



Emergence and Potential Extinction of Genetic Lineages of Human Metapneumovirus between 2005 and 2021

Kevin Groen,^a Stefan van Nieuwkoop,^a Adam Meijer,^b Bas van der Veer,^b Jeroen J. A. van Kampen,^a Pieter L. Fraaij,^{a,c} Ron A. M. Fouchier,^a  Bernadette G. van den Hoogen^a

^aDepartment of Viroscience, Erasmus MC, Rotterdam, the Netherlands

^bCenter for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, the Netherlands

^cPediatric Infectious Diseases & Immunology, Erasmus MC-Sophia Children's Hospital, Rotterdam, the Netherlands

ABSTRACT Human metapneumovirus (HMPV) is one of the leading causes of respiratory illness (RI), primarily in infants. Worldwide, two genetic lineages (A and B) of HMPV are circulating that are antigenically distinct and can each be further divided into genetic sublineages. Surveillance combined with large-scale whole-genome sequencing studies of HMPV are scarce but would help to identify viral evolutionary dynamics. Here, we analyzed 130 whole HMPV genome sequences obtained from samples collected from individuals hospitalized with RI and partial fusion ($n = 144$) and attachment ($n = 123$) protein gene sequences obtained from samples collected from patients with RI visiting general practitioners between 2005 and 2021 in the Netherlands. Phylogenetic analyses demonstrated that HMPV continued to group in the four sublineages described in 2004 (A1, A2, B1, and B2). However, one sublineage (A1) was no longer detected in the Netherlands after 2006, while the others continued to evolve. No differences were observed in dominant (sub)lineages between samples obtained from patients with RI being hospitalized and those consulting general practitioners. In both populations, viruses of lineage A2 carrying a 180-nucleotide or 111-nucleotide duplication in the attachment protein gene became the most frequently detected genotypes. In the past, different names for the newly emerging lineages have been proposed, demonstrating the need for a consistent naming convention. Here, criteria are proposed for the designation of new genetic lineages to aid in moving toward a systematic HMPV classification.

IMPORTANCE Human metapneumovirus (HMPV) is one of the major causative agents of human respiratory tract infections. Monitoring of virus evolution could aid toward the development of new antiviral treatments or vaccine designs. Here, we studied HMPV evolution between 2005 and 2021, with viruses obtained from samples collected from hospitalized individuals and patients with respiratory infections consulting general practitioners. Phylogenetic analyses demonstrated that HMPV continued to group in the four previously described sublineages (A1, A2, B1, and B2). However, one sublineage (A1) was no longer detected after 2006, while the others continued to evolve. No differences were observed in dominant (sub)lineages between patients being hospitalized and those consulting general practitioners. In both populations, viruses of lineage A2 carrying a 180-nucleotide or 111-nucleotide duplication in the attachment protein gene became the most frequently detected genotypes. These data were used to propose criteria for the designation of new genetic lineages to aid toward a systematic HMPV classification.

KEYWORDS human metapneumovirus, molecular evolution, next-generation sequencing, Nanopore technology, classification, phylogeny, duplication in G

Human metapneumovirus (HMPV) is a nonsegmented negative-strand RNA virus belonging to the *Pneumoviridae* family. Globally, HMPV is responsible for 5 to 15% of respiratory infections and is one of the leading causative agents of respiratory illness in young children,

Editor Peter Palese, Icahn School of Medicine at Mount Sinai

Copyright © 2022 Groen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Bernadette G. van den Hoogen, b.vandenhoogen@erasmusmc.nl.

The authors declare no conflict of interest.

Received 10 August 2022

Accepted 8 November 2022

Published 12 December 2022

only second to respiratory syncytial virus (RSV) (1–8). Clinical symptoms caused by HMPV range from mild respiratory illness to bronchiolitis and pneumonia, similar to those caused by RSV (reviewed in reference 9).

HMPV has a genome of approximately 13.3 kb in size and contains 8 genes, which encode 9 proteins. This includes three surface proteins, the short hydrophobic protein (SH), the attachment glycoprotein (G), and the fusion protein (F). These surface proteins can be used for virus genotyping, also in relation to antigenic variation of circulating HMPV (10, 11). In 2004, it was reported that HMPV isolates clustered in two main genetic lineages that are also antigenically distinct (A and B) and that each diverged in two sublineages (A1 and A2 and B1 and B2) (10). In 2006, a further split of sublineage A2 into A2a and A2b was reported (12). Although the presence of a third lineage (A2c) was reported, another study suggested that the A2b lineage should be further divided in A2b1 and A2b2 and that A2b2 can be mistaken for lineage A2c (13, 14). There were also suggestions that the B2 lineage should be further divided into lineages B2a and B2b, although a consensus on this matter has not been achieved (15). Recently, circulation of lineage A2 viruses carrying a 180-nucleotide (nt) or 111-nt duplication in the G gene was reported for the first time in Japan in 2014 and 2017, respectively (16, 17). In the following years, these viruses were also detected in Spain, Vietnam, and China (16–18). One of these genotypes, with a 111-nt duplication, became the dominant circulating genotype in Yokohama city in Japan (16–18).

Studies on HMPV evolution have focused mainly on the genetic variation of the F and G genes (10, 11, 13, 16, 17, 19–21). However, whole-genome sequencing of HMPV might provide insights into virus evolution that will be missed by analyzing (partial) F and G sequences. Recent genetic changes, such as sequence duplications in the HMPV G gene, highlight the importance of monitoring HMPV evolutionary dynamics, but partial genome sequencing will miss any evolutionary events that occur in genes other than F and G. With next-generation sequencing (NGS) platforms rapidly improving, obtaining viral whole-genome sequences has simplified considerably (22). Recent studies reported assays for HMPV whole-genome sequencing using Illumina sequencing, Ion Torrent, or Nanopore technology (23–26). We recently introduced an HMPV whole-genome sequencing method using MinION Nanopore technology (27). In this study, using this method, whole-genome sequences were generated from 130 HMPV strains obtained from samples collected from hospitalized individuals suffering from respiratory tract infections (RTIs) between 2005 and 2021. Additionally, partial F and G gene sequences were analyzed from 144 viruses obtained from patients with an RTI that visited general practitioners (GPs) between 2005 and 2021 to compare circulation of viruses between the hospitalized patients and patients consulting general practitioners. Phylogenetic analyses were performed to study recent evolutionary patterns and used to propose a robust HMPV classification system for emerging virus lineages.

RESULTS

Phylogenetic analyses of HMPV whole-genome sequences. Full-length genome sequences were generated for 130 viruses obtained from hospitalized individuals between 2005 and 2021 and for 13 additional viruses from earlier years (1993 to 2004) that were also not sequenced previously. Phylogenetic analysis of these 143 whole-genome sequences, in combination with all 56 whole-genome sequences available from GenBank (total of $n = 199$), demonstrated clustering of the viruses in the two main genetic lineages A and B (Fig. 1; Fig. S1 and Table S1 in the supplemental material). Within these lineages, separation into the four previously described sublineages A1, A2, B1, and B2 was still observed. Additional subdivisions within lineage A2 were clearly possible, as suggested before (12, 13), but clear criteria for such novel lineage designations are needed. Moreover, different names for the newly emerging lineages have been used in the past, demonstrating the need for a consistent naming convention. For simplicity, lineages from here forward are referred to as A1, A2, A2.1, A2.2, A2.2.1, A2.2.2, B1, and B2. Hereby, A and B are used to distinguish between different phenotypes, and phylogenetic clusters are defined numerically.

