



HAL
open science

Use of whole-genome sequencing in the molecular investigation of care-associated HCoV-OC43 infections in a hematopoietic stem cell transplant unit

Delphine Beury, Léa Fléchon, Florence Maurier, Ségolène Caboche, Jean-Stéphane Varré, Helene Touzet, Jean Dubuisson, David Hot, Benoît Guéry, Anne Goffard

► To cite this version:

Delphine Beury, Léa Fléchon, Florence Maurier, Ségolène Caboche, Jean-Stéphane Varré, et al.. Use of whole-genome sequencing in the molecular investigation of care-associated HCoV-OC43 infections in a hematopoietic stem cell transplant unit. *Journal of Clinical Virology*, 2020, 122, pp.104206. 10.1016/j.jcv.2019.104206 . hal-02403094

HAL Id: hal-02403094

<https://hal.science/hal-02403094>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Major article

Use of whole-genome sequencing in the molecular investigation of care-associated HCoV-OC43 infections in a hematopoietic stem cell transplant unit.

Authors: Delphine Beury^{1*}, Léa Fléchon^{2*}, Florence Maurier¹, Ségolène Caboche¹, Jean-Stéphane Varré³, Hélène Touzet³, Karine Faure⁴, Jean Dubuisson¹, David Hot¹, Benoit Guery⁵, Anne Goffard^{□□1}

*: Contribute equally to the work

¹Delphine Beury: Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR8204 – Center for Infection and Immunity of Lille, F-59000 Lille, France, Email address: delphine.beury@pasteur-lille.fr

²Léa Fléchon: Bilille, Inserm, Univ. Lille, CNRS, Inria, CHU Lille, Institut Pasteur de Lille, Lille, France. Email address: lea.flechon.etu@univ-lille.fr

¹Florence Maurier: Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR8204 – Center for Infection and Immunity of Lille, F-59000 Lille, France. Email address: florence.maurier@pasteur-lille.fr

¹Ségolène Caboche: Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR8204 – Center for Infection and Immunity of Lille, F-59000 Lille, France. Email address: segolene.caboche@pasteur-lille.fr

³Jean-Stéphane Varré: Univ. Lille, CNRS, Inria, UMR 9189 - CRISAL - Centre de Recherche en Informatique Signal et Automatique de Lille, Lille, France. Email address: jean-stephane.varre@lifl.fr

³Hélène Touzet: Univ. Lille, CNRS, Inria, UMR 9189 - CRISAL - Centre de Recherche en Informatique Signal et Automatique de Lille, Lille, France. Email address: helene.touzet@lifl.fr

⁴ Karine Faure: CHU Lille, Maladies Infectieuses, 59000 Lille, France; Université de Lille, EA 7366, Recherche translationnelle, relations hôte-pathogènes, 59000 Lille, France. Email address: karine.faure@univ-lille.fr

¹Jean Dubuisson: Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR8204 – Center for Infection and Immunity of Lille, F-59000 Lille, France. Email address: jean.dubuisson@ibl.cnrs.fr

¹David Hot: Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR8204 – Center for Infection and Immunity of Lille, F-59000 Lille, France. Email address: david.hot@pasteur-lille.fr

⁵ Benoit Guery: Infectious Diseases Unit, Department of Medicine, Lausanne University Hospital and University of Lausanne, Rue du Bugnon 46, 1011 Lausanne, Switzerland. Email address: Benoit.Guery@unil.ch

□□ ¹ Corresponding author: Anne Goffard, MD, PhD, Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 – UMR8204 – Center for Infection and Immunity of Lille, F-59000 Lille, France. Email address: anne.goffard@univ-lille.fr

Words count: 3502

Author's contributions

Ségolène Caboche (SC), David Hot (DH), Jean-Stephane Varre (JSV), Hélène Touzet (HT), Karine Faure (KF), Benoit Guéry (BG), Jean Dubuisson (JD) and Anne Goffard (AG): Conceptualization

AG, KF: Resources

Delphine Beury (DB), Léa Fléchon (LF), Florence Maurier (FM) and SC: Investigations

DB, LF, FM: Formal analyses

FM, SC, DH, JSV, HT: Data curation

JSV, HT, KF, SC, DH, AG: Writing original data

BG, JD: Writing, review and editing

BG, JD, AG: Funding acquisition

AG: Project administration

Major article

Background

1 Although highly pathogenic human coronaviruses (HCoV), like SARS-CoV and
2 MERS-CoV, have emerged recently, the majority of HCoV infections are caused by
3 HCoV-229E, HCoV-NL63, HCoV-OC43 and HCoV-HKU1. These coronaviruses are
4 mainly recognized as causative agents of community-acquired infections. With the
5 exception of SARS and MERS-CoV, there are not many published studies about
6 health-care associated HCoV infections [1–5]. HCoV-229E, HCoV-NL63, HCoV-
7 OC43 and HCoV-HKU1 are usually associated with mild diseases in
8 immunocompetent patients, but can cause severe respiratory tract infections in
9 fragile populations. Indeed, development of molecular detection tests for diagnosis
10 has shown that HCoV are clearly responsible for severe infections in
11 immunocompromised patients, including hematopoietic stem cell transplant
12 recipients (HSCT) [6–8].
13 HCoV-OC43, HCoV-229E, HCoV-NL63 and HCoV-HKU1 account for 6.7 to 15.4% of
14 the viruses detected during respiratory infections in HSCT recipients [9]. Currently,
15 HCoV-OC43 is mainly detected around the world, circulating all year long with a
16 slight predominance during winter in temperate countries. The efficient transmission
17 via small droplets and a prolonged shedding by HSCT recipient's contribute to virus
18 dissemination and highlight the need for a better understanding of hospital-acquired
19 HCoV infections [10–13]. Although an outbreak of HCoV-NL63 respiratory infections
20 in a long-term care facility has recently been reported, a limited amount data are
21 available on the overall genomic characteristics of HCoV-OC43 healthcare-
22 associated infections [14].

23 The main objective of this work was to identify potential clusters of infection that may
24 stand for a hospital transmission. Next-Generation Sequencing (NGS) was used to
25 generate whole-genome sequences from nasal swabs and phylogenetic analyses
26 were then applied to characterize sequences. The secondary objective was to
27 identify molecular signature of circulating strains.

28 **Methods**

29 **Patients.** Eight HSCT recipients with HCoV-OC43 detected in nasal swabs by
30 multiplex respiratory viral PCR using Anyplex II RVS 16 detection kit at the University
31 hospital of Lille between 2013 and 2015 were included in the study. Demographic
32 and clinical data were collected. All patients were informed that they were included in
33 a research cohort and agreed that their biological samples could be used for
34 research purposes. Nasal swabs studied here were included into the biological
35 collection of the University Hospital of Lille declared to the French Ministry of High
36 Education and Research (reference number DC-2008-642). According to the French
37 laws, this declaration implies acceptance by an ethic committee. Patients are named
38 as MDSX for “Maladies Du Sang”X.

39 **Sequencing.** Viral genome sequencing was performed as described by Maurier et
40 al. [15]. Briefly, extracted viral RNA was reverse transcribed, PCR amplified using
41 multiple primer pairs and pooled before preparing sequencing libraries using NxSeq
42 AmpFREE LowDNA (Lucigen) library kit and barcoded with Illumina-compatible
43 adaptors. Libraries were paired-end sequenced in 2x300 cycles on Illumina MiSeq.
44 Sequencing analysis and genome assembly was performed as described by Maurier
45 et al. [15]

46 **Genomes annotation.** Annotation of the 8 newly sequenced genomes MDS2,
47 MDS4, MDS6, MDS11, MDS12, MDS14, MDS15 and MDS16, was performed by

48 comparison with the HCoV-OC43 isolate from Mexico (Genbank accession
49 KX344031) [16] using BlastN [17]. This isolate served as a reference and allowed us
50 to compute the coordinates of the following functional elements: ORF1a, ORF1b,
51 ns2A, HE, S, M and N genes.

