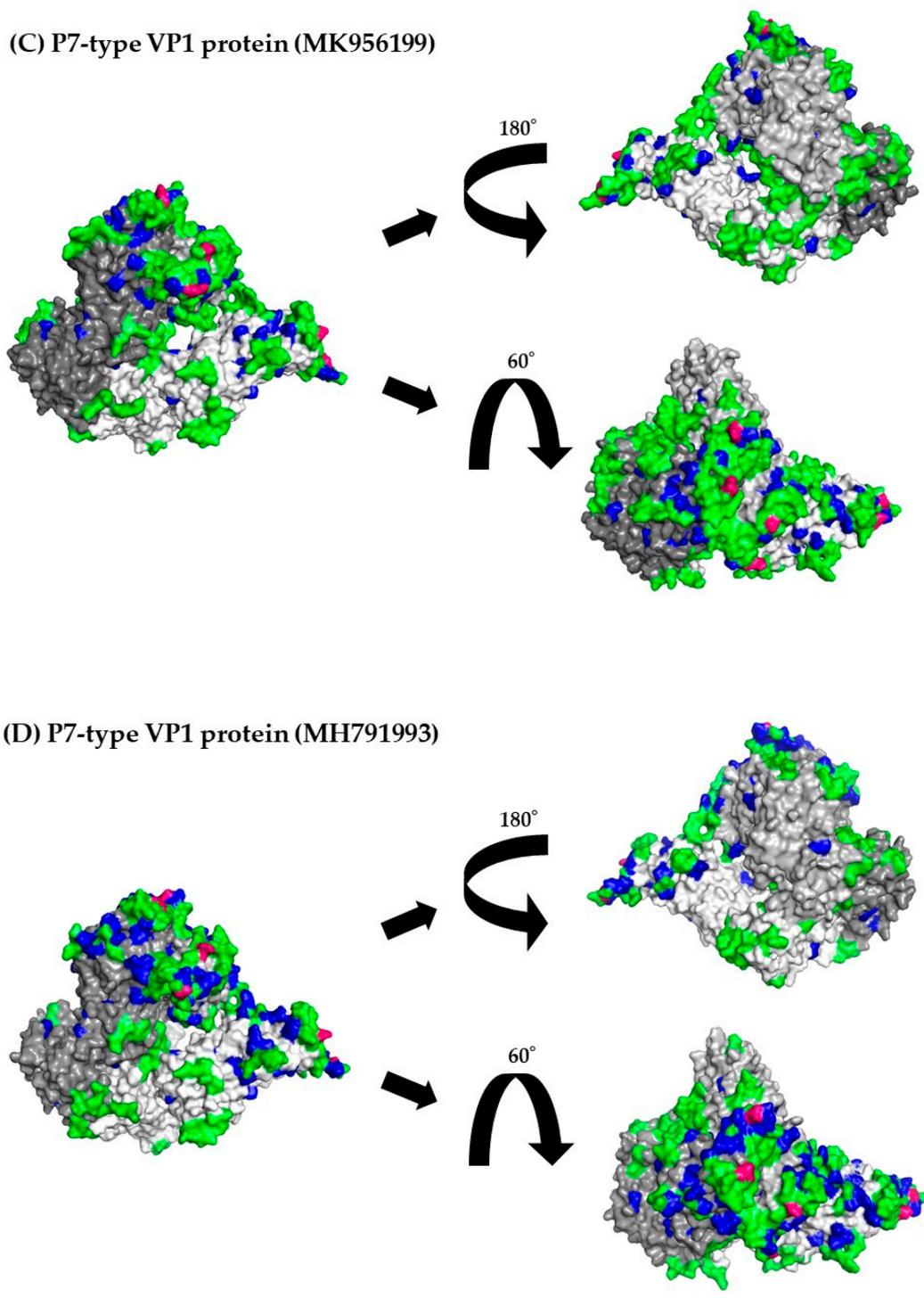
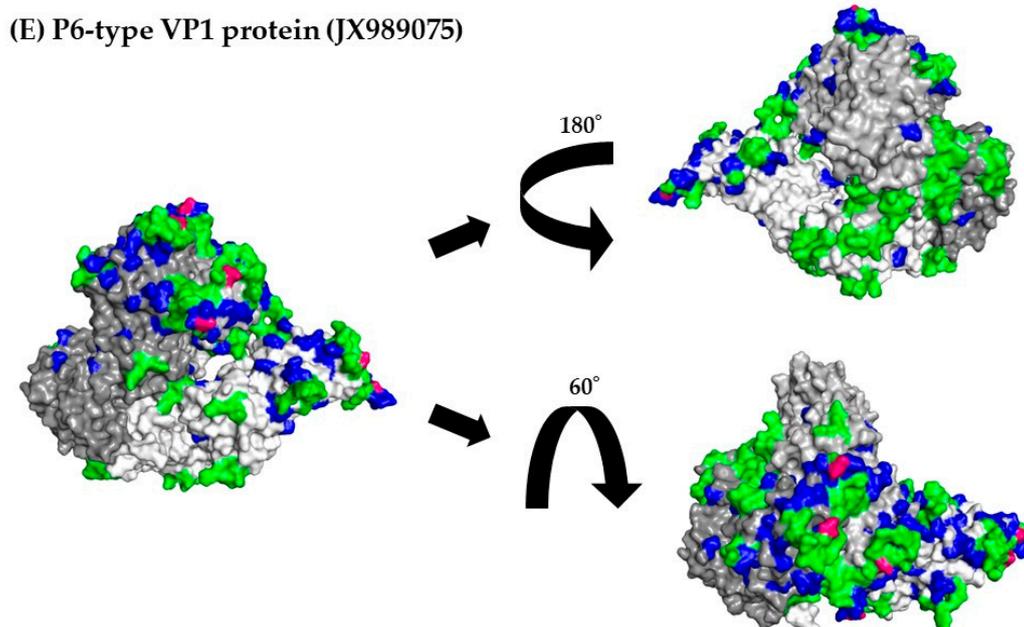