In this sample set, viruses with a duplication in the G gene were detected for the first time in 2015 (Table 1). Viruses containing either a 111-nucleotide or a 180-nucleotide

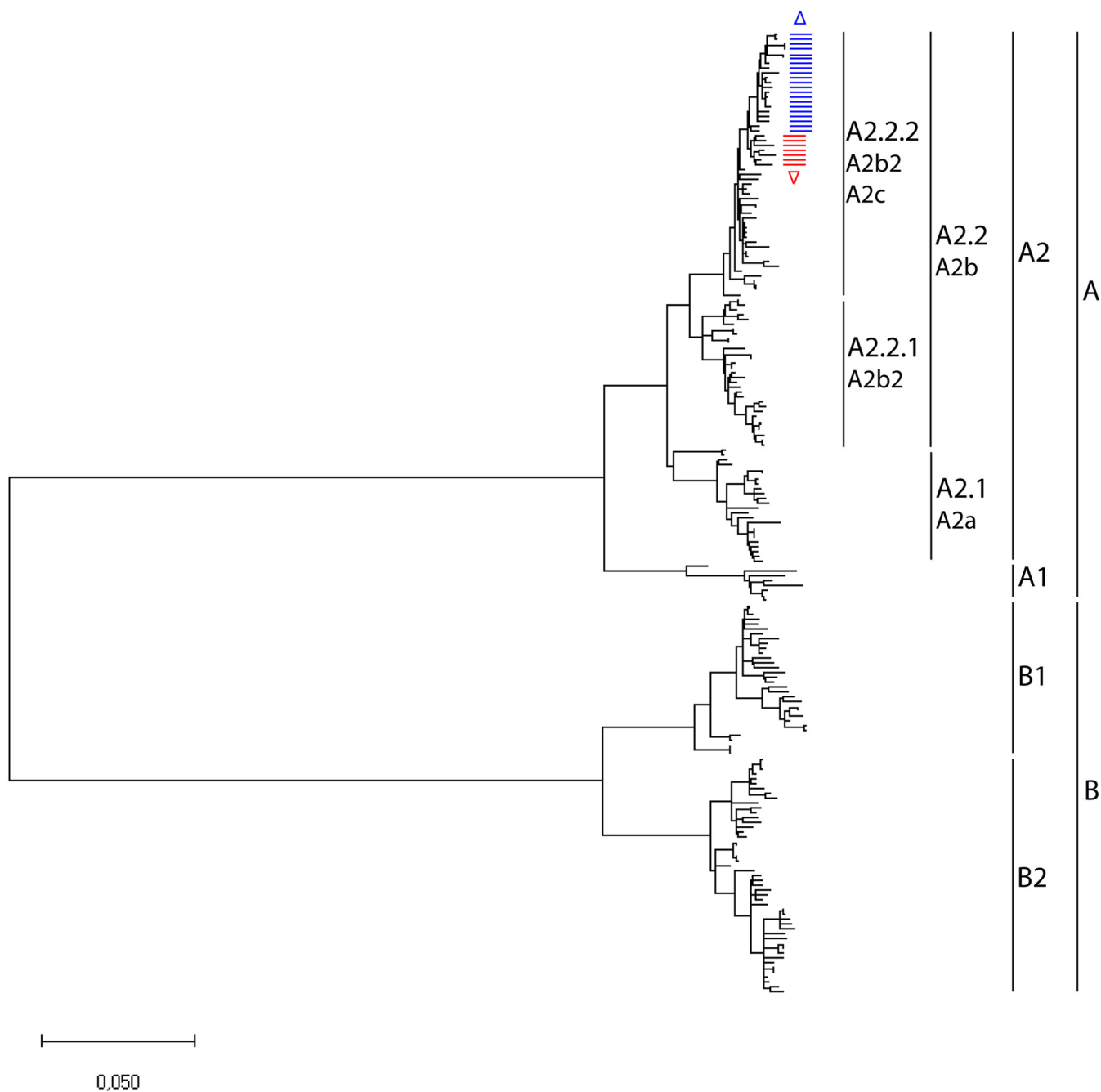


FIG 1 Maximum likelihood phylogenetic tree based on 199 whole HMPV genome sequences. The tree was reconstructed using the Maximum-Likelihood (GTR+G+I) substitution model with 1,000 bootstraps. The tree with bootstrap values and virus names is shown in Fig. S1 in the supplemental material. Markers on the right indicate the main virus lineages A and B, which could be divided in 6 sublineages (A1, A2.1, A2.2.1, A2.2.2, B1, and B2). Lineage A2.2.2 viruses with a 180-nt duplication (∇ , red line) or a 111-nt duplication (Δ , blue line) in the G gene are highlighted. Different names that have been used for each lineage are shown.

duplication in the G gene clustered within the A2.2.2 lineage in two smaller clusters (Fig. 1). The most recent virus with a 180-nucleotide duplication in the G gene was detected in 2018, while from 2019 onward, only viruses with a 111-nucleotide duplication in the G gene were detected. After 2018, no more lineage A2.2.2 viruses without a duplication in the G gene were detected.

Comparison of virus clustering based on whole-genome, full-length F, or partial F gene sequences. The majority of reported phylogenetic analyses have been based on either full-length or partial F gene sequences (in general nucleotide 778 to nucleotide 1226 of the F gene [10]). To compare clustering of viruses based on full-length genome sequences with

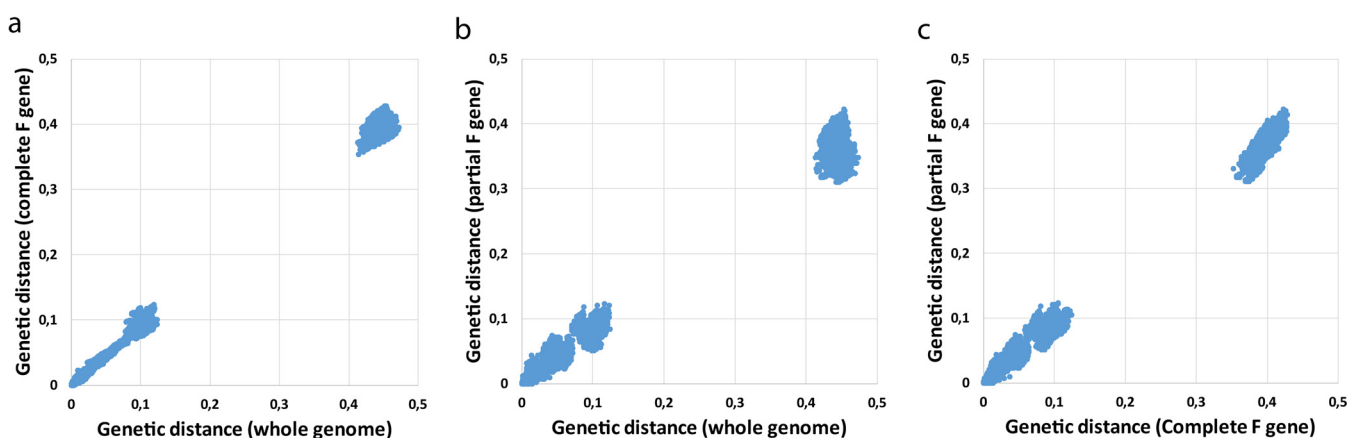
TABLE 1 Increase in incidence of viruses carrying 180- and 111-nucleotide duplications in the G gene from samples obtained from hospitalized patients within lineage A2.2.2 between 2014 and 2021

Yr	180-nt duplication ^a	111-nt duplication ^a	No duplication ^a	Total ^a
2014	–	–	5 (100%)	5
2015	–	2 (29%)	5 (71%)	7
2016	3 (38%)	2 (25%)	3 (38%)	8
2017	2 (29%)	3 (43%)	2 (29%)	7
2018	1 (50%)	1 (50%)	–	2
2019	–	2 (100%)	–	2
2020	–	–	–	–
2021	–	4 (100%)	–	4

^a–, no data.

that based on full-length F or partial F gene sequences, phylogenetic trees were reconstructed from nucleotide alignments of full-length F gene and partial F gene sequences (nucleotide 778 to nucleotide 1226 of the F gene) derived from the 143 full genome sequences described above. No difference in clustering of the viruses was observed between the three analyses, which demonstrates the consistency of lineage topology (Fig. S1, S2, and S3). Interestingly, viruses with a duplication in the G gene also clustered together based on full-length or partial F gene sequences, demonstrating that full-length or partial F gene sequences by themselves carry sufficient information to accurately genotype a virus. To assess whether the phylogenetic distances between viruses were similar in each of the three trees, scatterplots were generated from the *p* distances (the proportion of nucleotide sites at which sequences differ) between all viruses from one tree plotted against those from another tree. The variation in *p* distance between the trees reconstructed from whole-genome sequences and full-length F gene sequences was the smallest (Fig. 2a), while the variation in *p* distances between trees reconstructed from partial F gene sequences and either full-length F gene sequences or whole-genome sequences was slightly larger (Fig. 2b and c), indicating that full-length F gene sequences can be used for phylogenetic analysis with similar accuracy as whole-genome sequences.