52 **Phylogenetic analysis.** For the whole genome evolutionary analysis of MDS2,
53 MDS4, MDS6, MDS11, MDS12, MDS14, MDS15 and MDS16, we selected 40
54 published full-length human coronavirus genomes with diverse genotypes and
55 geographical origins, making a total of 48 sequences. We also considered more
56 specifically the phylogeny of *S* gene and included 12 additional sequences from
57 partially sequenced genomes from Caen hospital (France) and one additional
58 published sequence from China (Table supplementary data), making a total of 61
59 sequences. For each of the two datasets, three distantly related bovine
60 coronaviruses served as outgroup and multiple-sequence alignment was built using
61 MUSCLE from SeaView 4.6.4 followed by a manual correction step taking into
62 account protein coding sequences for coding regions [18,19]. Based on the multiple
63 sequence alignments, similarity between each pair of nucleotide sequences was
64 computed using the Sequence Manipulation Suite [20]. Phylogenetic trees were then
65 inferred using the maximum likelihood method implemented in PhyML 3.1 with a
66 general time reversible (GTR) nucleotide substitution model, a proportion of invariant
67 sites (+I) and a gamma rate heterogeneity (+ Γ) with 1000 bootstrap replicates [21].
68 The substitution model was chosen by optimizing the AIC score using smart model
69 selection [22]. A total of 30,815 nucleotides positions were included for the whole
70 genome analysis and 4,137 nucleotides positions for *S* gene. The software FigTree
71 1.4.4 was used to produce figures [23].

72 **Substitution analysis.** Amino acids comparative analyses for ORF1a, ORF1b,
73 ns2A, HE, S, M and N proteins were done on 58 sequences strains (8 MDS + 40
74 previously included and 10 others) for which complete genomes were available and
75 on 76 sequences for S gene only (8 MDS + 68 others) (Table Supplementary data).
76 For this analysis, we added new sequences compared to the phylogenetic analysis.
77 Indeed, we need to examine each position of the sequences separately, which
78 requires a larger support. By doing this, we reduced the chance to select erroneous
79 mutations. The goal is to ensure that the informative mutations found are
80 characteristic of a particular genotype: they are found only in this genotype and are
81 shared by all sequences known for this genotype. In both cases, the signature amino
82 acid substitutions were determined by selecting a subset of informative sites using
83 DIVEIN [24]. Colorization was performed manually to highlight conserved sites.

84 **Spatio-temporal and evolutionary dynamics.** We followed the methodology used
85 in [25]. Notably, we added Asiatic sequences that were also used in [25] and [26].
86 We estimated divergence times using BEAST 1.8.4 [12] on both whole genome and
87 S gene [27]. Exponential population size, relaxed clock with coalescent tree (BSP
88 distribution for S and uncorrelated exponential distribution for whole genome) was
89 tested to estimate the time of most recent common ancestor (tMRCA).

90 All information on the sequences used (identifiers, strains, bibliographical references)
91 is summed up in the supplementary Table S1.

92 Complete genome sequences were deposited in Genbank under accession numbers
93 MK303619 to MK303625, MK327281

94 **Results**

95 - Patients and viral characteristics

96 HCoV-OC43 isolates were detected by RT-PCR from nasal swabs collected from 8
97 HSCT recipients. Table 1 shows the characteristics of these patients. The median
98 time to HCoV-OC43 infection after HSCT was 239.5 days (range 13 – 1052 days). All
99 patients had graft-versus-host diseases for which they received a steroid daily dose
100 of higher than 1 mg/kg, except patient MDS4. All patients died except one, MDS4;
101 however, in all the cases, death was not directly related to HCoV-OC43 infection.
102 Figure 1 presents the timeline of events for patients involved in HCoV-OC43 clusters
103 of infections, that why patients MDS15 and MDS16 are not mentioned. Patients were
104 frequently admitted in hospital for consultations in the outpatient clinic or inpatient
105 hospitalizations. They shared the same hospital units and the same healthcare
106 worker teams. At the time of nasal swab collection, acute respiratory symptoms were
107 present for all the patients. Other respiratory pathogens were detected in nasal
108 swabs for only 2 patients (table 1). No concomitant co-infection with bacteria or fungi
109 was detected. All episodes occurred during winter.

110 - Characteristics of whole genome sequences

111 To characterize the overall diversity of HCoV-OC43 circulating among the 8 HSCT
112 patients, 7 full-length genome sequences were obtained from nasal swabs of HSCT
113 patients using NGS method. The sequence obtained from MDS15 patient was
114 incomplete since a part between positions 7473 and 7689 in orf1ab was missing.
115 Considering the small fragment of sequence missing, the partial full-length sequence
116 obtained from this patient was also included in all phylogenetic analyses. To
117 genotype the HCoV-OC43 isolates, phylogenetic trees were reconstructed by the
118 maximum-likelihood method using sequences of full-length genomes and S genes
119 obtained and compared to those retrieved from GenBank (Figure 2A/B). The tree
120 obtained from full-length genomes showed that MDS sequences are divided into 4

121 known genotypes B, F and G (Figure 2A) [25,26,28]. MDS4, MDS6, MDS14 and
122 MDS12 sequences, belonging to the genotype F, are grouped into the same cluster
123 separated from other sequences of genotype F with high bootstrap value (79.6%).
124 These sequences were obtained from samples collected in 2013 except for the
125 MDS12 sequence that was obtained from sample collected in 2014. This cluster was
126 named “cluster 2013” (Figure 2A). Among the cluster 2013, MDS6 and MDS14
127 sequences are closely related (69.2% bootstrap value), MDS12 sequence is
128 separated from the MDS 6 and MDS14 sequences (100% bootstrap value) and
129 MDS4 sequence is separated from the 3 other sequences (98.5% bootstrap value).
130 On the tree, MDS2 and MDS11 sequences are divided within the sequences of
131 genotype G with too low bootstrap values to separate firmly the sequences of this
132 cluster (Figure 2A). MDS2 and MDS11 sequences were collected during the year
133 2014 and will be named cluster 2014. Overall, the tree obtained from *S* gene
134 sequences shows the same distribution of MDS sequences. Notably, the cluster
135 2013 appears closely related with previous published sequences collected from
136 French patients in 2012 and 2013. In sum, the HCoV-OC43 sequences obtained
137 from clinical samples in this study belong to genotypes B, E, F and G, and 6 of them
138 appear clustered into two distinct clusters.

139 The sequences of the cluster 2013 are very conserved at the nucleotide level of the
140 full-length genomes (up to 99%). As observed on the tree, MDS 6 and 14 sequences
141 are the closest (with 99.93% of identity), MDS12 sequence is less similar to the
142 MDS6 and MDS14 sequences (99.11% of identity) and the MDS4 sequence is the
143 most divergent from the other of the cluster 2013 (99% of identity). However, MDS4
144 sequence presents a high level of homology with MDS6 and MD14 sequences
145 (99.81%). Likewise, MDS2 and MDS11 sequences are very similar with 99.86% of

146 identities (Table 2). These data reinforce the results of the phylogenetic trees and
 147 suggest that sequences of each cluster share the same origin.