Figure 5. Cont.



**Figure 5.** 3D mapping of the positive selection sites and conformational epitopes in the VP1 protein (trimer). Illustration shows the 3D structure of the VP1 protein in the prototype and the most recent strain for each cluster. The strains in each figure are as follows: (A) a P7-type prototype strain (AB039777); (B) a P7-type strain (LC122916) in cluster 1; (C) a P7-type strain (MK956199) in cluster 2; (D) a P7-type strain (MH791993) in cluster 3; (E) a P6-type strain (JX989075) in cluster 1. Chains of the trimeric structures are coloured in dark grey (chain A), light grey (chain B), and white (chain C). Conformational epitopes of each strain are indicated in green. Amino acid substitutions of each strain are indicated in blue. Positive selection sites are coloured red. When amino acid substitutions overlapped with conformational epitopes, the amino acid substitutions were given priority and coloured blue. The amino acid sequences and details are provided in Supplementary Table S4.

**Table 3.** Number of amino acid residues of predicted positive and negative selection sites in HuNoVGII.6.

Region/Gene	Number of Negative Selection Sites	Number of Positive Selection Sites	Estimated as Positive Selective Sites
P6 type and P7 type <i>RdRp</i> region	258	0	—
P6 type <i>RdRp</i> region	3	0	—
P7 type <i>RdRp</i> region	205	1	126Lys, Lys126Arg
GII.P6-GII.6 and GII.P7-GII.6 <i>VP1</i> gene	298	2	354Pro, 390Asn
GII.P6-GII.6 <i>VP1</i> gene	8	1	Lys386His
GII.P7-GII.6 <i>VP1</i> gene	274	2	Pro354Thr, Ser and Gln, Asn390Thr and Asp

#### 4. Discussion

To better understand the evolution of HuNoV GII.6 strains with different *RdRp* types (P6 and P7), we analyzed both the *RdRp* region and *VP1* gene using various authentic bioinformatics technologies. First, a time-scaled phylogenetic tree showed that a common ancestor of the P6- and P7-type *RdRp* region emerged approximately 50 years ago and uniquely evolved and formed clusters. A common ancestor of P6- and P7-type GII.6 *VP1* gene emerged approximately 110 years ago and formed clusters. The dominant type for both the *RdRp* region and *VP1* gene was P7-type (Figure 1A,B). Secondly, the evolutionary

rates of both the P6-type *RdRp* region and *VP1* gene were faster than those of the P7-type *RdRp* region and *VP1* gene (Table 1). Next, the phylogenetic distances of the P7-type *RdRp* region and *VP1* gene were wider than those of the P6-type *RdRp* region and *VP1* gene. Furthermore, phylodynamic data showed that the *RdRp* region and *VP1* gene population sizes fluctuated after 2000 (Figure 3A–F). Some positive selection sites in the *VP1* proteins were estimated, and these were located in the antigenicity-related P2 domain. Among these, the positive selection sites in the P7-type *VP1* protein overlapped with the conformational epitopes (Figure 5A–D and Table S4). These data imply that the GII.6 *VP1* gene and *VP1* protein uniquely evolved because of recombination between the P6- and P7-type *RdRp* regions in the HuNoV GII.P6-GII.6 and GII.P7-GII.6 genomes.