Designation of HMPV lineages. Although the majority of HMPV evolution studies are based on analysis of F gene sequences, some studies rely on analysis of the more variable G gene, while the SH gene is rarely used for these studies. To compare genetic clustering based on the F, G, and SH genes, an alignment was generated from HMPV whole-genome sequences obtained in this study combined with those available from GenBank (*n* = 199). From this alignment, the F, G, and SH genes were selected and used to analyze genetic clustering. For all three proteins, genetic maps were constructed from the sequence alignments. Genetic maps were generated by multidimensional scaling of genetic distance matrixes, which contain all pairwise numbers of nucleotide substitutions between viruses. Distances in the

**FIG 2** (a to c) Scatterplots of *p* distances between viruses from phylogenetic trees reconstructed from whole-genome sequences versus full-length F gene sequences (a), whole-genome sequences versus partial F gene sequences (b), and full-length F gene sequences versus partial F gene sequences (c). The trees from which *p* distances were compared are shown in Fig. S1, S2, and S3.

map, represented by a grid, demonstrate the number of nucleotide substitutions between viruses, with each line of the grid representing 10 substitutions. The genetic maps demonstrated clear clustering into six distinct lineages: A1, A2.1, A2.2.1, A2.2.2, B1, and B2 (Fig. S4). Clustering based on F gene sequences resulted in the highest resolution between these six lineages, which supported the use of F gene sequences for designation of lineages.

To propose a robust HMPV classification system, genetic evolution was analyzed based on alignment of full-length F gene sequences obtained in this study combined with those from GenBank ($n = 744$). Analysis of a genetic map and a phylogenetic tree reconstructed from this alignment demonstrated that the viruses clustered in six distinct lineages (Fig. 3a to c; Fig. S5), identical to that described above. The A2.2 lineage displayed gradual evolution over time, while in other lineages, directional evolution was not as clear (Fig. 3b). In the genetic map, there appeared to be a segregation within the B2 lineage, although the border between these two lineages was not clear. This was in agreement with the phylogenetic tree, in which the distance between these two lineages was relatively small (Fig. 3c). To accurately designate the different sublineages, the ranges of nucleotide substitutions between viruses within and between lineages were calculated (Table 2). Overall, the genetic distances within the main lineages A and B were comparable (0 to 129 nucleotide and 0 to 126 nucleotide substitutions between viruses within lineage A and B, respectively). Between lineage A1 and A2, the genetic distance ranged from 68 to 129 nucleotide substitutions. The genetic differences between viruses within the B1 and B2 lineage ranged from 0 to 88 and 0 to 73 nucleotide substitutions, respectively, while between B1 and B2, the genetic distance ranged between 62 and 126 nucleotide substitutions. The genetic distances between viruses within the lineages A2.1, A2.2.1, and A2.2.2 ranged between 0 and 59 nucleotide substitutions, while between A2.1 and the two A2.2 lineages (A2.2.1 and A2.2.2), this distance was 21 to 75 and 38 to 87 for A2.2.1 and A2.2.2 lineage viruses, respectively. The distance between viruses from lineages A2.2.1 and A2.2.2 ranged between 26 and 70 nucleotide substitutions. Although genetic distances within lineage A and B were comparable, two new lineages had evolved within lineage A, while no new lineages have evolved in lineage B.

Based on this analysis, we suggest that the nucleotide substitutions between a newly sequenced virus and viruses from each known lineage can be used to determine to which lineage the new virus belongs. For example, a virus from lineage A2.2.2 should contain between 0 and 41 nucleotide substitutions compared to other lineage A2.2.2 viruses and between 26 and 70 nucleotide substitutions compared to A2.2.1 lineage viruses, etc. If the number of nucleotide substitutions falls within the ranges of each lineage as described in Table 2, the newly sequenced virus can be classified accordingly. Alternatively, if the genetic distance of a newly isolated virus exceeds the ranges of nucleotide substitutions between lineages as described in Table 2, this virus might be considered part of a new lineage. These calculations provide a means to determine which lineage a newly isolated virus belongs to.

Circulation of virus lineages over time. Among viruses obtained from hospitalized individuals between 2005 and 2021, one virus belonged to lineage A1, 15 to lineage A2.1, 21 to lineage A2.2.1, 44 to lineage A2.2.2, 14 to B1, and 35 to B2 (Fig. 4a). The earliest virus of lineage A2.2.2 was obtained in 2008, similar to that described by others (13). The last virus belonging to the A1 lineage was collected in 2006. In this data set, significantly more lineage A viruses (81/130, 62.3%) were detected than lineage B viruses (49/130, 37.7%; χ^2 , $P = 0.046$) (Table 3). It is possible that this is due to the selection bias of samples based on a cutoff quantitative real-time-PCR (qRT-PCR) C_T value of ≤ 25.0 . To investigate whether the underrepresentation of B lineage viruses was due to this C_T cutoff value, 33 additional samples with a C_T value of > 25.0 (C_T range of 25.3 to 30.7) were sequenced and genotyped based on lineage-specific amino acid residues of partial F gene sequences (10). Eleven of these viruses (33.3%) belonged to lineage A (1 A1 virus from 2006 and 10 A2 viruses) and 22 (66.6%) to lineage B (8 B1 and 14 B2). Thus, more lineage B viruses were detected in samples with a C_T value of > 25.0 than lineage A viruses.

Prevalence of HMPV lineages in patients with RTI consulting general practitioners. To address circulation of HMPV lineages among patients with RTIs consulting general practitioners, 144 viruses collected from samples obtained from patients with RTIs consulting general practitioners in the Netherlands between 2005 and 2021 were genotyped by sequencing