% identity	MDS2	MDS4	MDS6	MDS1 1	MDS12	MDS14	MDS15	MDS16
MSD2		99.68	99.66	99.86	98.85	99.65	96.14	99.14
MSD4			99.81	99.69	99	99.81	96.22	99.2
MSD6				99.67	99.11	99.93	96.19	99.19
MSD11					98.86	99.66	96.16	99.15
MSD12						99.11	95.4	98.4
MSD14							96.18	99.19
MDS15								96.64
MSD16								

148 Table 2. Percent amino acid Identity of the 8 full-length genome sequences of MDS
 149 patients.

150 To reinforce the hypothesis of common origin of clinical isolates, the divergence
 151 times of MDS sequences were estimated using BEAST method. The estimated mean
 152 evolutionary rate was 4.0×10^{-4} ($3.4 \times 10^{-4} - 4.5 \times 10^{-4}$) substitutions/site/year for the
 153 full-length genome and, for the *S* gene, 5.5×10^{-4} ($4.5 \times 10^{-4} - 6.5 \times 10^{-4}$)
 154 substitutions/site/year. Based on the full-length genome and the *S* gene data, the
 155 time of emergence for genotype A was estimated in the 1960s and in the late 1990s
 156 to early 2000s for genotypes B to D (table 3). Similarly, emergence of genotypes D to
 157 G was estimated in the late 2000s to early 2010s. Based on the evolutionary
 158 estimates of the full-length genome and the *S* gene, the common ancestor of the
 159 sequences of the cluster 2013 was dated back to 2010.9 (2009.6 – 2012.2) and to
 160 2010.9 (2009.8 – 2011.7), respectively. Furthermore, the common ancestor of the
 161 sequences of the cluster 2014 was dated back to 2011.5 (2010.7 – 2012.5) and to
 162 2011.0 (2010.1 – 2011.7) for the full-length genome and the *S* gene, respectively.
 163 These estimates obtained from two different databases were consistent and show
 164 that the strains of the clusters 2013 emerged from a common ancestor in 2010.9 and
 165 those of 2014 emerged from another in 2011 +/-0.5.

tMRCA (95% HPD)		
Genotype	S gene	Full-length genome
A	1963.1 (1956.6 – 1966.7)	1967.0 (1966.9 – 1967.1)
B	1999.3 (1995.1 – 2002.3)	2001.6 (2000.3 – 2002.7)
C	1993.2 (1988.7 – 1996.5)	2002.9 (2001.8 – 2003.7)
D	1998.1 (1994.8 – 2001.2)	2003.0 (2002.0 – 2003.7)
E	2006.5 (2005.7 – 2007.5)	2007.2 (2005.9 – 2008.4)
F	2007.7 (2007.0 – 2008.3)	2009.4 (2008.0 – 2010.5)
G	2010.8 (2010.0 – 2011.4)	2010.9 (2010.3 – 2011.5)
Cluster 2013	2010.9 (2009.8 – 2011.7)	2010.9 (2009.6 – 2012.2)
Cluster 2014	2011.0 (2010.1 – 2011.7)	2011.5 (2010.7 – 2012.5)
Mean evolutionary rate (nt subst/site/year)	5.5×10^{-4} (4.5×10^{-4} – 6.5×10^{-4})	4.0×10^{-4} (3.4×10^{-4} – 4.5×10^{-4})

166 Table 3. Time of the most recent common ancestors (tMRCA) with 95 % highest
167 posterior density (95 % HPD) for HCoV-OC43 genotypes A to G and clusters based
168 on the spike (S) gene and full-length genome.

169

170 Using the prototype ATCC VR759 as the reference strain for amino acid positions,
171 sequence MDS16 shares 19 amino acid substitutions with other sequences of
172 genotype B and presents 18 specific amino acid substitutions mapped across the
173 whole genome relative to the other sequences of the genotype B (Figure 3A).
174 Likewise, sequences of the cluster 2013 show the specific substitution of the
175 genotype F (Y177H) and 9 additional amino acids substitutions relative to the others
176 of the genotype F [25]. Notably, MDS6, MDS12 and MDS14 sequences show 12
177 additional substitutions relative to the genotype F, which are absent on the MDS4
178 sequence. MDS15 sequence, for which only the S region could be analyzed, shares
179 the same amino acid substitutions with the other sequences of the genotype E and
180 shows only 1 specific substitution relative to the other sequences of the same
181 genotype. MDS2 and MDS11 sequences appear similar to the others of genotype G.

182 - Characteristics of HCoV-OC43 strains

183 Since S protein sequences of other French HCoV-OC43 isolates have been
184 previously published, phylogenetic analyses show that MDS sequences of the cluster

185 2013 are clustered to other French sequences on the phylogenetic tree, whereas
186 sequences of the cluster 2014 appear more related to Malaysian, Chinese and
187 American sequences (Figure 2B) [29]. When signature amino acid substitutions were
188 analyzed, for the genotype B, 3 amino acid substitutions appear on the sequence
189 MDS16: absence of substitution D471N, substitutions L482F that is common to other
190 sequences of genotype B, C, F and G and L505S that is also observed on
191 sequences of genotype C. For the genotype E, the amino acid substitution N767I is
192 observed on the sequence MDS15 and had never been described previously. For the
193 genotype F, the amino acid substitution P38S is observed on the MDS4, MDS6,
194 MDS12 and MDS14 sequences which are shared with genotype B and G. The
195 substitutions D1264H and C1319S are absent from the MDS4, MDS6, MDS12 and
196 MDS14 sequences as from the other French sequences K963240, K963243 and
197 K963244. Finally, 2 amino acid substitutions (A1054S and D1252G) are observed on
198 MDS6, MDS12 and MDS14 sequences and also on previously published sequences
199 K963243 and K963244 (Figure 3B).

200 These data show that any amino acid substitution appear specific to sequences of
201 genotype B and G circulating in France. But, sequences circulating among French
202 people belonging to the genotype F had a peculiar molecular signature (absence of
203 substitution D1264H and C1319S, substitution A1054S and D1252G).

204 **Discussion**

205 In this work, we describe retrospectively 8 HCoV-OC43 infections having occurred
206 during the winters 2013 and 2014 at the HSCT Unit of the Lille University hospital.
207 Six patients were divided into 2 clusters, respectively named 2013 and 2014 on the
208 basis of the year of occurrence and genomic characteristics.

209 HCoV-OC43 cases described in this report could be classified as healthcare-
210 associated infections defined as a group of infections emerging among patients who
211 come from a community with a history of previous exposure to healthcare, but do not
212 fit the nosocomial infection criteria [30]. The 6 patients divided into the two clusters of
213 cases regularly attended the hospital for in- and outpatient cares, as consultations,
214 day and inpatient hospitalizations. So, they were frequently in contact with health
215 care professionals and hospital visitors that may be considered as multiple sources
216 of HCoV infections, during periods of usual circulation of HCoV-OC43 in France [31].
217 Outside the hospital, they resided in geographically non-related places and no
218 common activity could be identified. So, the hospital admission seems the only link
219 between the patients, even though we were unable to identify any direct contact
220 between the patients. HCoV-OC43 can be transmitted either by an asymptomatic
221 patient with prolonged shedding, by healthcare worker asymptomatic or with an
222 upper respiratory tract infection [10,32]. Phylogenetic analyses showed that whole-
223 genome sequences included in a same cluster were very closely related, had very
224 high levels of similarities and shared same common ancestors. All these elements
225 provide additional evidence that these infections could be classified as healthcare-
226 associated infections. Similarly, to what is observed for influenza viruses,
227 parainfluenza virus type 3 and respiratory syncytial virus, our data show that HCoV-
228 OC43 is involved in healthcare-associated infections [33–35].
229 Because we were unable to identify the index case of HCoV-OC43 infections of the
230 cluster 2013, the MDS12 patient infection during the winter 2014 with a strain highly
231 related to those circulating the previous year is surprising. However, previous studies
232 have shown that genomes of other high and low pathogenic coronaviruses presented

233 low variation rates during hospital-acquired outbreaks [36,37]. Our data show that
234 sequences of HCoV-OC43 appear also very stable.