A previous report regarding the evolutionary analyses of the *RdRp* region of various HuNoV genotypes showed that the P6- and P7-type *RdRp* region diverged from a common ancestor of other *RdRp* genotypes, including P18, P15, and P20 [42]. This report also estimated that the divergence year of the P6- and P7-types of the *RdRp* region was in the 1960s [42]. This may be compatible with the present data (December 1966). Moreover, the topologies of the previous time-scaled evolutionary tree and our tree were similar [42]. Although this and other reports did not show the evolutionary rates of each *RdRp* genotype, the evolutionary rates of various *RdRp* genotypes were estimated as  $2.52 \times 10^{-3}$  s/s/y to  $3.12 \times 10^{-3}$  s/s/y. The present data are also compatible with the data from a previous report [42]. These results suggested that the P6- and P7-type *RdRp* regions are genetically related.

Next, the HuNoV *RdRp* region/*RdRp* protein may have affected the evolution of the *VP1* gene/*VP1* protein [22,23]. As shown in Figure 1, the phylogeny of the *VP1* gene in GII.P6-GII.6 and GII.P7-GII.6 was clearly divided and evolved uniquely. In contrast, in the present study, the topology of the evolutionary tree of the *RdRp* region was similar to that of the *VP1* gene. Notably, the phylogeny of ORF1 and ORF2 was topologically similar in other genogroups and genotypes [43]. Previous reports have also suggested that recombination between the HuNoV genome ORF1, incorporating the *RdRp* region and ORF2, incorporating *VP1* gene, affects *VP1* gene/*VP1* protein evolution and HuNoV antigenicity [8,22,23]. For example, during the 2016/2017 season, recombination between different lineages of the P16-type *RdRp* region in the GII.P16-GII.2 strains occurred, and the recombinant caused large outbreaks of acute gastroenteritis in various countries [44–46]. Moreover, the GII.4 genotype caused a gastroenteritis pandemic between 2006 and 2012 [43,47]. Outbreaks may also be associated with recombination between ORF1 and ORF2 in GII.4 strains [48]. Based on previous and the present results, the prevalence of GII.P7-GII.2 strains was due to the recombination of P6- and P7-type *RdRp* regions. Moreover, the acquisition of new polymerases in recombinant strains may alter the evolution rate of the *VP1* gene [49]. Among Kawasaki 2014 type-detected cases, a norovirus GII.17 variant that was predominant in Hong Kong from 2014 to 2015 was significantly more common than GII.4 in elderly cases. GII.17 *VP1* protein evolution is estimated to be faster (by an order of magnitude) than *VP1* gene evolution in GII.4 [50]. This suggests that recombination may alter the susceptibility and evolutionary rate of the *VP1* protein in GII.17. Although we did not analyse *VP1* functional changes pre- and post-recombination between GII.P6 and GII.P7 in this study, a similar effect may have occurred in the GII.P7-GII.6 and GII.P6-GII.6 strains. Furthermore, the evolutionary rates of the *VP1* gene combined with P6- and P7-type *RdRp* regions were estimated as  $5.063 \times 10^{-3}$  s/s/y and  $3.022 \times 10^{-3}$  s/s/y, respectively. Previous data estimated the mean rates of various GII.2 genotype strains (GII.1 to GII.22) as  $3.21 \times 10^{-3}$  to  $4.30 \times 10^{-3}$  s/s/y [51]. Thus, these values and the present data may be similar [52]. Taken together, these findings provide information on the evolutionary history of these viral strains and suggest that recombination events may have played a pivotal role in their evolution.

We estimated the genetic divergence of the P6- and P7-type *RdRp* regions and P6- and P7-type *VP1* genes in the present strains. First, a larger divergence of P7-type *RdRp* regions and P7-type *VP1* genes was estimated compared to that of the P6-type *VP1* gene. In the

present study, the number of P6-type strains was relatively small (15 strains), although statistical analyses were performed.

We also analyzed the phylodynamics of the *RdRp* region and *VP1* gene. The results showed that the genome population size of GII.P6-GII.6 increased around 2000–2003, while the genome population size of GII.P7-GII.6 increased after 2010 and peaked around 2014. Previous epidemiological studies conducted in Shanghai, China and East Java, Indonesia showed that the detection of the P7-type in HuNoV cases peaked in 2014–2015, although the sample size was small [53,54]. Moreover, another molecular epidemiological study suggested that GII.6 had a biphasic prevalence between 2000 and 2005 and 2007 and 2010. Thus, the present phylodynamic data may reflect the prevalence of GII.P6-GII.6 and GII.P7-GII.6, although we could not determine the factors underlying these changes.

To evaluate the functional and evolutionary characteristics of the P6- and P7-type *RdRp* proteins, we constructed 3D dimeric *RdRp* proteins and mapped them with amino acid substitutions (Figure 4). Several amino acid substitutions were also identified. Previous reports have suggested that some amino acid substitutions are associated with replication efficacy [8,48]. For example, the efficacy of HuNoV genome replication is increased by amino acid substitutions (291Thr or 291Val) in various *RdRp* proteins [52]. However, no substitutions in the active sites were found in the P6- and P7-type *RdRp* proteins. Furthermore, both GII.P6-GII.6 and GII.P7-GII.6 *RdRp* proteins had a relatively small number of amino acid substitutions and no identifiable positive selection site. These results are compatible with previous reports investigating other genogroups (GI) and genotypes [42,55]. This is partly because *RdRp*, which is not a target of neutralising antibodies, undergoes less selective pressure from host immune systems than *VP1*.