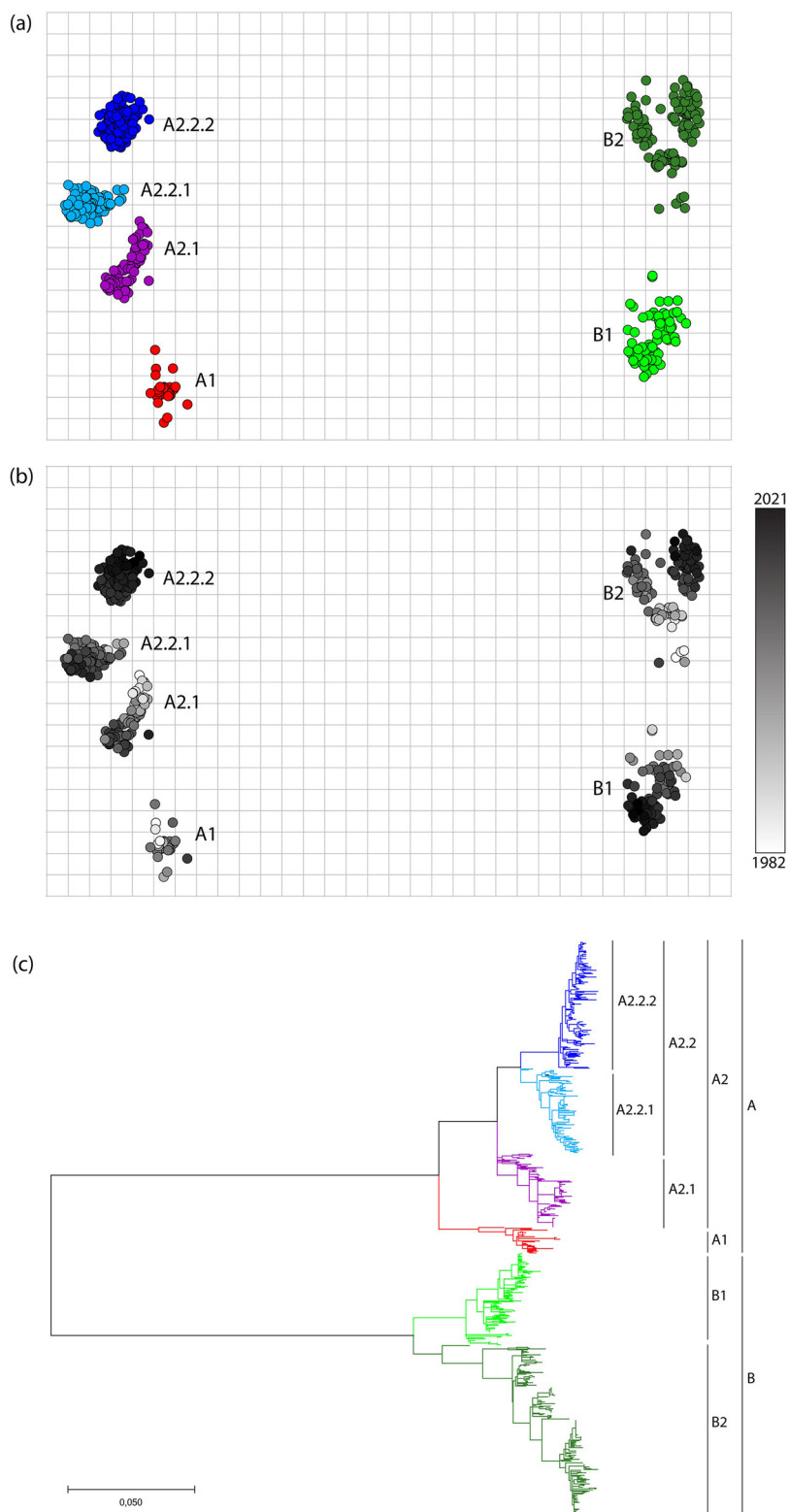


FIG 3 Genetic evolution of HMPV based on full-length F gene sequences. (a and b) A genetic map was generated from full-length F gene sequences obtained in this study combined with those from GenBank ($n = 744$) with color coding per lineage (a) or year of isolation (b). The axes represent genetic distance, with each square of the grid representing a 10-nucleotide distance. Sequences were color coded by year of isolation using a greyscale. (c) Maximum likelihood phylogenetic tree based on 744 full-length HMPV F gene sequences. The tree was reconstructed using the GTR +G +I substitution model with 1,000 bootstraps. The tree is color coded according to the colors used in the map shown in a. The tree with bootstrap values and isolate names is shown in Fig. S5. Markers on the right indicate virus lineages and sublineages.

TABLE 2 Nucleotide substitutions within and between all HMPV lineages including the proposed A2.2.1 and A2.2.2 lineages

Lineage	A1	A2.1	A2.2.1	A2.2.2	B1	B2
A1	0–57	68–114	68–122	84–129	234–275	236–264
A2.1		0–59	21–75	39–87	239–278	240–275
A2.2.1			0–45	26–70	247–288	248–284
A2.2.2				0–41	245–281	240–274
B1					0–88	62–126
B2						0–73

of partial F and G genes. The C_T values of these samples determined by qRT-PCR ranged from 18.6 to 38.1. Phylogenetic analysis of these sequences indicated clustering of the viruses in the six lineages similar to the viruses obtained from hospitalized patients (Fig. 4b and 5; Fig. S6). Viruses belonging to lineage A1 were not detected. From the 144 viruses, 71 (49.3%) clustered in lineage A and 73 (50.7%) in lineage B (Table 3). In the samples with C_T values of ≤ 25.0 , the numbers of A and B lineage viruses were not significantly different (20/34 [58.8%] lineage A and 14/34 [41.2%] lineage B, $P = 0.4651$). In samples with a C_T value of >25.0 , the numbers of lineage A (52/110, 47.2%) and lineage B viruses (58/110, 52.8%) were not significantly different ($P \geq 0.05$), although a slightly higher number of lineage B viruses was detected. This contrasted with that of samples isolated from hospitalized patients, in which a significantly higher number of lineage B viruses was detected among samples with a C_T value of >25.0 (Table 3).

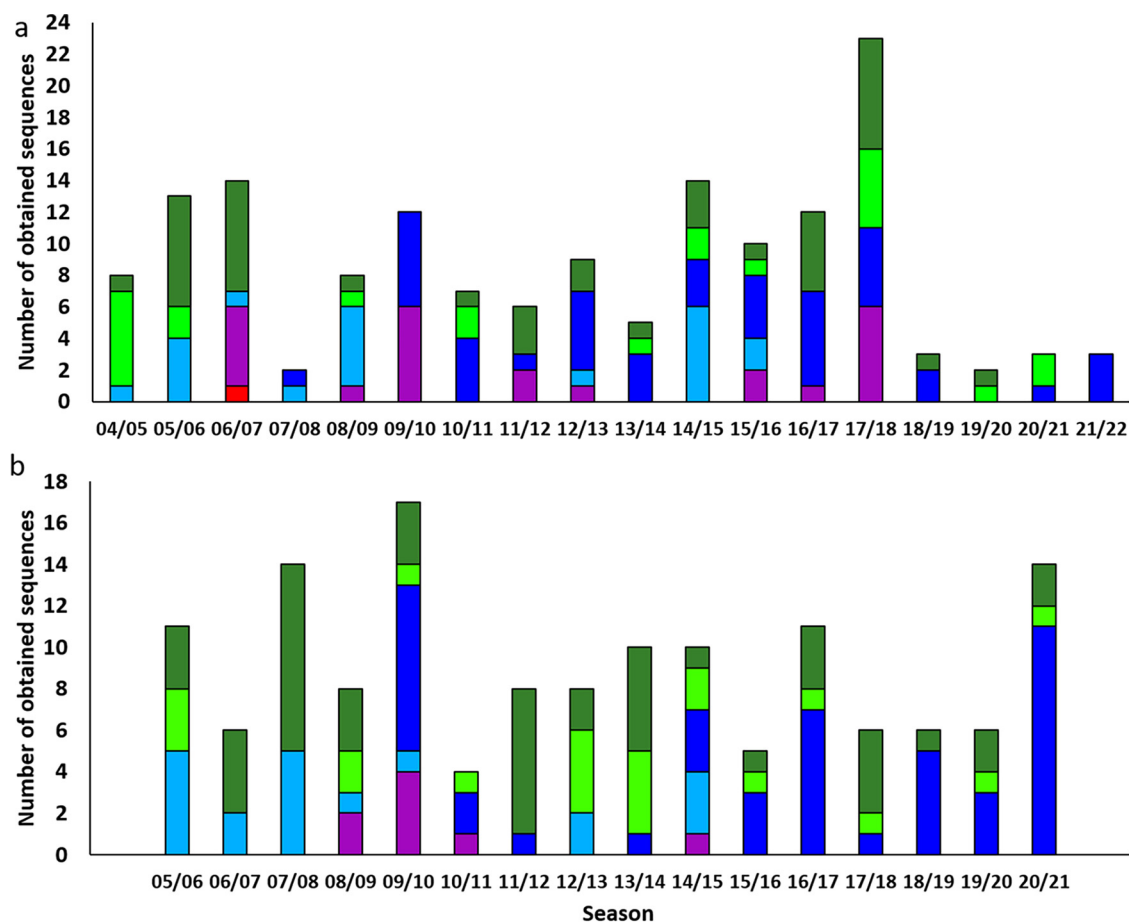


FIG 4 (a) The number of obtained sequences from each lineage of viruses isolated from hospitalized patients suffering from RTIs per epidemiologic season. Seasons started in week 40 and ended in week 39 of the next year. Samples were selected based on a C_T value of ≤ 25.0 , with exceptions as described in the Materials and Methods section, and genotyped by phylogenetic analyses of whole-genome sequences. (b) Clustering of viruses isolated from patients suffering from RTIs visiting general practitioners per year. Samples were selected as described in the Materials and Methods section and genotyped based on phylogenetic analyses of partial F and G gene sequences.

TABLE 3 Numbers of viruses belonging to lineage A or lineage B sorted by C_T values (≤ 25.0 and > 25.0) from both hospitalized patients and people visiting general practitioners

Virus lineage	$C_T \leq 25.0$		$C_T > 25.0$		Total	
	A	B	A	B	A	B
Hospitalized patients	81/130 (62.3%)	49/130 (37.6%)	11/33 (33.3%)	22/33 (66.7%)	92/163 (56.4%)	71/163 (43.5%)
General practitioners	20/34 (58.8%)	14/34 (41.2%)	52/110 (47.2%)	58/110 (52.8%)	72/144 (50.0%)	72/144 (50.0%)
Total	101/164 (61.6%)	63/164 (38.4%)	62/143 (43.4%)	80/143 (56.6%)	— ^a	—

^a—, no data.

DISCUSSION

After the identification of HMPV in 2001, it was reported that worldwide viruses clustered into two genetic lineages that were also antigenically distinct (A and B), which both split into two genetic lineages (A1, A2, B1, and B2) (10). Here, we investigated the evolution of HMPV over the last decades in the Netherlands using phylogenetic analysis of full genome sequences. In this study, phylogenetic analysis of 130 full-length genomes of viruses collected from hospitalized patients with respiratory tract infections between 2005 and 2021 indicated that viruses within our sample set clustered in six genetic lineages: A1, A2.1, A2.2.1, A2.2.2, B1, and B2. Additionally, these data indicate that since 2003, a split in the A2 lineage into lineages A2.1 and A2.2 was observed, and since 2008, lineage A2.2 split into lineages A2.2.1 and A2.2.2. All viruses in lineage A2.2.2 were relatively recent viruses, with the first collected in 2008, and since 2015, this lineage is primarily defined by viruses with a duplication in the G gene. A gradual evolution over time was only clear for lineage A2.2.2 and not so much for the other lineages.