235 Phylogenetic trees based on the whole genome and *S* gene HCoV-OC43 show that
236 genotypes B, G, E and F currently circulate among French patients, in agreement
237 with other report [29]. Even if HCoV-OC43 was previously detected among patients
238 with haematology malignancy, the involvement of HCoV-OC43 genotypes in this type
239 of infection was rarely investigated [38–40]. Indeed, the analysis of more HCoV-
240 OC43 strains from other healthcare-associated infections will reveal the relative
241 prevalence of each genotype in this type of infections in different localities.

242 The divergence time of sequences included in the phylogenetic analyses was
243 calculated to accumulate to collect additional information showing the common origin
244 of clinical isolates. Our results are similar to those previously described, validating
245 the tMRCA method used [25,28]. Indeed, our results show that each cluster of
246 sequences has diverged from a distinct ancestor in 2010.9 for the cluster 2013 and in
247 2011.5 for the cluster 2014.

248 Due to limited HCoV-OC43 full-length genome published in databases, phylogenetic
249 analyses were conducted on *S* gene sequences. All results obtained from whole
250 genome and *S* gene sequences were similar. However, phylogenetic analyses from
251 the numerous published *S* gene sequences allowed comparison of HCoV-OC43
252 sequences isolated in Lille with those isolated in different parts of the world. Strains
253 described here represent probably a part of strains currently circulating among
254 European people. However, our data show that both kinds of HCoV-OC43 strains
255 circulate among patients included in the study: those belonging to the genotype G
256 closely related to Asian strains and other belonging to the genotype F more related to
257 other European strains previously described [25,26,28,29,41]. Amino acid

258 substitution analyses showed that some substitutions especially in the S protein
259 could be specific of strains circulating in European population but the number of
260 strains analyzed here and the number of sequences obtained from European isolates
261 of HCoV-OC43 are too small and further studies will be needed to confirm our data.
262 We are aware that our study presents some limitations. As previously mentioned, we
263 were unable to identify the index case of the HCoV-OC43 infections in the HSCT unit
264 and determine precisely the contact history of the patients. Thus, we cannot exclude
265 that other patients were involved in the HCoV-OC43 healthcare-associated infections
266 described. In the same way, we were unable to detect infection or asymptomatic
267 portage in healthcare workers of the HSCT unit. Some reports have shown that
268 different HCoV-OC43 strains circulated during a same winter, suggesting that there
269 may be multiple channels of HCoV introduction in hospital [25,42,43]. In fact, the
270 HCoV strains of healthcare-associated infections were probably contracted from one
271 of the strains circulating in the community at that time. For all these reasons, cases
272 reported in this study did not correspond to the definition of nosocomial transmission
273 of pathogens and we used the term of coronavirus healthcare-associated infections.
274 Another weakness of our work concerns the determination of pattern of amino acid
275 substitutions specific of strains circulating among French people. Indeed, our data
276 suggest that some substitutions might be specific to strains circulating in Europe but
277 the analysis of a larger number of strains would be necessary to confirm our
278 hypothesis.

279 To conclude, we retrospectively described the molecular investigation of HCoV-OC43
280 infections in a HSCT unit using whole-genome sequencing combined with advanced
281 phylogenetic analyses. These tools allowed us to associate HCoV-OC43 infections
282 with healthcare. Moreover, we suggest that two kinds of HCoV-OC43 strains circulate

283 among the French population, one sharing common ancestors with Asian strains and
284 the other closely related to European lineage.

285 **Acknowledgements.** The authors thank J. Ogiez for the management of the
286 biological collection.

287 **Funding sources.** This work was supported by a Lille Hospital University grant (EPI-
288 CoV) and by CNRS and Lille University (PEPS 2015).

289

290 Table 1. Clinical features of patients with HCoV-OC43 infections.

291 Figure 1. Timeline of events. NS+: Nasal swab positive for HCoV-OC43

292 Figure 2. Maximum-likelihood trees of HCoV-OC43 strains. A, B, C, D, E, F, G
293 genotypes/clades were represented. MDS sequences are in bold. A. Tree obtained
294 from whole-genome sequences. B. Tree obtained from S gene sequences.

295 Figure 3. Signature amino acid substitution differences. A. Amino acid substitution
296 differences across the full-length genome. B. Amino acid substitution differences
297 among the S gene sequences.

298 **References**

- 299 [1] A. Gagneur, J. Sizun, S. Vallet, M.C. Legr, B. Picard, P.J. Talbot, Coronavirus-
300 related nosocomial viral respiratory infections in a neonatal and paediatric intensive
301 care unit: a prospective study, *J. Hosp. Infect.* 51 (2002) 59–64.
302 doi:10.1053/jhin.2002.1179.
- 303 [2] N. Lee, D. Hui, A. Wu, P. Chan, P. Cameron, G.M. Joynt, A. Ahuja, M.Y.
304 Yung, C.B. Leung, K.F. To, S.F. Lui, C.C. Szeto, S. Chung, J.J.Y. Sung, A Major
305 Outbreak of Severe Acute Respiratory Syndrome in Hong Kong, *N. Engl. J. Med.* 348
306 (2003) 1986–1994. doi:10.1056/NEJMoa030685.
- 307 [3] A. Gagneur, S. Vallet, P.J. Talbot, M.-C. Legrand-Quillien, B. Picard, C.
308 Payan, J. Sizun, Outbreaks of human coronavirus in a pediatric and neonatal
309 intensive care unit, *Eur. J. Pediatr.* 167 (2008) 1427–1434. doi:10.1007/s00431-008-
310 0687-0.
- 311 [4] F.S. Alhamlan, M.S. Majumder, J.S. Brownstein, J. Hawkins, H.M. Al-Abdely,
312 A. Alzahrani, D.A. Obaid, M.N. Al-Ahdal, A. BinSaeed, Case characteristics among
313 Middle East respiratory syndrome coronavirus outbreak and non-outbreak cases in
314 Saudi Arabia from 2012 to 2015, *BMJ Open.* 7 (2017) e011865.
315 doi:10.1136/bmjopen-2016-011865.
- 316 [5] M.S. Majumder, J.S. Brownstein, S.N. Finkelstein, R.C. Larson, L. Bourouiba,
317 Nosocomial amplification of MERS-coronavirus in South Korea, 2015, *Trans. R. Soc.*
318 *Trop. Med. Hyg.* 111 (2017) 261–269. doi:10.1093/trstmh/trx046.
- 319 [6] J. Garbino, S. Crespo, J.-D. Aubert, T. Rochat, B. Ninet, C. Deffernez, W.
320 Wunderli, J.-C. Pache, P.M. Socal, L. Kaiser, A prospective hospital-based study of
321 the clinical impact of non-severe acute respiratory syndrome (Non-SARS)-related
322 human coronavirus infection, *Clin. Infect. Dis.* 43 (2006) 1009–1015.

323 doi:10.1086/507898.

324 [7] M. Hakki, R.M. Rattray, R.D. Press, The clinical impact of coronavirus infection
325 in patients with hematologic malignancies and hematopoietic stem cell transplant
326 recipients, *J. Clin. Virol.* 68 (2015) 1–5. doi:10.1016/j.jcv.2015.04.012.

327 [8] J.L. Piñana, S. Madrid, A. Pérez, J.C. Hernández-Boluda, E. Giménez, M.J.
328 Terol, M. Calabuig, D. Navarro, C. Solano, Epidemiologic and Clinical Characteristics
329 of Coronavirus and Bocavirus Respiratory Infections after Allogeneic Stem Cell
330 Transplantation: A Prospective Single-Center Study, *Biol. Blood Marrow Transplant.*
331 *J. Am. Soc. Blood Marrow Transplant.* 24 (2018) 563–570.
332 doi:10.1016/j.bbmt.2017.11.001.