We also constructed P6- and P7-type 3D trimeric *VP1* proteins (Figure 5). Previous reports have shown that the P2 domain may act not only as a host cell-binding site, but also as a major part of the HuNoV antigen [56,57]. Therefore, amino acid substitutions in this domain may be associated with infectivity and antigenicity [56,57]. Moreover, positively selected sites may function as escape mutations in the host [58]. In the present study, some conformational epitopes were identified in both *RdRp*-type *VP1* proteins. Some of these were located in the P2 domain. Positively selected sites were also identified. Moreover, the amino acid positions of the conformational epitopes and positive selections between the P6- and P7-type *VP1* proteins were distinct. These results suggested that the P6- and P7-type *VP1* proteins have distinct antigenicity, and both may undergo distinct selective pressure from host defence systems (i.e., host immunity). To our knowledge, the present study is the first to analyse the differences in antigenicity between GII.P7-GII.6 and GII.P6-GII.6, although we did not examine this *in vitro*.

Next, we evaluated negative selection sites for the *RdRp* and *VP1* proteins (Tables 3 and S3). Many negative selection sites in P7-type *RdRp* (205 sites) and *VP1* proteins (274 sites) were estimated, while P6-type *RdRp* and *VP1* proteins were small. In general, negative selection sites play a role in preventing the deterioration of viral protein function [46]. Thus, the present negative selection data may indicate the maintenance of *RdRp* and *VP1* protein function. Furthermore, a small number of negative selection sites in P6-type *RdRp* and *VP1* proteins were estimated. This may be because of the relatively small number of strains used in this study (15 strains).

The present study has some limitations. First, it lacks *in vitro* and *in vivo* approaches. Bioinformatics techniques are crucial in molecular evolution analyses, but they do not always reflect actual evolutionary trajectories. Our antigenicity analyses should be validated using *in vitro* and *in vivo* approaches in the future. Second, the study included a relatively small number of GII.P6-GII.6 strains. Despite collecting additional HuNoV GII.6 data from each local government public health institution, the number of P6- and P7-type *RdRp* strains that we collected and analysed was 15 and 126, respectively. Therefore, our results regarding the P6-type may be biased. Considering this limitation, further evolutionary analyses should be conducted after additional P6-type strains are detected.

## 5. Conclusions

In this study, to better understand the evolution of the HuNoV GII.P6-GII.6 and GII.P7-GII.6 strains, we performed a detailed analysis of both the *RdRp* region and *VP1* gene in these viruses using various bioinformatics methods. A common ancestor of the P6- and P7-type *RdRp* region emerged approximately 50 years ago and formed clusters. A common ancestor of the P6- and P7-type *VP1* gene emerged approximately 110 years ago. Moreover, both *RdRp* region and *VP1* gene have evolved uniquely. The evolutionary rates of the P6-type *RdRp* region and P6-type *VP1* gene were faster than the evolutionary rates of the P7-type *RdRp* region and *VP1* genes. More genetic divergence was observed in the P7-type *RdRp* region and *VP1* gene than in the P6-type *RdRp* region and *VP1* gene. The phylodynamics of the *RdRp* region and *VP1* fluctuated after 2000. Some positive selection sites in the *VP1* proteins were located in the antigenicity-related P2 domain, and these sites in the P7-type *VP1* protein overlapped with the conformational epitopes. Taken together, the GII.6 *VP1* and *VP1* proteins evolved uniquely due to recombination between the P6- and P7-type *RdRp* regions in the HuNoV GII.P6-GII.6 and GII.P7-GII.6 strains.

**Supplementary Materials:** The following supporting information can be downloaded from: <https://www.mdpi.com/article/10.3390/v15071497/s1>, Table S1. List of GII.P6-GII.6 and GII.P7-GII.6 strains used in this study. Table S2. Parameters used in Bayesian Markov chain Monte Carlo (MCMC) analyses. Table S3a. Details of P6- and P7-type *RdRp* region negative selection sites in comparing with each genotype. Table S3b. Details of GII.P6-GII.6 and GII.P7-GII.6 *VP1* gene negative selection sites in comparison with each genotype. Table S4. Detailed amino acid sequences of the *VP1* gene for representative and prototype strains of each cluster used in this study.