In this study, HMPV evolution based on full genome sequences was studied using viruses circulating in the Netherlands between 2005 and 2021. However, analysis of the F gene sequences of these Dutch isolates together with those circulating worldwide (Fig. 2; Fig. S1 to S3 in the supplemental material) demonstrated that the Dutch strain clustered among those collected worldwide, indicating that the circulation of different genotypes is not specific for the Netherlands.

Viruses clustering in the A2.2.2 lineage with a 111-nucleotide or 180-nucleotide duplication in the G gene have been reported to circulate worldwide since 2014 and 2017, respectively, and the prevalence of viruses carrying a duplication in the G gene has increased over the years (16–18, 20, 28). In our study, although the number of viruses was limited, all A2.2.2 viruses circulating between 2018 and 2021 carried a 111-nucleotide or 180-nucleotide duplication in the G gene.

In line with other studies, no viruses belonging to lineage A1 were detected after 2006. As most qRT-PCR assays used for HMPV diagnostics target the N gene, one explanation for this could be that viruses from this lineage are still circulating but became undetectable due to mutations in the region used in diagnostic assays. More likely, this virus lineage is now extinct or circulating in an area without surveillance.

The samples for the study in hospitalized patients were selected on the basis of C_T values of ≤ 25 and a maximum of 10 samples per year; therefore, no conclusions can be drawn on the yearly incidence of viruses from the different lineages. However, our study does show that viruses from the different lineages circulate in the same year. Among viruses obtained from hospitalized patients in the Netherlands, a higher number of lineage A than lineage B viruses was detected for most years. This is in contrast to the study among patients with RTIs visiting general practitioners, where lineage A and B viruses were equally often detected. This could be explained by the selection bias of samples from hospitalized patients with a C_T value of ≤ 25.0 , as in samples with a C_T value of > 25.0 a higher number of lineage B than lineage A viruses was detected. Additionally, among all samples obtained in this study (from both hospitalized individuals and people visiting general practitioners combined), there was a significantly lower number of B lineage viruses among samples with a C_T value of < 25.0 (101/164 [61.6%] lineage A, 63/164 [38.4%] lineage B; $P = 0.0347$). This indicates that, in general, in samples with a C_T value of ≤ 25.0 , lineage A viruses are more often

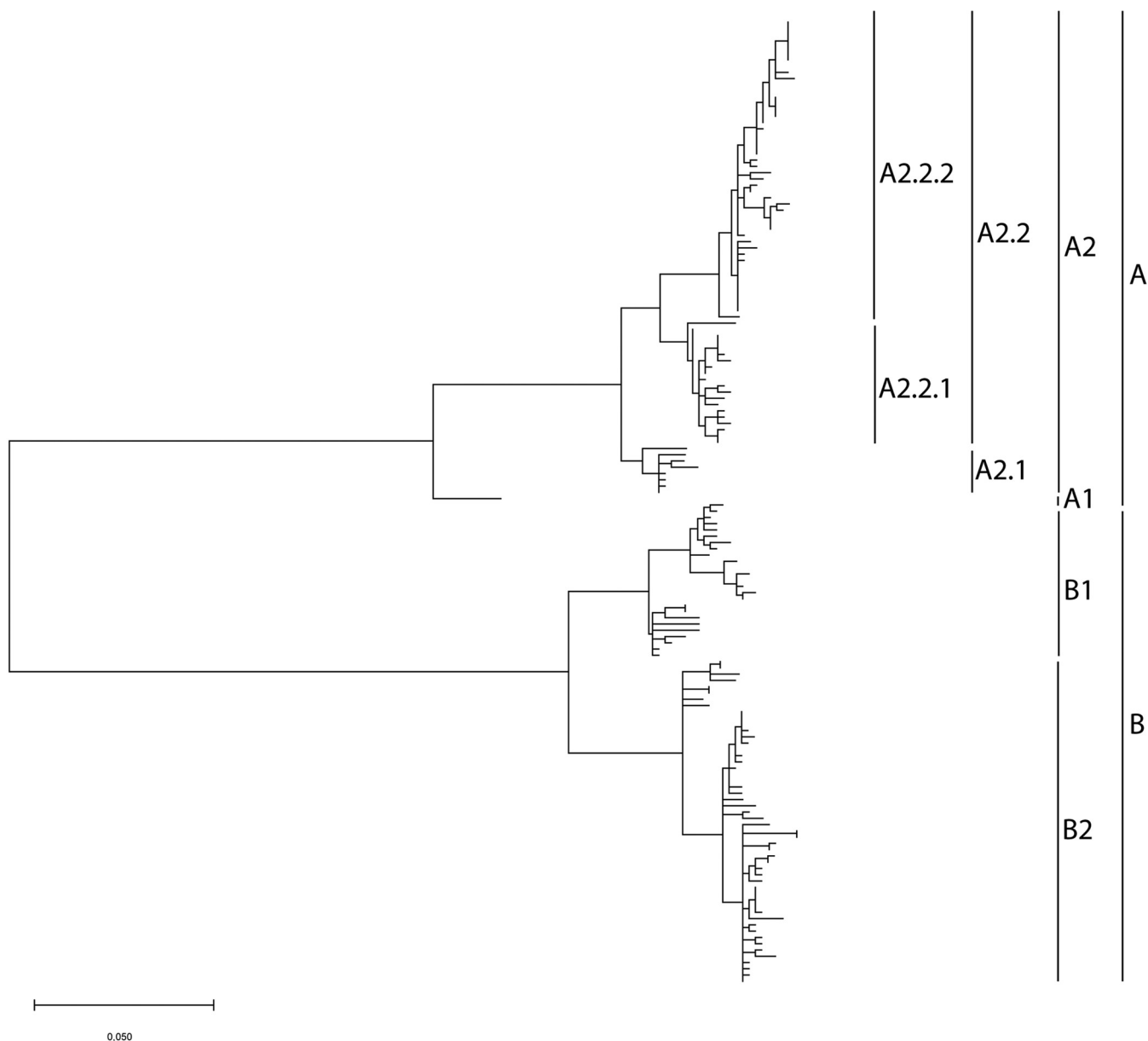


FIG 5 Maximum likelihood phylogenetic tree reconstructed from partial F gene sequences from viruses obtained from patients with RTIs visiting general practitioners combined with those of six viruses obtained from hospitalized individuals. The tree was reconstructed using the GTR +G +I substitution model with 1,000 bootstraps. The tree with bootstrap values and isolate names is shown in Fig. S6. Lines indicate virus lineages (e.g., A2).

detected than lineage B viruses. Whether this difference in C_T values between lineage A and B viruses corresponds to differences in clinical impact remains to be investigated. Of note, since 2019, diagnostics for detection of HMPV for hospitalized patients has switched from in-house primers and probes to the Panther Fusion System (Hologic), of which the primer and probe sequences are unknown. Primers and probes used to detect HMPV in samples from patients consulting general practitioners were similar to those used to detect HMPV in samples from hospitalized individuals before the introduction of the Panther Fusion System. These differences in assays used for virus diagnostics might result in different sensitivities to detect A or B lineage viruses after 2019.

One of the challenges of comparing HMPV molecular evolution studies is the use of different classification criteria for designating lineages. For other viruses, criteria have been used according to a more unified system, such as for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and influenza virus. For SARS-CoV-2, the Pangolin nomenclature system provides detailed cluster information (29), similar to the Global Initiative on Sharing

Avian Influenza Data (GISAID) nomenclature system (<https://www.gisaid.org/resources/statements-clarifications/clade-and-lineage-nomenclature-aids-in-genomic-epidemiology-of-active-hcov-19-viruses/>), while the Nextstrain nomenclature system offers a more large-scale view of clade trends (<https://nextstrain.org/blog/2020-06-02-SARSCoV2-clade-naming>). The influenza A(H5) nomenclature system designates new lineages based on monophyletic clusters with bootstrap support values greater than 70% with defined average pairwise p distances within and between clades (30). Since 2016, a nomenclature system was proposed for swine influenza A viruses that is in concordance with the nomenclature system for influenza A(H5) viruses (31). While a robust nomenclature system for these viruses is in place, attempts to unify a nomenclature system for RSV have recently started (32). Here, we propose a simple, yet robust, HMPV classification system based on genetic clustering of F gene sequences. The number of nucleotide substitutions of a newly sequenced virus compared to the ranges of nucleotide substitutions within and between the known virus lineages can be used to determine to which lineage a virus belongs. A newly sequenced virus belonging to lineage A2.2.2 should fall within the range of nucleotide substitutions compared to A2.2.1, A2.1, A1, B1, and B2, as described in this study. Alternatively, a virus of which the number of nucleotide substitutions exceeds the ranges of nucleotide substitutions between each of these lineages could be considered part of a new lineage. In addition to this proposed numerical system, to achieve a uniform system for naming new HMPV isolates, we suggest using the guidelines of the International Committee on Taxonomy of Viruses (ICTV) for other viruses, including country codes according to the ISO 3166-1 alpha-2 standards, that is, HMPV/country/isolate number/year/lineage (e.g., HMPV/NL/17/06/A).