333 [9] F. Milano, A.P. Campbell, K.A. Guthrie, J. Kuypers, J.A. Englund, L. Corey, M.
334 Boeckh, Human rhinovirus and coronavirus detection among allogeneic
335 hematopoietic stem cell transplantation recipients, *Blood.* 115 (2010) 2088–2094.
336 doi:10.1182/blood-2009-09-244152.

337 [10] C. Ogimi, A.L. Greninger, A.A. Waghmare, J.M. Kuypers, R.C. Shean, H. Xie,
338 W.M. Leisenring, T.L. Stevens-Ayers, K.R. Jerome, J.A. Englund, M. Boeckh,
339 Prolonged Shedding of Human Coronavirus in Hematopoietic Cell Transplant
340 Recipients: Risk Factors and Viral Genome Evolution, *J. Infect. Dis.* 216 (2017) 203–
341 209. doi:10.1093/infdis/jix264.

342 [11] G. Gerna, G. Campanini, F. Rovida, E. Percivalle, A. Sarasini, A. Marchi, F.
343 Baldanti, Genetic variability of human coronavirus OC43-, 229E-, and NL63-like
344 strains and their association with lower respiratory tract infections of hospitalized
345 infants and immunocompromised patients, *J. Med. Virol.* 78 (2006) 938–949.
346 doi:10.1002/jmv.20645.

347 [12] I.M. Mackay, K.E. Arden, D.J. Speicher, N.T. O’Neil, P.K. McErlean, R.M.

348 Greer, M.D. Nissen, T.P. Sloots, Co-circulation of four human coronaviruses (HCoVs)
349 in Queensland children with acute respiratory tract illnesses in 2004, *Viruses*. 4
350 (2012) 637–653. doi:10.3390/v4040637.

351 [13] C.S. Silva, L.B. Mullis, O. Pereira, L.J. Saif, A. Vlasova, X. Zhang, R.J. Owens,
352 D. Paulson, D. Taylor, L.M. Haynes, M.P. Azevedo, Human Respiratory
353 Coronaviruses Detected In Patients with Influenza-Like Illness in Arkansas, USA,
354 *Virol. Mycol. Infect. Dis.* 2014 (2014). doi:10.4172/2161-0517.S2-004.

355 [14] J. Hand, EB. Rose, A. Salinas, X. Lu, SK Sakthivel, E. Schneider, JT. Watson.
356 Severe Respiratory Illness Outbreak Associated with Human Coronavirus NL63 in a
357 Long-Term Care Facility - Volume 24, Number 10—October 2018 - Emerging
358 Infectious Diseases journal - CDC, (n.d.). doi:10.3201/eid2410.180862.

359 [15] F. Maurier, D. Beury, L. Fléchon, J.-S. Varré, H. Touzet, A. Goffard, D. Hot, S.
360 Caboche, A complete protocol for whole-genome sequencing of virus from clinical
361 samples: Application to coronavirus OC43, *Virology*. 531 (2019) 141–148.
362 doi:10.1016/j.virol.2019.03.006.

363 [16] B.T. Taboada, P. Isa, M.A. Espinoza, F.E. Aponte, M.A. Arias-Ortiz, J. Monge-
364 Martínez, R. Rodríguez-Vázquez, F. Díaz-Hernández, F. Zárate-Vidal, R.M. Wong-
365 Chew, V. Firo-Reyes, C.N. Del Río-Almendárez, J. Gaitán-Meza, A. Villaseñor-
366 Sierra, G. Martínez-Aguilar, M. García-Borjas, D.E. Noyola, L.F. Pérez-González, S.
367 López, J.I. Santos-Preciado, C.F. Arias, Complete Genome Sequence of Human
368 Coronavirus OC43 Isolated from Mexico, *Genome Announc.* 4 (2016).
369 doi:10.1128/genomeA.01256-16.

370 [17] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local
371 alignment search tool, *J. Mol. Biol.* 215 (1990) 403–410. doi:10.1016/S0022-
372 2836(05)80360-2.

- 373 [18] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and
374 high throughput, *Nucleic Acids Res.* 32 (2004) 1792–1797. doi:10.1093/nar/gkh340.
- 375 [19] M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform
376 graphical user interface for sequence alignment and phylogenetic tree building, *Mol.*
377 *Biol. Evol.* 27 (2010) 221–224. doi:10.1093/molbev/msp259.
- 378 [20] P. Stothard, The sequence manipulation suite: JavaScript programs for
379 analyzing and formatting protein and DNA sequences, *BioTechniques.* 28 (2000)
380 1102, 1104. doi:10.2144/00286ir01.
- 381 [21] S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel,
382 New algorithms and methods to estimate maximum-likelihood phylogenies:
383 assessing the performance of PhyML 3.0, *Syst. Biol.* 59 (2010) 307–321.
384 doi:10.1093/sysbio/syq010.
- 385 [22] V. Lefort, J.-E. Longueville, O. Gascuel, SMS: Smart Model Selection in
386 PhyML, *Mol. Biol. Evol.* 34 (2017) 2422–2424. doi:10.1093/molbev/msx149.
- 387 [23] FigTree, (n.d.). <http://tree.bio.ed.ac.uk/software/figtree/> (accessed January 28,
388 2019).
- 389 [24] W. Deng, B.S. Maust, D.C. Nickle, G.H. Learn, Y. Liu, L. Heath, S.L.
390 Kosakovsky Pond, J.I. Mullins, DIVEIN: a web server to analyze phylogenies,
391 sequence divergence, diversity, and informative sites, *BioTechniques.* 48 (2010)
392 405–408. doi:10.2144/000113370.
- 393 [25] X.Y. Oong, K.T. Ng, Y. Takebe, L.J. Ng, K.G. Chan, J.B. Chook, A.
394 Kamarulzaman, K.K. Tee, Identification and evolutionary dynamics of two novel
395 human coronavirus OC43 genotypes associated with acute respiratory infections:
396 phylogenetic, spatiotemporal and transmission network analyses, *Emerg. Microbes*
397 *Infect.* 6 (2017) e3. doi:10.1038/emi.2016.132.

398 [26] Y. Zhang, J. Li, Y. Xiao, J. Zhang, Y. Wang, L. Chen, G. Paranhos-Baccalà, L.
399 Ren, J. Wang, Genotype shift in human coronavirus OC43 and emergence of a novel
400 genotype by natural recombination, *J. Infect.* 70 (2015) 641–650.
401 doi:10.1016/j.jinf.2014.12.005.

402 [27] A.J. Drummond, M.A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics
403 with BEAUti and the BEAST 1.7, *Mol. Biol. Evol.* 29 (2012) 1969–1973.
404 doi:10.1093/molbev/mss075.

405 [28] S.K.P. Lau, P. Lee, A.K.L. Tsang, C.C.Y. Yip, H. Tse, R.A. Lee, L.-Y. So, Y.-L.
406 Lau, K.-H. Chan, P.C.Y. Woo, K.-Y. Yuen, Molecular epidemiology of human
407 coronavirus OC43 reveals evolution of different genotypes over time and recent
408 emergence of a novel genotype due to natural recombination, *J. Virol.* 85 (2011)
409 11325–11337. doi:10.1128/JVI.05512-11.

410 [29] N. Kin, F. Mischczak, W. Lin, M.A. Gouilh, A. Vabret, E. Consortium, Genomic
411 Analysis of 15 Human Coronaviruses OC43 (HCoV-OC43s) Circulating in France
412 from 2001 to 2013 Reveals a High Intra-Specific Diversity with New Recombinant
413 Genotypes, *Viruses.* 7 (2015) 2358–2377. doi:10.3390/v7052358.

414 [30] T. Cardoso, M. Almeida, N.D. Friedman, I. Aragão, A. Costa-Pereira, A.E.
415 Sarmiento, L. Azevedo, Classification of healthcare-associated infection: a systematic
416 review 10 years after the first proposal, *BMC Med.* 12 (2014) 40. doi:10.1186/1741-
417 7015-12-40.

