

# Genetic variability and circulation pattern of human metapneumovirus isolated in Italy over five epidemic seasons

Alessandra Lo Presti<sup>1</sup>, Roberta Cammarota<sup>2</sup>, Paola Apostoli<sup>3</sup>, Eleonora Cella<sup>1</sup>, Simona Fiorentini<sup>3</sup>, Muhammed Babakir-Mina<sup>3</sup>, Marco Ciotti<sup>4</sup>, Massimo Ciccozzi<sup>1</sup>

<sup>1</sup>Department of Infectious, Parasitic and Immunomediated Diseases, Istituto Superiore di Sanità, Rome, Italy;

<sup>2</sup>Virology, Department of Experimental Medicine and Biochemistry, Tor Vergata University, Rome, Italy;

<sup>3</sup>Department of Experimental and Applied Medicine, Section of Microbiology,

University of Brescia Medical School and Pediatric Unit, Brescia Civic Hospital, Brescia, Italy;

<sup>4</sup>Laboratory of Molecular Virology, Foundation Policlinic Tor Vergata, Rome, Italy

## SUMMARY

Human metapneumovirus (hMPV) is an important aetiological agent of respiratory tract infection (RTI) in infants. Based on genetic differences, hMPVs are separated into two main groups, A and B, and two subgroups 1 and 2. To better understand the genetic diversity of hMPV, we analyzed 435 bp fragments of the F gene from 49 isolates obtained from a paediatric clinical centre in Northern Italy from 2005 to 2009. The phylogenetic analysis showed that our hMPV sequences clustered into five main clades (A1, A2a, A2b, B1, B2) statistically supported. Most of the strains belong to A2 (49%) and to B1 (28.5%) lineages. The intermixing in the phylogenetic tree showed no seasonal distribution between samples collected over a 5 year period. Mean genetic distances showed that clade A2 was more heterogeneous than others. In our F glycoprotein dataset we observed only two positively selected sites with an  $\omega$  value of 1.408 and 1.429, respectively, and 27 negatively selected sites. The two positively selected sites could be considered evolutionary "hot spots" because they were under positive selection and/or relaxed selective constraints. The abundant negatively selected sites reflect a high degree of conservation of F protein, probably necessary for viral infection.

**KEY WORDS:** hMPV, Genetic variability, F gene, Selective pressure

Received March 04, 2011

Accepted May 17, 2011

## INTRODUCTION

Human metapneumovirus (hMPV), originally isolated in The Netherlands in 2001 (van den Hoogen *et al.*, 2001), is an important aetiological agent of respiratory tract infection (RTI) in infants. Clinical symptoms range from influenza-like symptoms (i.e. fever, cough, rhinorrhea) to severe lower respiratory tract infection (Bastien *et al.*, 2003; van den Hoogen *et al.*, 2001, 2003;

Williams *et al.*, 2004). The virus has been also linked to acute respiratory disease in both the elderly and immunocompromised patients (Boivin *et al.*, 2002; Pelletier *et al.*, 2002; Falsey *et al.*, 2003; Boivin *et al.*, 2003).

Like human respiratory syncytial virus (hRSV), hMPV seems to be ubiquitous and may cause repeated respiratory tract infections throughout life probably because of incomplete immunity or genetic variability (Pelletier *et al.*, 2002). Moreover, the two viruses exhibit the same seasonal epidemic distribution with a high prevalence mostly in the winter-spring period (Kashiwa *et al.*, 2004; Caracciolo *et al.*, 2008).

hMPV and avian metapneumovirus (aMPV), members of the *Metapneumovirus* genus, were found to be closely related. Both belong to the

Corresponding author

Massimo Ciccozzi

Department of Infectious

Parasitic and Immunomediated Diseases

Istituto Superiore di Sanità

Viale Regina Elena, 299 - 00161 Rome, Italy

E-mail: massimo.ciccozzi@iss.it Rome, Italy

*Paramyxoviridae* family, *Pneumovirinae* subfamily as well as hRSV. The aMPVs have been classified into four groups: A to D. The aMPV subgroup C, first detected in the USA in 1996, appears more closely related to hMPV than any other aMPV subgroups (Govindarajan & Samal, 2005; Govindarajan *et al.*, 2004; Toquin *et al.*, 2003). It has been hypothesized that the hMPV probably originated from birds and the common ancestor of aMPV-C and hMPV existed around 200 years ago (de Graaf *et al.*, 2008).

hMPV is an enveloped, non-segmented single stranded negative sense RNA virus. The genome is approximately 13 Kb in length, and the 3'-to-5' gene order is N-P-M-F-M2-1/M2-2-SH-G-L, which is the same as that of the aMPV (van den Hoogan *et al.*, 2001).