The proposed classification system described above supports the split between the A2.2.1 and A2.2.2 lineages. Although a split within lineage B2 seems to be developing, the data thus far do not support this split yet. These findings are in accordance with a study reporting division of the A2.2 lineage into A2.2.1 and A2.2.2 (13), which were referred to as A2b1 and A2b2. However, for HMPV nomenclature, the A and B lineages were initially named as such to define two lineages that were antigenically distinct. The A and B lineages both split into two phylogenetic lineages, which were not antigenically different, named A1, A2, B1, and B2. To prevent confusion about antigenic differences in HMPV nomenclature, we propose to use the nomenclature that follows the guidelines of the ICTV for other viruses, for example, A2.1 instead of A2a and A2.2 instead of A2b (30, 33). There have been reports about the presence of an A2.3 lineage (named A2c in other reports [14]); however, viruses from this proposed lineage cluster within the A2.2.2 lineage in other studies as well as in ours.

Monitoring HMPV evolution is generally done by analyzing sequences of (partial) F or G genes (10, 11, 13, 16, 17, 20, 21, 34). More recently, analyzing whole-genome sequences is being explored as it adds important information compared to partial genome sequences, and the cost-efficiencies of obtaining full genomes by next-generation sequencing have increased considerably (22). Here, we show that phylogenetic clustering based on whole-genome sequences is similar to that of full-length F or partial F gene sequences, although trees reconstructed from partial F gene sequences are slightly more variable.

In conclusion, this study reports on the emergence and potential extinction of genetic HMPV lineages throughout recent decades in the Netherlands. This study suggests that using whole-genome sequences of HMPV captures similar phylogenetic information as the use of full-length F gene sequences. However, new evolutionary events, such as the reported duplication in the G gene, will be missed when only the full-length F gene is analyzed. Finally, phylogenetic analyses were used to propose robust classification criteria for future HMPV lineage designation and virus nomenclature.

MATERIALS AND METHODS

Sample selections from hospitalized patients. HMPV-positive samples obtained from patients hospitalized for RTIs between 2005 and 2021 with a qRT-PCR C_T value of ≤ 25.0 were selected. For samples obtained before the year 2019, qRT-PCR was performed as described previously (35). For samples obtained after 2019, qRT-PCR was performed using a Panther Fusion System (Hologic), of which the primer and probe sequences are unknown. To obtain an even distribution throughout the years, 10 samples were selected from each year. For seasons with more than 10 samples with a C_T value of ≤ 25.0 , the samples with the lowest C_T values were selected. For seasons with less than 10 samples with a C_T value

of ≤ 25.0 , additional samples were selected with a C_T value between 25.0 and 29.0. For some years, either a total number of 10 samples could not be obtained or 10 full-genome sequences could not be obtained. Nine additional samples obtained between 1993 and 2004 were selected based on available materials. Already available full-length genome sequences of prototype viruses for each lineage (e.g., NL/1/00 for lineage A1 and CAN83-97 for lineage A2.1) were also included for references. Patient data were anonymized before receiving the samples, and as a consequence, additional medical ethics review was not needed as consented by the Erasmus MC Medical Ethical Board (MEC-2015-306).

Selection of samples obtained from patients with RTIs consulting general practitioners. The general practitioner (GP) sentinel surveillance by the Nivel Netherlands Institute for Health Services Research Primary Care Database and the National Institute for Public Health and the Environment (Rijksinstituut voor Volksgezondheid en milieu; RIVM) reference laboratory (about 40 GP practices spread throughout the Netherlands representing about 0.7% of the Dutch population) collected nasopharyngeal and oropharyngeal swabs combined in one tube of virus transport medium from patients consulting the GP for an RTI (including influenza-like illness). Based on weekly virological records of about 20 clinical diagnostic laboratories in the Netherlands under the auspices of the Dutch Working Group for Clinical Virology (<https://www.rivm.nl/virologische-weekstaten>), the HMPV epidemic period for each season from 2005 to 2020 was determined. From these periods, a mean of 95 samples (range of 84 to 100) per epidemic period evenly spread over the epidemic period were selected from the GP sentinel surveillance biobank that were negative by qRT-PCR for influenza virus, RSV, rhinovirus, and enterovirus. These samples were subjected to qRT-PCR for HMPV as described previously (10), following nucleic acid isolation by MagNA Pure 96 (Roche) extraction of 200 μ L of specimen using a DNA and viral NA small volume kit (Roche). A mean of 10 (range of 4 to 19) samples per epidemic period were found to be positive for HMPV. In 2020/2021, there was no HMPV season due to coronavirus disease 2019 (COVID-19) measures. From January 2021 onward, all collected samples in the GP sentinel surveillance were prospectively tested for HMPV by qRT-PCR. A total of 15 HMPV-positive samples were detected up to October 2021 that were included in the study. Due to lower viral loads in the GP sentinel surveillance patients than in hospitalized patients, all qRT-PCR HMPV-positive specimens irrespective of C_T value (median C_T value of 29; range of 19 to 39) were subjected to partial F gene and G gene PCR and Sanger sequencing. For these samples, a nested PCR was performed to increase efficiency for lower viral load specimens (F PCR: forward primer 5'-TGTTGTGCGGCAR TTTTCAGAC-3' and reverse primers 5'-GCACTGCTTAGGATTCTGTTTGA-3' and 5'-TGCACTGTTAGRATTTTGTTC GAC-3'; nested F PCR: forward primers 5'-TGGAATAACACCAGCAATATCCTT-3' and 5'-AATGCAGGGATAACACCA GCAAT-3' and reverse primer 5'-CTGTCRCYCAACTTGTCT-3'; G PCR: forward primer 5'-TACAAAACAAGAACA TGGGACAAG-3' and reverse primers 5'-GTTTCACTGAAAGATATTACTCCTT-3' and 5'-GTCTCAAAAAGAAATC ACTCCTT-3'; nested G PCR: forward primers 5'-ATGGAGGTGAAAGTGGAGAACATTC-3' and 5'-ATGGAAGTAAG AGTGGAGAACATTC-3' and reverse primer 5'-GAGATAGACATTAACAGTGGATTC-3'). The primers used for Sanger sequencing of PCR products were the same as those used for the nested PCR.

Virus nomenclature. Nomenclature of viruses followed the guidelines of the ICTV for other viruses, for example, HMPV/country/isolate number/year/lineage (i.e., HMPV/NL/17/06/A) (33).

MinION sequencing and PCR of the G gene and partial HMPV F gene. HMPV-positive samples were used for RNA isolation and MinION sequencing as described previously (27). PCR of the HMPV F gene was performed as described previously (27). The presence or absence of 111-nucleotide and 180-nucleotide duplications was confirmed by Sanger sequencing as described previously (27). Gaps in consensus sequences were resolved by PCR amplification of the missing amplicon followed by Sanger sequencing as described previously (27). Sequencing of partial F genes (nucleotide positions 778 to 1226 of the F open reading frame [ORF]) for genotyping of samples obtained from hospitalized individuals with a C_T value of >25.0 was performed as described previously (10). All full-length F gene sequences available from GenBank were downloaded as a FASTA alignment file.

Phylogenetic analysis. Nucleotide sequence alignments were generated using MAFFT (<https://mafft.cbrc.jp/alignment/server/>). Maximum likelihood phylogenetic analyses were conducted using the MEGA-X software with the best fit DNA model as determined by the MEGA-X software with 1,000 bootstraps. The best fit model for each phylogenetic tree is described in the figure legends. Pairwise distances were generated by MEGA-X software.