418 [31] A. Hassoun, M.D. Huff, D. Weisman, K. Chahal, E. Asis, D. Stalons, E.
419 Grigorenko, J. Green, L.L. Malone, S. Clemmons, S. Lu, Seasonal variation of
420 respiratory pathogen colonization in asymptomatic health care professionals: A
421 single-center, cross-sectional, 2-season observational study, *Am. J. Infect. Control.*
422 43 (2015) 865–870. doi:10.1016/j.ajic.2015.04.195.

423 [32] J.C. Esbenshade, K.M. Edwards, A.J. Esbenshade, V.E. Rodriguez, H.K.
424 Talbot, M.F. Joseph, S.K. Nwosu, J.D. Chappell, J.E. Gern, J.V. Williams, T.R.
425 Talbot, Respiratory virus shedding in a cohort of on-duty healthcare workers
426 undergoing prospective surveillance, *Infect. Control Hosp. Epidemiol.* 34 (2013) 373–
427 378. doi:10.1086/669857.

428 [33] J.H. Choi, E.H. Choi, H.J. Kang, K.D. Park, S.S. Park, H.Y. Shin, H.J. Lee,
429 H.S. Ahn, Respiratory viral infections after hematopoietic stem cell transplantation in
430 children, *J. Korean Med. Sci.* 28 (2013) 36–41. doi:10.3346/jkms.2013.28.1.36.

431 [34] K.M. Weedon, A.H. Rupp, A.C. Heffron, S.F. Kelly, X. Zheng, S.T. Shulman,
432 P. Gutman, D. Wang, Y. Zhou, G.A. Noskin, E.J. Anderson, The impact of infection
433 control upon hospital-acquired influenza and respiratory syncytial virus, *Scand. J.*
434 *Infect. Dis.* 45 (2013) 297–303. doi:10.3109/00365548.2012.726738.

435 [35] P. Loubet, G. Voiriot, N. Houhou-Fidouh, M. Neuville, L. Bouadma, F.-X.
436 Lescure, D. Descamps, J.-F. Timsit, Y. Yazdanpanah, B. Visseaux, Impact of
437 respiratory viruses in hospital-acquired pneumonia in the intensive care unit: A
438 single-center retrospective study, *J. Clin. Virol.* 91 (2017) 52–57.
439 doi:10.1016/j.jcv.2017.04.001.

440 [36] J.P. Hays, S.H. Myint, PCR sequencing of the spike genes of geographically
441 and chronologically distinct human coronaviruses 229E, *J. Virol. Methods.* 75 (1998)
442 179–193. doi:10.1016/S0166-0934(98)00116-5.

443 [37] S. Tong, J.R. Lingappa, Q. Chen, B. Shu, A.C. LaMonte, B.T. Cook, C. Birge,
444 S. Wang Chern, X. Liu, R. Galloway, M. Le Quynh, F.N. Wai, J.-Y. Yang, J. Butany,
445 J.A. Comer, S.S. Monroe, S.R. Beard, T.G. Ksiazek, D. Erdman, P.A. Rota, M.A.
446 Pallansch, L.J. Anderson, Direct Sequencing of SARS-Coronavirus S and N Genes
447 from Clinical Specimens Shows Limited Variation, *J. Infect. Dis.* 190 (2004) 1127–

448 1131. doi:10.1086/422849.

449 [38] M. Mikulska, V. Del Bono, N. Gandolfo, S. Dini, A. Dominietto, C. Di Grazia, S.
450 Bregante, R. Varaldo, A. Orsi, F. Ansaldi, A. Bacigalupo, C. Viscoli, Epidemiology of
451 viral respiratory tract infections in an outpatient haematology facility, *Ann. Hematol.*
452 93 (2014) 669–676. doi:10.1007/s00277-013-1912-0.

453 [39] S. Hutspardol, M. Essa, S. Richardson, T. Schechter, M. Ali, J. Krueger, H.
454 Fujii, R.M. Egeler, A. Gassas, Significant Transplantation-Related Mortality from
455 Respiratory Virus Infections within the First One Hundred Days in Children after
456 Hematopoietic Stem Cell Transplantation, *Biol. Blood Marrow Transplant.* 21 (2015)
457 1802–1807. doi:10.1016/j.bbmt.2015.06.015.

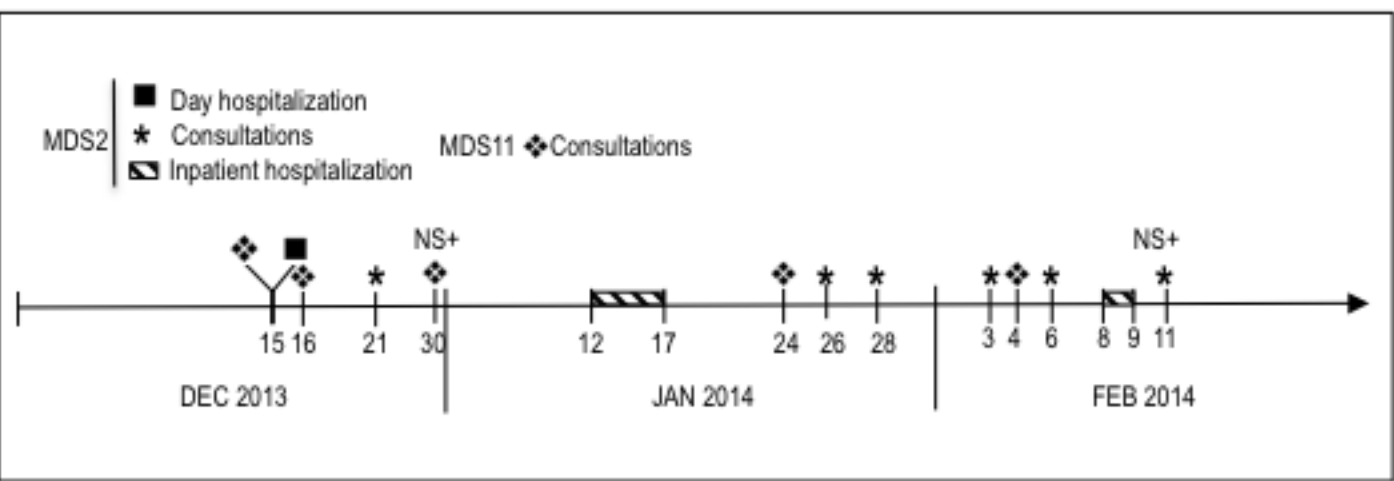
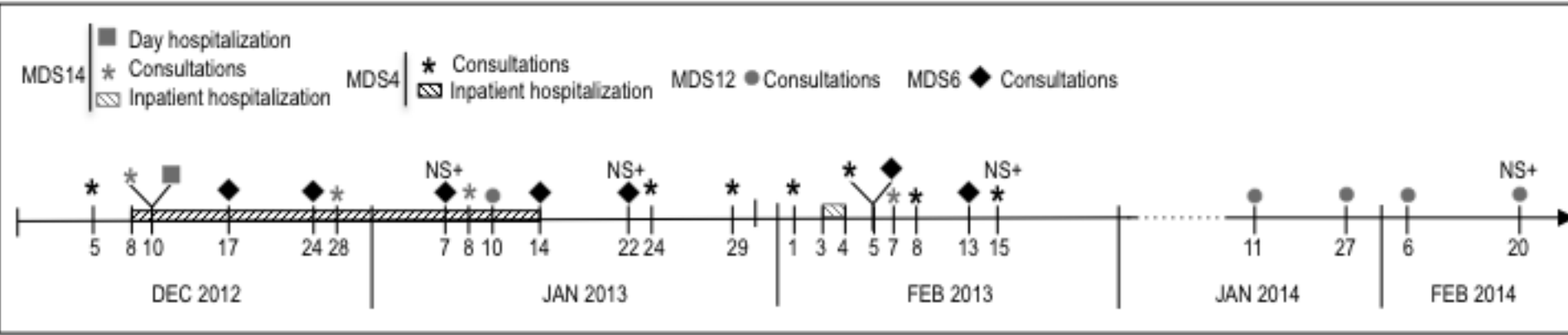
458 [40] C. Ogimi, A.A. Waghmare, J.M. Kuypers, H. Xie, C.C. Yeung, W.M.
459 Leisenring, S. Seo, S.-M. Choi, K.R. Jerome, J.A. Englund, M. Boeckh, Clinical
460 Significance of Human Coronavirus in Bronchoalveolar Lavage Samples From
461 Hematopoietic Cell Transplant Recipients and Patients With Hematologic
462 Malignancies, *Clin. Infect. Dis.* 64 (2017) 1532–1539. doi:10.1093/cid/cix160.