**Author Contributions:** Conceptualization, H.K., T.T. and K.K.; methodology, T.T., R.K., T.S. (Tatsuya Shirai), M.S., T.S. (Toshiyuki Sugai), K.H., F.M., K.O., Y.H. and M.K.; writing—original draft preparation, T.K., Y.S., K.K. and H.K.; review and editing, K.M., K.I., H.I., A.R., K.F., K.K. and H.K.; visualization, T.T., R.K. and Y.M.; funding acquisition, H.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Japan Agency for Medical Research and Development, AMED (<https://www.amed.go.jp/>, accessed on 1 April 2023), under Grant Number JP23fk0108667. The funders played no role in the study design, data collection, analysis, decision to publish, or manuscript preparation.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets generated and/or analyzed in the present study are available from the corresponding author upon reasonable request.

**Acknowledgments:** We thank Miki Kawaji for skillful supports in data processing.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Robilotti, E.; Deresinski, S.; Pinsky, B.A. Norovirus. *Clin. Microbiol. Rev.* **2015**, *28*, 134–164. [[CrossRef](#)] [[PubMed](#)]
2. Mans, J. Norovirus Infections and disease in lower-middle and low-income countries, 1997–2018. *Viruses* **2019**, *11*, 341. [[CrossRef](#)] [[PubMed](#)]
3. Chan-It, W.; Thongprachum, A.; Khamrin, P.; Kobayashi, M.; Okitsu, S.; Mizuguchi, M.; Ushijima, H. Emergence of a new norovirus GII.6 variant in Japan, 2008–2009. *J. Med. Virol.* **2012**, *84*, 1089–1096. [[CrossRef](#)] [[PubMed](#)]
4. Patel, M.M.; Hall, A.J.; Vinjé, J.; Parashar, U.D. Noroviruses: A comprehensive review. *J. Clin. Virol.* **2009**, *44*, 1–8. [[CrossRef](#)]
5. Patel, M.M.; Widdowson, M.A.; Glass, R.I.; Akazawa, K.; Vinjé, J.; Parashar, U.D. Systematic literature review of role of noroviruses in sporadic gastroenteritis. *Emerg. Infect. Dis.* **2008**, *14*, 1224–1231. [[CrossRef](#)]
6. Parrón, I.; Álvarez, J.; Jané, M.; Cornejo Sánchez, T.; Razquin, E.; Guix, S.; Camps, G.; Pérez, C.; Domínguez, À.; The Working Group for the Study of Outbreaks of Acute Gastroenteritis in Catalonia. A foodborne norovirus outbreak in a nursing home and spread to staff and their household contacts. *Epidemiol. Infect.* **2019**, *147*, e225. [[CrossRef](#)]
7. Winder, N.; Gohar, S.; Muthana, M. Norovirus: An overview of virology and preventative measures. *Viruses* **2022**, *14*, 2811. [[CrossRef](#)]