Metapneumoviruses encode for three membrane glycoproteins, the attachment (G), the small hydrophobic (SH) and the fusion (F) protein. The hMPV and aMPV-C SH and G proteins are type II membrane proteins that are not essential for virus replication *in vitro* (Biacchesi *et al.*, 2004; Naylor *et al.*, 2004). Moreover, recombinant hMPV from which the SH gene has been deleted replicates with normal kinetics in hamsters and monkeys (Biacchesi *et al.*, 2004). In contrast, hMPV from which the G gene has been deleted is viable but attenuated (Biacchesi *et al.*, 2004). Thus, it appears that the F protein by itself can promote attachment and fusion. The F protein is a type I membrane protein and is synthesized as an inactive precursor F<sub>0</sub> that is cleaved by host proteases into the functional subunits F<sub>1</sub> and F<sub>2</sub> (Lamb *et al.*, 2006; Russell *et al.*, 2001). The F protein mediates fusion of the virus envelope with the cell membrane during viral entry and induces syncytium formation in infected cells (Herfst *et al.*, 2008; Schowalter *et al.*, 2006). For bovine respiratory syncytial virus (bRSV) and hRSV, the F protein determines the cellular host range (Schlender *et al.*, 2003). The F protein of hMPV is highly immunogenic and protective, whereas G and SH are not (Herfst *et al.*, 2007; Skiadopoulos *et al.*, 2006).

Based on genetic differences, hMPVs are separated into two groups, A and B, with each group divided into genetic subgroups 1 and 2 (Caracciolo *et al.*, 2008). More recently, phylogenetic analyses showed a further bipartition of subgroup A2 into two new genetic clusters (sub-

lineages), designated A2a and A2b (Caracciolo *et al.*, 2008). Among all the sub-lineages of hMPV the A2 sub-lineage shows the greatest diversity (Huck *et al.*, 2006).

For our analysis we chose the F surface glycoprotein because it is highly conserved (Biacchesi *et al.*, 2003) and represent the most important target of protective immunity (Williams *et al.*, 2007). To date, only few Italian studies have analysed the sequence data obtained during hMPV epidemics, and the molecular evolution of the virus over time remains an open question. To better understand how hMPV evolves, in terms of time and mode, we analyzed 435 bp fragments of the F gene from 49 isolates obtained from a paediatric clinical centre in Northern Italy during five consecutive years (2005 to 2009).

## MATERIAL AND METHODS

### RNA extraction, RT-PCR and sequencing

Forty-nine samples obtained from hMPV positive hospitalized children with respiratory symptoms in the urban area of Brescia were analyzed. Nasopharyngeal samples were collected from infants 1 month to 13 months old (mean age 12 months) during the period from 2005-2009. All the specimens were subjected to RT-PCR for the F gene as described below.

Viral RNA was extracted from 200 µl of nasopharyngeal aspirates with the RNAeasy Kit (QIAGEN, Milan, Italy), according to the manufacturer's instructions and the cDNA was synthesized from 20 µl of eluted RNA using the cDNA Archive mini kit (Applied Biosystems, Monza, Italy).

Primers used for the amplification of the hMPV fusion F gene were as follows: F-f (5'-GTY AGC TTC AGT CAA TTC AAC AGA AG-3'); F-r (5'-CCT GTG CTG ACT TTG CAT GGG-3') (Caracciolo *et al.*, 2008). The amplified products were analyzed by electrophoresis on 2% agarose gel stained with ethidium bromide, and the size of the amplicons was compared with that of standard molecular weight markers. To validate the amplification process and to exclude the presence of carryover contamination, positive and negative controls were run in each PCR.