463 [41] N. KIN, F. Mischczak, L. Diancourt, V. Caro, F. Moutou, A. Vabret, M. Ar
464 Gouilh, Comparative molecular epidemiology of two closely related coronaviruses,
465 bovine coronavirus (BCoV) and human coronavirus OC43 (HCoV-OC43), reveals a
466 different evolutionary pattern, *Infect. Genet. Evol.* 40 (2016) 186–191.
467 doi:10.1016/j.meegid.2016.03.006.

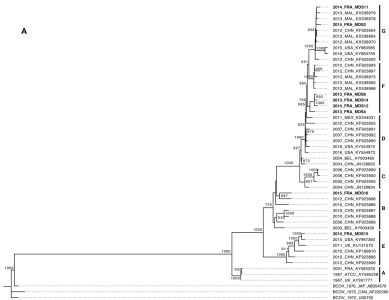
468 [42] Y. Zhu, C. Li, L. Chen, B. Xu, Y. Zhou, L. Cao, Y. Shang, Z. Fu, A. Chen, L.
469 Deng, Y. Bao, Y. Sun, L. Ning, C. Liu, J. Yin, Z. Xie, K. Shen, A novel human
470 coronavirus OC43 genotype detected in mainland China, *Emerg. Microbes Infect.* 7
471 (2018). doi:10.1038/s41426-018-0171-5.

472 [43] S.-F. Zhang, J.-L. Tuo, X.-B. Huang, X. Zhu, D.-M. Zhang, K. Zhou, L. Yuan,

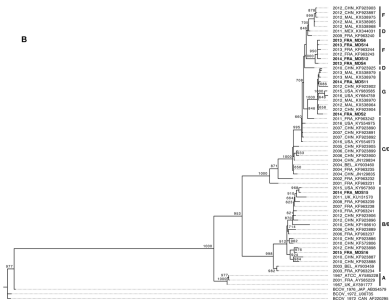
473 H.-J. Luo, B.-J. Zheng, K.-Y. Yuen, M.-F. Li, K.-Y. Cao, L. Xu, *Epidemiology*
474 characteristics of human coronaviruses in patients with respiratory infection
475 symptoms and phylogenetic analysis of HCoV-OC43 during 2010-2015 in
476 Guangzhou, *PloS One*. 13 (2018) e0191789. doi:10.1371/journal.pone.0191789.

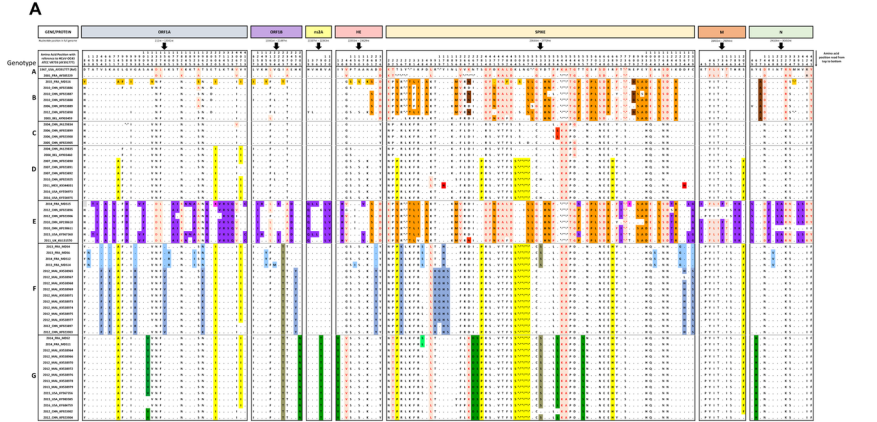


A



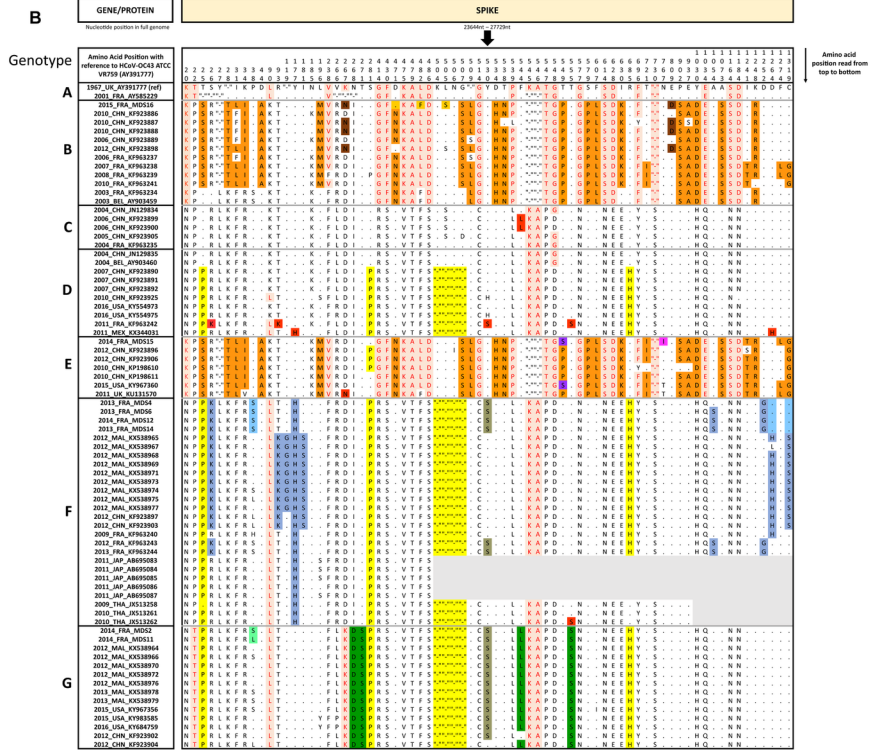
B





Signature amino acid substitution differences across the whole genome.

Signatures were identified by using Signatures software. Amino acid positions were aligned with Clustal2 in the BioEdit software. Informative sites were obtained by using DMSV website (<http://dmsv.msu.edu/ncvtool/ncvtool.cgi?seqtype=1>).



Signature amino acid substitution differences across the Spike gene (S gene).