8. de Graaf, M.; van Beek, J.; Koopmans, M.P. Human norovirus transmission and evolution in a changing world. *Nat. Rev. Microbiol.* **2016**, *14*, 421–433. [[CrossRef](#)]
9. Jiang, X.; Wang, M.; Wang, K.; Estes, M.K. Sequence and genomic organization of Norwalk virus. *Virology* **1993**, *195*, 51–61. [[CrossRef](#)]
10. Chhabra, P.; de Graaf, M.; Parra, G.I.; Chan, M.C.; Green, K.; Martella, V.; Wang, Q.; White, P.A.; Katayama, K.; Vennema, H.; et al. Updated classification of norovirus genogroups and genotypes. *J. Gen. Virol.* **2019**, *100*, 1393–1406. [[CrossRef](#)]
11. Jiang, X.; Espul, C.; Zhong, W.M.; Cuello, H.; Matson, D.O. Characterization of a novel human calicivirus that may be a naturally occurring recombinant. *Arch. Virol.* **1999**, *144*, 2377–2387. [[CrossRef](#)]
12. Ludwig-Begall, L.F.; Mauroy, A.; Thiry, E. Norovirus recombinants: Recurrent in the field, recalcitrant in the lab—A scoping review of recombination and recombinant types of noroviruses. *J. Gen. Virol.* **2018**, *99*, 970–988. [[CrossRef](#)] [[PubMed](#)]
13. Kroneman, A.; Vega, E.; Vennema, H.; Vinjé, J.; White, P.A.; Hansman, G.; Green, K.; Martella, V.; Katayama, K.; Koopmans, M. Proposal for a unified norovirus nomenclature and genotyping. *Arch. Virol.* **2013**, *158*, 2059–2068. [[CrossRef](#)] [[PubMed](#)]
14. Bhavanam; Freedman, S.; Lee, S.B.; Zhuo, B.E.; Qiu, R.; Chui, Y.; Xie, L.; Ali, J.; Vanderkooi, S.; Pang, O.G.; et al. Differences in Illness Severity among Circulating Norovirus Genotypes in a Large Pediatric Cohort with Acute Gastroenteritis. *Microorganisms* **2020**, *8*, 1873. [[CrossRef](#)] [[PubMed](#)]
15. van Beek, J.; de Graaf, M.; Al-Hello, H.; Allen, D.J.; Ambert-Balay, K.; Botteldoorn, N.; Brytting, M.; Buesa, J.; Cabrerizo, M.; Chan, M.; et al. Molecular surveillance of norovirus, 2005–16: An epidemiological analysis of data collected from the NoroNet network. *Lancet Infect. Dis.* **2018**, *18*, 545–553. [[CrossRef](#)] [[PubMed](#)]
16. Khamrin, P.; Kumthip, K.; Yodmeeklin, A.; Jampanil, N.; Phengma, P.; Yamsakul, P.; Okitsu, S.; Kobayashi, T.; Ushijima, H.; Maneekarn, N. Changing Predominance of Norovirus Recombinant Strains GII.2[P16] to GII.4[P16] and GII.4[P31] in Thailand, 2017 to 2018. *Microbiol. Spectr.* **2022**, *10*, e0044822. [[CrossRef](#)]
17. Vega, E.; Barclay, L.; Gregoricus, N.; Shirley, S.H.; Lee, D.; Vinj, J. Genotypic and epidemiologic trends of norovirus outbreaks in the United States, 2009 to 2013. *J. Clin. Microbiol.* **2014**, *52*, 147–155. [[CrossRef](#)]
18. Cannon, J.L.; Bonifacio, J.; Bucardo, F.; Buesa, J.; Bruggink, L.; Chan, M.C.-W.; Fumian, T.M.; Giri, S.; Gonzalez, M.D.; Hewitt, J.; et al. Global Trends in Norovirus Genotype Distribution among Children with Acute Gastroenteritis. *Emerg. Infect. Dis.* **2021**, *27*, 1438–1445. [[CrossRef](#)]
19. Li, J.; Zhang, L.; Zou, W.; Yang, Z.; Zhan, J.; Cheng, J. Epidemiology and genetic diversity of norovirus GII genogroups among children in Hubei, China, 2017–2019. *Virol. Sin.* **2023**, *38*, 351–362. [[CrossRef](#)]