Clustering of HMPV in genetic maps. To visualize the clustering of HMPV sequences based on the number of nucleotide substitutions between them, genetic maps representing HMPV nucleotide sequences were generated essentially as described previously (36, 37). Nucleotide sequence alignments were generated for various HMPV open reading frames using MAFFT, and a genetic distance matrix was generated with all pairwise numbers of nucleotide substitutions. A multidimensional scaling algorithm was used to generate two-dimensional representations of the matrices with Euclidian distances (36).

Data availability. All full genome sequences from viruses isolated from hospitalized individuals generated in this study were submitted to GenBank with accession numbers [OL794355](https://doi.org/10.1128/mbio.02280-22) to [OL794483](https://doi.org/10.1128/mbio.02280-22). Partial F and G gene sequences from viruses isolated from patients visiting general practitioners were submitted to GenBank with accession numbers [ON461481](https://doi.org/10.1128/mbio.02280-22) to [ON461747](https://doi.org/10.1128/mbio.02280-22).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, TIF file, 2.3 MB.

FIG S2, TIF file, 1.9 MB.

FIG S3, TIF file, 2.1 MB.

FIG S4, TIF file, 0.9 MB.

FIG S5, PDF file, 1.2 MB.

FIG S6, TIF file, 2.8 MB.

TABLE S1, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

We thank Euníce Then for helping with qRT-PCR analysis of clinical specimens and sequencing of viruses obtained from specimens from patients visiting general practitioners; Gé Donker, Mariëtte Hooiveld, and Janneke Hendriksen from Nivel and the Nivel Primary Care Database GP's for collecting specimens from patients with RTI consulting the GP's; and the Erasmus MC virology diagnostics department for performing qRT-PCR analysis on clinical specimens from hospitalized patients. We thank the Dutch Working Group for Clinical Virology for the opportunity to use the virology weekly records database to determine HMPV epidemic periods in the Netherlands.

This work was supported in part by Horizon 2020 research and innovation program grant number 874735 (VEO) from the European Union. The Dutch Ministry of Health, Welfare and Sport funded the GP sentinel surveillance.

REFERENCES

1. Van Den Hoogen BG, De Jong JC, Groen J, Kuiken T, De Groot R, Fouchier RAM, Osterhaus ADME. 2001. A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat Med* 7: 719–724. <https://doi.org/10.1038/89098>.
2. Williams JV, Wang CK, Yang C, Tollefson SJ, House FS, Heck JM, Chu M, Brown JB, Lintao LD, Quinto JD, Chu D, Spaete RR, Edwards KM, Wright PF, Crowe JE, Jr. 2006. The role of human metapneumovirus in upper respiratory tract infections in children: a 20-year experience. *J Infect Dis* 193: 387–395. <https://doi.org/10.1086/499274>.
3. Jain S, Self WH, Wunderink RG, Fakhran S, Balk R, Bramley AM, Reed C, Grijalva CG, Anderson EJ, Courtney DM, Chappell JD, Qi C, Hart EM, Carroll F, Trabue C, Donnelly HK, Williams DJ, Zhu Y, Arnold SR, Ampofo K, Waterer GW, Levine M, Lindstrom S, Winchell JM, Katz JM, Erdman D, Schneider E, Hicks LA, McCullers JA, Pavia AT, Edwards KM, Finelli L, CDC EPIC Study Team. 2015. Community-acquired pneumonia requiring hospitalization among U.S. adults. *N Engl J Med* 373:415–427. <https://doi.org/10.1056/NEJMoa1500245>.
4. O'Brien KL, Baggett HC, Brooks WA, Feikin DR, Hammitt LL, Higdon MM, Howie SR, Deloria Knoll M, Kotloff KL, Levine OS, Madhi SA, Murdoch DR, Prosperi C, Scott JAG, Shi Q, Thea DM, Wu Z, Zeger SL, Adrian PV, Akarasewi P, Anderson TP, Antonio M, Awori JO, Baillie VL, Bunthi C, Chipeta J, Chisti MJ, Crawley J, DeLuca AN, Driscoll AJ, Ebruke BE, Endtz HP, Fancourt N, Fu W, Goswami D, Groome MJ, Haddix M, Hossain L, Jahan Y, Kagucia EW, Kamau A, Karron RA, Kazungu S, Kourouma N, Kuwanda L, Kwenda G, Li M, Machuka EM, Mackenzie G, Mahomed N, et al. 2019. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *Lancet* 394:757–779. [https://doi.org/10.1016/S0140-6736\(19\)30721-4](https://doi.org/10.1016/S0140-6736(19)30721-4).
5. O'Bryant SC, Momin Z, Camp E, Jones J, Meskill S. 2022. Longitudinal evaluation of pediatric respiratory infections. *J Clin Virol* 148:105084. <https://doi.org/10.1016/j.jcv.2022.105084>.
6. Loubet P, Mathieu P, Lenzi N, Galtier F, Lainé F, Lesieur Z, Vanhems P, Duval X, Postil D, Amour S, Rogez S, Lagathu G, L'Honneur A-S, Foulongne V, Houhou N, Lina B, Carrat F, Launay O, Fluvac study group. 2021. Characteristics of human metapneumovirus infection in adults hospitalized for community-acquired influenza-like illness in France, 2012–2018: a retrospective observational study. *Clin Microbiol Infect* 27:127.e1–127.e6. <https://doi.org/10.1016/j.cmi.2020.04.005>.
7. Barrera-Badillo G, Olivares-Flores B, Ruiz-López A, Fierro-Valdez MÁ, Gutiérrez-Vargas RI, López-Martínez I. 2020. Human metapneumovirus: etiological agent of severe acute respiratory infections in hospitalized and deceased patients with a negative diagnosis of influenza. *Pathogens* 9:85. <https://doi.org/10.3390/pathogens9020085>.
8. Wang X, Li Y, Deloria-Knoll M, Madhi SA, Cohen C, Ali A, Basnet S, Bassat Q, Brooks WA, Chittaganpitch M, Echavarría M, Fasce RA, Goswami D, Hirve S, Homaira N, Howie SRC, Kotloff KL, Khuri-Bulos N, Krishnan A, Lucero MG, Lupisan S, Mira-Iglesias A, Moore DP, Moraleda C, Nunes M, Oshitani H, Owor BE, Polack FP, O'Brien KL, Rasmussen ZA, Rath BA, Salimi V, Scott JAG, Simões EAF, Strand TA, Thea DM, Treurnicht FK, Vaccari LC, Yoshida L-M, Zar HJ, Campbell H, Nair H, Libster R, Otieno G, Joundi I, Broor S, Nicol M, Amarchand R, Shi T, López-Labrador FX, et al. 2021. Global burden of acute lower respiratory infection associated with human metapneumovirus in children under 5 years in 2018: a systematic review and modelling study. *Lancet Glob Heal* 9:e33–e43. [https://doi.org/10.1016/S2214-109X\(20\)30393-4](https://doi.org/10.1016/S2214-109X(20)30393-4).
9. Shafagati N, Williams J. 2018. Human metapneumovirus—what we know now. *F1000Res* 7:135. <https://doi.org/10.12688/f1000research.12625.1>.
10. Van Den Hoogen BG, Herfst S, Sprong L, Cane PA, Forleo-Neto E, De Swart RL, Osterhaus ADME, Fouchier RAM. 2004. Antigenic and genetic variability of human metapneumoviruses. *Emerg Infect Dis* 10:658–666. <https://doi.org/10.3201/eid1004.030393>.
11. Williams JV, Harris PA, Tollefson SJ, Halburnt-Rush LL, Pingsterhaus JM, Edwards KM, Wright PF, Crowe JE. 2004. Human metapneumovirus and lower respiratory tract disease in otherwise healthy infants and children. *N Engl J Med* 350:443–450. <https://doi.org/10.1056/NEJMoa025472>.
12. Huck B, Scharf G, Neumann-Haefelin D, Puppe W, Weigl J, Falcone V. 2006. Novel human metapneumovirus sublineage. *Emerg Infect Dis* 12: 147–150. <https://doi.org/10.3201/eid1201.050772>.
13. Nao N, Saikusa M, Sato K, Sekizuka T, Usuku S, Tanaka N, Nishimura H, Takeda M. 2020. Recent molecular evolution of human metapneumovirus (HMPV): subdivision of HMPV A2b strains. *Microorganisms* 8:1280. <https://doi.org/10.3390/microorganisms8091280>.
14. Nidaira M, Taira K, Hamabata H, Kawaki T, Gushi K, Mahoe Y, Maeshiro N, Azama Y, Okano S, Kyan H, Kudaka J, Hiroyuki T, Noda M, Kimura H. 2013. Molecular epidemiology of human metapneumovirus from 2009 to 2011 in Okinawa, Japan. *Jpn J Infect Dis* 66:337–340. <https://doi.org/10.7883/yoken.65.337>.
15. Carr MJ, Waters A, Fenwick F, Toms GL, Hall WW, O'Kelly E. 2008. Molecular epidemiology of human metapneumovirus in Ireland. *J Med Virol* 80: 510–516. <https://doi.org/10.1002/jmv.21081>.
16. Saikusa M, Nao N, Kawakami C, Usuku S, Sasao T, Toyozawa T, Takeda M, Okubo I. 2017. A novel 111-nucleotide duplication in the G gene of human metapneumovirus. *Microbiol Immunol* 61:507–512. <https://doi.org/10.1111/1348-0421.12543>.
17. Saikusa M, Kawakami C, Nao N, Takeda M, Usuku S, Sasao T, Nishimoto K, Toyozawa T. 2017. 180-Nucleotide duplication in the G gene of human metapneumovirus A2b subgroup strains circulating in Yokohama City, Japan, since 2014. *Front Microbiol* 8:402. <https://doi.org/10.3389/fmicb.2017.00402>.
18. Piñana M, Vila J, Gimferrer L, Valls M, Andrés C, Codina MG, Ramón J, Martín MC, Fuentes F, Saiz R, Alcubilla P, Rodrigo C, Pumarola T, Antón A. 2017. Novel human metapneumovirus with a 180-nucleotide duplication in the G gene. *Future Microbiol* 12:565–571. <https://doi.org/10.2217/fmb-2016-0211>.
19. Boivin G, Mackay I, Sloots TP, Madhi S, Freymuth F, Wolf D, Shemer-Avni Y, Ludewick H, Gray GC, LeBlanc É. 2004. Global genetic diversity of human metapneumovirus fusion gene. *Emerg Infect Dis* 10:1154–1157. <https://doi.org/10.3201/eid1006.031097>.
20. Saikusa M, Nao N, Kawakami C, Usuku S, Tanaka N, Tahara M, Takeda M, Okubo I. 2019. Predominant detection of the subgroup A2b human metapneumovirus strain with a 111-nucleotide duplication in the G gene in Yokohama city, Japan in 2018. *Jpn J Infect Dis* 72:350–352. <https://doi.org/10.7883/yoken.JIID.2019.124>.
21. Arshad Y, Rana MS, Ikram A, Salman M, Aamir UB, Zaidi SSZ, Alam MM, Sharif S, Shaikat S, Khurshid A, Hakim R, Mujtaba G, Umair M, Sattar S,