01/10/2021

code	Nom, Prénom	N° MOLIS	DDN	UF	Diagnostic	Date de plt	Ech	Pi	RNA Haoran	RNA CHRU
MDS-1	ROGIE, Cedric	13074C5163	04.09.1986	1408	OC43	14/03/2013	Ec nasal		0 10 µl	50 µl
MDS-2	EL MORABITI,SAFIYA	14111C4470	14.03.1995	1408	OC43	11/03/2014	Ec Nasal		0	50 µl
MDS-4	HERNAULT,CHRISTOPHE	13075C4299	04.02.1983	1408	OC43	15/02/2013	Ec Nasal	CHR	10 µl	
MDS-5	BEHONH TOLLY,GEORGES	13123C4707	15.03.1988	1408	OC43	20/03/2013	Ec Nasal	CHR	10 µl	
MDS-6	DELANNOY, Micheline, ep. Monel	13042C4722	26.10.1949	1408	OC43	22/01/2013	Ec Nasal	CHR	10 µl	
MDS-11	DIAFI,NOURIA	14054C3961	10.03.1969	1408	OC43+++	30/01/2014	Ec Nasal		0 10 µl	
MDS-12	KINGET,JOSETTE ep FAUQUEMBERGUE	14084C4874	22.06.1949	1408	Grippe A++;OC43++	20/02/2014	Ec Nasal		0 10 µl	
MDS-14	SECQ,EMILIE ep LAIGNEL	13022C3862	28.12.1980	1408	Grippe A++;OC43++;BocaV	08/01/2013	Ec Nasal	IBL	10 µl	50 µl
MDS-15	FONTAINE,SYLVIA	14072C4436	05.04.1974	1408	Grippe A++;OC43++;VRS A	11/02/2014	Ec Nasal		0 5 µl	
MDS-16	BOUDARA, KAIS	15035C4408	15.04.1982	1408	OC43++	16/01/2015	Ec Nasal	IBL	10 µl	50 µl
PR2	REYMBAUT,MAGALIE	14095c3755	11.07.1980	5328	OC43	28/02/2014	LBA		0	
	REYMBAUT,MAGALIE	14111C6659		5344	OC43	10/03/2014	LBA	IBL	10 µl	50 µl

code	Nom, Prénom	N° MOLIS	DDN	âge	UF	Diagnostic	Date de plt	Ech	date hospitalisation	Cause hospitalisation	Microbiologie	Hb (g/dL)	Plaquettes (x10exp9/L)	GB (x10exp9/L)	PNN (x10exp9/L)	Lymphocytes (x10exp9/L)	Monocytes (x10exp9/L)	Diagnostic initial	Greffe	Conditionnement	date greffe	ATCDS	Traitement	
MDS-2	EL MORABITI,Safya	14111C4470	14.03.1995	19	1408	OC43	11/03/2014	Ec Nasal	Mar-14		10/03/14 : CMV (pp65) : 12 cellules +, 03/14 : sérologie fongique : aspergillase neg, serologie candida = neg, pas bactériologie demandée	11	196	6.45	3.6	1.3	1.4	LAM 1, trisomie 8	Alogreffe de MO		2014		methotrexate, ciclosporine	
MDS-4	HERNAULT,Christophe	13075C4299	04.02.1983	30	1408	OC43	15/02/2013	Ec Nasal	Feb-13		EBV = CV moyenne, Hémoc stériles, Ag CMV neg, serologies fongiques neg, 19/01/13 = ECBU = E. coli	11.3	71	6.35	4.5	1	0.7	Myelome multiple IgA kappa	Alogreffe de MO	Endoxan, Busulfan, SAL	20/12/2012		ciclosporine, posaconazole	
MDS-6	DELANNOY, Micheline, ep. Monel	13042C4722	26.10.1949	64	1408	OC43	22/01/2013	Ec Nasal	Jan-13		22/01 : HHV6 + faible	8.7	14	1.59	1	0.4	0.1		Alogreffe de MO		01/07/2012			
MDS-11	DIAFINOURIA	14054C3961	10.03.1969	45	1408	OC43+++	30/01/2014	Ec Nasal	Jan-14		02/01/14, 30/01/14 : charge virale faible à EBV, 30/01/14 : ECBC = flore bucco-pharyngée,	8.6	21	1.49	0.3	1.1	0.1	Sd myeloprolifératif atypique, alogreffe			09/08/2013		ciclosporine	
MDS-12	KINGET, Josette ep FAUQUEMBERGUE	14084C4874	22.06.1949	65	1408	Grippe A+-OC43++	20/02/2014	Ec Nasal	Feb-14		23/01 : ec nasal Rhinovirus +, 20/02 : sero fongiques négatives	10.1	73	1.26	0.7	0.4	0.1	Rechute de myélome multiple IgG lambda	Alogreffe de MO	protocole de Nantes	26/10/2012		Pomalidomide, Endoxan, dexamethasone/ radiothérapie externe dans les jours précédents	
MDS-14	SECO,EMILIE ep LAIGNEL	13022C3862	28.12.1980	33	1408	Grippe A+-OC43+-BocaV	08/01/2013	Ec Nasal	Jan-13			13.7	102	6.92	4.8	1.3	0.6	Rechute LAL	Alogreffe de MO	Vincristine, Methotrexate, Purinethol, Dexamethasone	02/03/2012	09/2012 : pneumopathie, 10/12 : pneumocystose pulm		
MDS-16	BOUDARA, Kais	15035C4408	15.04.1982	33	1408	OC43++	16/01/2015	Ec Nasal	Jan-15		RAS	14.6	256	14.05	11.3	2.1	0.7	Rechute de mycosis fongicide	Alogreffe de MO familiale	Thiotépa, Fludarabine, Busilex, SAL	18/11/2012		neoral, corticoïdes	

code	DDN	Sexe	date de prélèvement	type d'échantillon	Lieu
MDS-2	14.03.1995	F	11/03/2014	Ecouvillon Nasal	Lille
MDS-4	04.02.1983	M	15/02/2013	Ecouvillon Nasal	Lille
MDS-6	26.10.1949	F	22/01/2013	Ecouvillon Nasal	Lille
MDS-11	10.03.1969	F	30/01/2014	Ecouvillon Nasal	Lille
MDS-12	22.06.1949	F	20/02/2014	Ecouvillon Nasal	Lille
MDS-14	28.12.1980	F	08/01/2013	Ecouvillon Nasal	Lille
MDS-16	15.04.1982	M	16/01/2015	Ecouvillon Nasal	Lille

	Age	Sex	Hematology feature	Transplant numbers	Donor type	Days between transplant and HCoV-infection	date of HCoV detection	Cq value	Respiratory or infectious symptoms	Platelet count	Leukocyte count	Neutrophil count	Lymphocyte count	Monocyte count	steroids	GVH	Co-detection	Outcome
	y					m/d/y	m/d/y			x10exp9/L	x10exp9/L	x10exp9/L	x10exp9/L	x10exp9/L				
MDS-2	19	F	acute myeloid leukemia	1	unrelated	2014	3/11/14	35	yes	196	6.45	3.6	1.3	1.4	y	y	no	death
MDS-4	30	M	Multiple myeloma	2	auto/unrelated	13	2/15/13	34	yes	71	6.35	4.5	1	0.7	n	y	no	alive
MDS-6	64	F	myeloproliferative syndrome	1	unrelated	359	1/22/13	37	yes	14	1.59	1	0.4	0.1	y	y	no	death
MDS-11	45	F	atypical myeloproliferative syndrome	1	unrelated	144	1/30/14	20	yes	21	1.49	0.3	1.1	0.1	y	y	no	death
MDS-12	65	F	relapse of multiple myeloma	1	unrelated	430	2/20/14	27	yes	73	1.26	0.7	0.4	0.1	y	y	Influenza A	death
MDS-14	33	F	relapse of lymphoblastic acute leukemia	2	unrelated	335	1/8/13	25	yes	102	6.92	4.8	1.3	0.6	y	y	Influenza A + Bocavirus	death
MDS-16	33	M	relapse of mycosis fungoides	1	related	1052	1/16/15	27	yes	256	14.05	11.3	2.1	0.7	y	y	no	death
MDS 15	40	F	acute myeloid leukemia	1	unrelated	103	2/11/14	24	yes	125	2.04	0.5	1.1	0.5	y	n	no	
Means						494.4				95.3	4.5	3.0	1.0	0.5				