20. Pybus, O.G.; Rambaut, A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat. Rev. Genet.* **2009**, *10*, 540–550. [[CrossRef](#)]
21. Matsushima, Y.; Mizukoshi, F.; Sakon, N.; Doan, Y.H.; Ueki, Y.; Ogawa, Y.; Motoya, T.; Tsukagoshi, H.; Nakamura, N.; Shigemoto, N.; et al. Evolutionary Analysis of the VP1 and RNA-dependent RNA polymerase regions of human norovirus GII.P17-GII.17 in 2013–2017. *Front. Microbiol.* **2019**, *10*, 2189. [[CrossRef](#)]
22. Mizukoshi, F.; Nagasawa, K.; Doan, Y.H.; Haga, K.; Yoshizumi, S.; Ueki, Y.; Shinohara, M.; Ishikawa, M.; Sakon, N.; Shigemoto, N.; et al. Molecular Evolution of the RNA-Dependent RNA Polymerase and Capsid Genes of Human Norovirus Genotype GII.2 in Japan during 2004–2015. *Front. Microbiol.* **2017**, *8*, 705. [[CrossRef](#)]
23. Nagasawa, K.; Matsushima, Y.; Motoya, T.; Mizukoshi, F.; Ueki, Y.; Sakon, N.; Murakami, K.; Shimizu, T.; Okabe, N.; Nagata, N.; et al. Phylogeny and Immunoreactivity of Norovirus GII.P16-GII.2, Japan, Winter 2016–17. *Emerg. Infect. Dis.* **2018**, *24*, 144–148. [[CrossRef](#)]
24. Diakoudi, G.; Lanave, G.; Catella, C.; Medici, M.C.; De Conto, F.; Calderaro, A.; Loconsole, D.; Chironna, M.; Bonura, F.; Giammanco, G.M.; et al. Analysis of GII.P7 and GII.6 noroviruses circulating in Italy during 2011–2016 reveals a replacement of lineages and complex recombination history. *Infect. Genet. Evol.* **2019**, *75*, 103991. [[CrossRef](#)] [[PubMed](#)]
25. Yang, Z.; Vinj, J.; Elkins, C.A.; Kulka, M. Complete Genome Sequence of Human Norovirus Strain GII.P7-GII.6 Detected in a Patient in the United States in 2014. *Genome Announc.* **2016**, *4*, e01211-16. [[CrossRef](#)] [[PubMed](#)]
26. Zhou, N.; Li, M.; Zhou, L.U.; Huang, Y. Genetic characterizations and molecular evolution of human norovirus GII.6 genotype during the past five decades. *J. Med. Virol.* **2023**, *95*, e28876. [[CrossRef](#)]
27. Bouckaert, R.; Heled, J.; Kühnert, D.; Vaughan, T.; Wu, C.-H.; Xie, D.; Suchard, M.A.; Rambaut, A.; Drummond, A.J. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comp. Biol.* **2014**, *10*, e1003537. [[CrossRef](#)]
28. Saito, M.; Tsukagoshi, H.; Sada, M.; Sunagawa, S.; Shirai, T.; Okayama, K.; Sugai, T.; Tsugawa, T.; Hayashi, Y.; Ryo, A.; et al. Detailed Evolutionary Analyses of the F gene in the respiratory syncytial virus subgroup A. *Viruses* **2021**, *13*, 2525. [[CrossRef](#)]
29. Darriba, D.; Taboada, G.L.; Doallo, R.; Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods* **2012**, *9*, 772. [[CrossRef](#)] [[PubMed](#)]
30. Kanda, Y. Investigation of the freely available easy-to-use software ‘EZR’ for medical statistics. *Bone Marrow Transpl.* **2013**, *48*, 452–458. [[CrossRef](#)] [[PubMed](#)]
31. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [[CrossRef](#)] [[PubMed](#)]
32. Fourment, M.; Gibbs, M.J. PATRISTIC: A program for calculating patristic distances and graphically comparing the components of genetic change. *BMC Evol. Biol.* **2006**, *6*, 1. [[CrossRef](#)]