- Bostan N. 2022. Molecular detection and genetic characterization of human metapneumovirus strains circulating in Islamabad, Pakistan. *Sci Rep* 12:2790. <https://doi.org/10.1038/s41598-022-06537-5>.
22. Dudas G, Bedford T. 2019. The ability of single genes vs full genomes to resolve time and space in outbreak analysis. *BMC Evol Biol* 19:232. <https://doi.org/10.1186/s12862-019-1567-0>.
23. Pollett S, Trovão NS, Tan Y, Eden JS, Halpin RA, Bera J, Das SR, Wentworth D, Ocaña V, Mendocilla SM, Álvarez C, Calisto ME, García J, Halsey E, Ampuero JS, Nelson MI, Leguia M. 2018. The transmission dynamics and diversity of human metapneumovirus in Peru. *Influenza Other Respir Viruses* 12:508–513. <https://doi.org/10.1111/irv.12537>.
24. Xu Y, Lewandowski K, Jeffery K, Downs LO, Foster D, Sanderson ND, Kavanagh J, Vaughan A, Salvagno C, Vipond R, Carroll M, Danby R, Peto T, Crook D, Walker AS, Matthews PC, Pullan ST. 2020. Nanopore metagenomic sequencing to investigate nosocomial transmission of human metapneumovirus from a unique genetic group among haematology patients in the United Kingdom. *J Infect* 80:571–577. <https://doi.org/10.1016/j.jinf.2020.02.003>.
25. Xu Y, Lewandowski K, Lumley S, Sanderson N, Ali V, Vipond R, Carroll M, Jeffery K, Foster D, Walker AS, Peto T, Crook D, Pullan ST, Matthews PC. 2018. Nanopore metagenomic sequencing of full length human metapneumovirus (HMPV) within a unique sub-lineage. *bioRxiv*. <https://doi.org/10.1101/496687>.
26. Kamau E, Oketch JW, De Laurent ZR, Phan MVT, Agoti CN, Nokes DJ, Cotten M. 2020. Whole genome sequencing and phylogenetic analysis of human metapneumovirus strains from Kenya and Zambia. *BMC Genomics* 21: 5. <https://doi.org/10.1186/s12864-019-6400-z>.
27. Groen K, van Nieuwkoop S, Bestebroer TM, Fraaij PL, Fouchier RAM, van den Hoogen BG. 2021. Whole genome sequencing of human metapneumoviruses from clinical specimens using MinION Nanopore technology. *Virus Res* 302:198490. <https://doi.org/10.1016/j.virusres.2021.198490>.
28. Pinana M, Vila J, Maldonado C, Galano-Frutos JJ, Valls M, Sancho J, Andres FXNC, Martín-Gomez MT, Esperalba J, Codina MG, Pumarola T, Anton A. 2020. Insights into immune evasion of human metapneumovirus: novel 180- and 111-nucleotide duplications within viral G gene throughout 2014–2017 seasons in Barcelona, Spain. *J Clin Virol* 132:104590. <https://doi.org/10.1016/j.jcv.2020.104590>.
29. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, Du Plessis L, Pybus OG. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 5:1403–1407. <https://doi.org/10.1038/s41564-020-0770-5>.
30. Smith GJD, Donis RO, World Health Organization/World Organisation for Animal Health/Food and Agriculture Organization (WHO/OIE/FAO) H5 Evolution Working Group. 2015. Nomenclature updates resulting from the evolution of avian influenza A(H5) virus clades 2.1.3.2a, 2.2.1, and 2.3.4 during 2013–2014. *Influenza Other Respi Viruses* 9:271–276. <https://doi.org/10.1111/irv.12324>.
31. Anderson TK, Macken CA, Lewis NS, Scheuermann RH, Van Reeth K, Brown IH, Swenson SL, Simon G, Saito T, Berhane Y, Ciacci-Zanella J, Pereda A, Davis CT, Donis RO, Webby RJ, Vincent AL. 2016. A phylogeny-based global system and automated annotation tool for H1 influenza A viruses. *mSphere* 1:e00275-16. <https://doi.org/10.1128/mSphere.00275-16>.
32. Ramaekers K, Rector A, Cuypers L, Lemey P, Keyaerts E, van Ranst M. 2020. Towards a unified classification for human respiratory syncytial virus genotypes. *Virus Evol* 6:veaa052. <https://doi.org/10.1093/ve/veaa052>.
33. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. 2018. Virus taxonomy: the database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res* 46:D708–D717. <https://doi.org/10.1093/nar/gkx932>.
34. Perchetti GA, Wilcox N, Chu HY, Katz J, Khatri SK, LeClerq SC, Tielsch JM, Jerome KR, Englund JA, Kuypers J. 2021. Human metapneumovirus infection and genotyping of infants in rural Nepal. *J Pediatric Infect Dis Soc* 10: 408–416. <https://doi.org/10.1093/jpids/piaa118>.
35. Hoek RAS, Paats MS, Pas SD, Bakker M, Hoogsteden HC, Boucher CAB, Van Der Eerden MM. 2013. Incidence of viral respiratory pathogens causing exacerbations in adult cystic fibrosis patients. *Scand J Infect Dis* 45:65–69. <https://doi.org/10.3109/00365548.2012.708942>.
36. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Rimmelzwaan GF, Osterhaus ADME, Fouchier RAM. 2004. Mapping the antigenic and genetic evolution of influenza virus. *Science* 305:371–376. <https://doi.org/10.1126/science.1097211>.
37. Higgins DG. 1992. Sequence ordinations: a multivariate analysis approach to analysing large sequence data sets. *Comput Appl Biosci* 8:15–22. <https://doi.org/10.1093/bioinformatics/8.1.15>.