33. Rambaut, A.; Drummond, A.J.; Xie, D.; Baele, G.; Suchard, M.A. Posterior Summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **2018**, *67*, 901–904. [[CrossRef](#)]
34. Weaver, S.; Shank, S.D.; Spielman, S.J.; Li, M.; Muse, S.V.; Kosakovsky Pond, S.L. Datamonkey 2.0: A modern web application for characterizing selective and other evolutionary processes. *Mol. Biol. Evol.* **2018**, *35*, 773–777. [[CrossRef](#)] [[PubMed](#)]
35. Webb, B.; Sali, A. Protein structure modeling with MODELLER. *Methods Mol. Biol.* **2014**, *1137*, 1–15. [[CrossRef](#)]
36. Emsley, P.; Lohkamp, B.; Scott, W.G.; Cowtan, K. Features and development of coot. *Acta. Crystallogr. D. Biol. Crystallogr.* **2010**, *66*, 486–501. [[CrossRef](#)] [[PubMed](#)]
37. Guex, N.; Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **1997**, *18*, 2714–2723. [[CrossRef](#)]
38. Kringelum, J.V.; Lundegaard, C.; Lund, O.; Nielsen, M. Reliable B cell epitope predictions: Impacts of method development and improved benchmarking. *PLOS Comput. Biol.* **2012**, *8*, e1002829. [[CrossRef](#)]
39. Ponomarenko, J.; Bui, H.H.; Li, W.; Füsseder, N.; Bourne, P.E.; Sette, A.; Peters, B. ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinform.* **2008**, *9*, 514. [[CrossRef](#)]
40. Shashkova, T.I.; Umerenkov, D.; Salnikov, M.; Strashnov, P.V.; Konstantinova, A.V.; Lebed, I.; Shcherbinin, D.N.; Asatryan, M.N.; Kardymon, O.L.; Ivanisenko, N.V. SEMA: Antigen B-cell conformational epitope prediction using deep transfer learning. *Front. Immunol.* **2022**, *13*, 960985. [[CrossRef](#)]
41. Zhou, C.; Chen, Z.; Zhang, L.; Yan, D.; Mao, T.; Tang, K.; Qiu, T.; Cao, Z. SEPPA 3.0-enhanced spatial epitope prediction enabling glycoprotein antigens. *Nucleic Acids Res.* **2019**, *47*, W388–W394. [[CrossRef](#)] [[PubMed](#)]
42. Ozaki, K.; Matsushima, Y.; Nagasawa, K.; Motoya, T.; Ryo, A.; Kuroda, M.; Katayama, K.; Kimura, H. Molecular evolutionary analyses of the RNA-dependent RNA polymerase region in norovirus genogroup II. *Front. Microbiol.* **2018**, *9*, 3070. [[CrossRef](#)] [[PubMed](#)]
43. Vinjé, J. Advances in laboratory methods for detection and typing of norovirus. *J. Clin. Microbiol.* **2015**, *53*, 373–381. [[CrossRef](#)]
44. Bidalot, M.; Théry, L.; Kaplon, J.; De Rougemont, A.; Ambert-Balay, K. Emergence of new recombinant noroviruses GII.p16-GII.4 and GII.p16-GII.2, France, winter 2016 to 2017. *Eurosurveillance* **2017**, *22*, 30508. [[CrossRef](#)]
45. Ao, Y.; Wang, J.; Ling, H.; He, Y.; Dong, X.; Wang, X.; Peng, J.; Zhang, H.; Jin, M.; Duan, Z. Norovirus GII.P16/GII.2-Associated Gastroenteritis, China, 2016. *Emerg. Infect. Dis.* **2017**, *23*, 1172–1175. [[CrossRef](#)]
46. Liu, L.T.; Kuo, T.Y.; Wu, C.Y.; Liao, W.T.; Hall, A.J.; Wu, F.T. Recombinant GII.P16-GII.2 Norovirus, Taiwan, 2016. *Emerg. Infect. Dis.* **2017**, *23*, 1180–1183. [[CrossRef](#)]
47. Siebenga, J.J.; Vennema, H.; Zheng, D.P.; Vinjé, J.; Lee, B.E.; Pang, X.L.; Ho, E.C.; Lim, W.; Choudekar, A.; Broor, S.; et al. Norovirus illness is a global problem: Emergence and spread of norovirus GII.4 variants, 2001–2007. *J. Infect. Dis.* **2009**, *200*, 802–812. [[CrossRef](#)] [[PubMed](#)]
48. Eden, J.S.; Tanaka, M.M.; Boni, M.F.; Rawlinson, W.D.; White, P.A. Recombination within the pandemic norovirus GII.4 lineage. *J. Virol.* **2013**, *87*, 6270–6282. [[CrossRef](#)]
49. Smertina, E.; Urakova, N.; Strive, T.; Frese, M. Calicivirus RNA-Dependent RNA Polymerases: Evolution, Structure, Protein Dynamics, and Function. *Front. Microbiol.* **2019**, *10*, 1280. [[CrossRef](#)]
50. Chan, M.C.W.; Lee, N.; Hung, T.-N.; Kwok, K.; Cheung, K.; Tin, E.K.Y.; Lai, R.W.M.; Nelson, E.A.S.; Leung, T.F. Rapid emergence and predominance of a broadly recognizing and fast-evolving norovirus GII.17 variant in late 2014. *Nat. Commun.* **2015**, *6*, 10061. [[CrossRef](#)]
51. Kobayashi, M.; Matsushima, Y.; Motoya, T.; Sakon, N.; Shigemoto, N.; Okamoto-Nakagawa, R.; Nishimura, K.; Yamashita, Y.; Kuroda, M.; Saruki, N.; et al. Molecular evolution of the capsid gene in human norovirus genogroup II. *Sci. Rep.* **2016**, *6*, 29400. [[CrossRef](#)]
52. Bull, R.A.; Eden, J.S.; Rawlinson, W.D.; White, P.A. Rapid evolution of pandemic noroviruses of the GII.4 lineage. *PLoS Pathog.* **2010**, *6*, e1000831. [[CrossRef](#)]
53. Lu, L.; Zhong, H.; Xu, M.; Su, L.; Cao, L.; Jia, R.; Xu, J. Genetic diversity and epidemiology of Genogroup II noroviruses in children with acute sporadic gastroenteritis in Shanghai, China, 2012–2017. *Bmc Infect. Dis.* **2019**, *19*, 736. [[CrossRef](#)] [[PubMed](#)]
54. Utsumi; Lusida; Dinana, T.; Wahyuni, M.I.; Soegijanto, Z.; Soetjijto, R.M.; Athiyyah, S.; Sudarmo, A.F.; Ranuh, S.M.; Darma, R.G.; et al. Molecular epidemiology and genetic diversity of norovirus infection in children hospitalized with acute gastroenteritis in East Java, Indonesia in 2015–2019. *Infect. Genet. Evol.* **2021**, *88*, 104703. [[CrossRef](#)]
55. Zhou, N.; Huang, Y.; Zhou, L.U.; Li, M.; Jin, H. Molecular Evolution of RNA-Dependent RNA Polymerase Region in Norovirus Genogroup I. *Viruses* **2023**, *15*, 166. [[CrossRef](#)] [[PubMed](#)]
56. Hutson, A.M.; Atmar, R.L.; Marcus, D.M.; Estes, M.K. Norwalk virus-like particle hemagglutination by binding to h histo-blood group antigens. *J. Virol.* **2003**, *77*, 405–415. [[CrossRef](#)] [[PubMed](#)]

57. Schroten, H.; Hanisch, F.G.; Hansman, G.S. Human norovirus interactions with histo-blood group antigens and human milk oligosaccharides. *J. Virol.* **2016**, *90*, 5855–5859. [[CrossRef](#)]
58. Holmes, E.C. Virus evolution. In *Fields Virology*, 6th ed.; Knipe, D.M., Howley, P.M., Cohen, J.I., Griffin, D.E., Lamb, R.A., Martin, M.A., Racaniello, V.D., Roizman, B., Eds.; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2013; Volume 1, pp. 286–313.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.