The amplified fragment was purified with Wizard SV Gel and PCR Clean-Up System (Promega, Milan, Italy) and sequenced at the CRIBI BMR

TABLE 1 - Reference sequences downloaded from GenBank (<http://www.ncbi.nlm.nih.gov>).

Viral strains	Accession numbers
HMPV isolate TN/99/4-6	EU857545
HMPV isolate CAN97-83	AY145296
HMPV isolate NL/00/17	FJ168779
HMPV isolate JPY88-12	AY622381
HMPV isolate TN/89/7-13	EU857569
HMPV isolate TN/99/4-19	EU857581
HMPV isolate JPS03-240	AY530095
HMPV isolate CAN97-83	AY297749
HMPV isolate 00-1	AF371337
HMPV isolate NL/1/99	AY525843
HMPV isolate CAN98-75	AY297748

Genomics Sequence Facility (Padua, Italy). The obtained sequences (49; 435nt) were labelled using a letter code representing the geographic area of isolation (i.e. "BS" represents Brescia) followed by the hospital ID. Reference sequences are reported in Table 1. The Italian isolates are labelled from BS1 to BS49 and referred at Acc. Number from JF907707 to JF907755.

### Inferring phylogenetic tree

The phylogenetic analysis was carried out on a 435 bp fragment of the F gene (amino acid positions: 222 to 367) of the 49 hMPV isolates and compared with 11 reference sequences downloaded from the NCBI data base (<http://www.ncbi.nlm.nih.gov>), Table 1. All the sequences in the data set were aligned using CLUSTAL X software (Thompson *et al.*, 1994), then manually edited with the Bioedit software (version 7.0.9) (Hall *et al.*, 1999). A phylogenetic tree was estimated using the PAUP\* package (Swofford, 2002). The Tamura-Nei (TrN93) model of nucleotide substitution, incorporating Maximum Likelihood (ML) estimates of base composition and the shape parameter ( $\alpha$ ) of a gamma distribution ( $\Gamma$ ) model of among-site rate variation as it consistently gave much higher like-

lihood value using Modeltest v.3.7 implemented in PAUP\* was employed (Posada and Crandall 1998; Posada and Buckley, 2004). This model assumes different nucleotide frequencies, a transition/transversion bias, the equality of substitution rates over sites modelled by a gamma distribution (Yang *et al.*, 1994) and also allows purine (A $\leftrightarrow$ G) and pyrimidine (C $\leftrightarrow$ T) transition to occur at different rates (Tamura *et al.*, 1993). A Maximum Likelihood tree was estimated under this model using tree bisection-reconnection (TBR) branch swapping. The statistical robustness and reliability of the branching order within each phylogenetic tree were confirmed by bootstrap analysis using 1,000 replicates for the Neighbor-Joining (NJ) tree and by the Zero Branch Length Test for the ML tree (Swofford, 2002).

The software MEGA 4.1 (Tamura *et al.*, 2007) allowed to calculate the genetic distances among different clades using the model of nucleotide maximum composite likelihood (MCL) including transitions and transversions. The substitution rates were set equal among sites but different among lineages. Gaps were treated by pairwise-deletion, i.e. ignoring only the gaps which are involved in the comparison of a pair of sequences.

### Selective pressure analysis

The  $d_N/d_S$  rate ( $\omega$ ) was estimated by ML approach implemented in the program HyPhy (Pond and Muse, 2005). In particular, the global (assuming a single selective pressure for all branches) and the local (allowing the selective pressure to change along every branch) models were compared by likelihood ratio test (LRT). Site specific positive and negative selections were estimated by two different algorithms: the fixed-effects likelihood approach (FEL) which fits an  $\omega$  rate to every site and use likelihood ratio to test if  $d_N \neq d_S$ ; and a random effect likelihood (REL), a variant of the Nielsen-Yang approach (Nielsen & Yang, 1998), which assumes that exists a discrete distribution of rates across sites and allows both  $d_S$  and  $d_N$  to vary site-by-site independently (Kosakovsky Pond and Frost, 2005). Sites under selective pressure were selected assuming a p value of  $\leq 0.1$  or a posterior probability of  $\geq 0.9$ . Some analyses were carried out using the Web based interface Datamonkey (<http://www.datamonkey.org/>) (Pond and Frost, 2005).

For the evolutionary analysis, the hMPV isolate

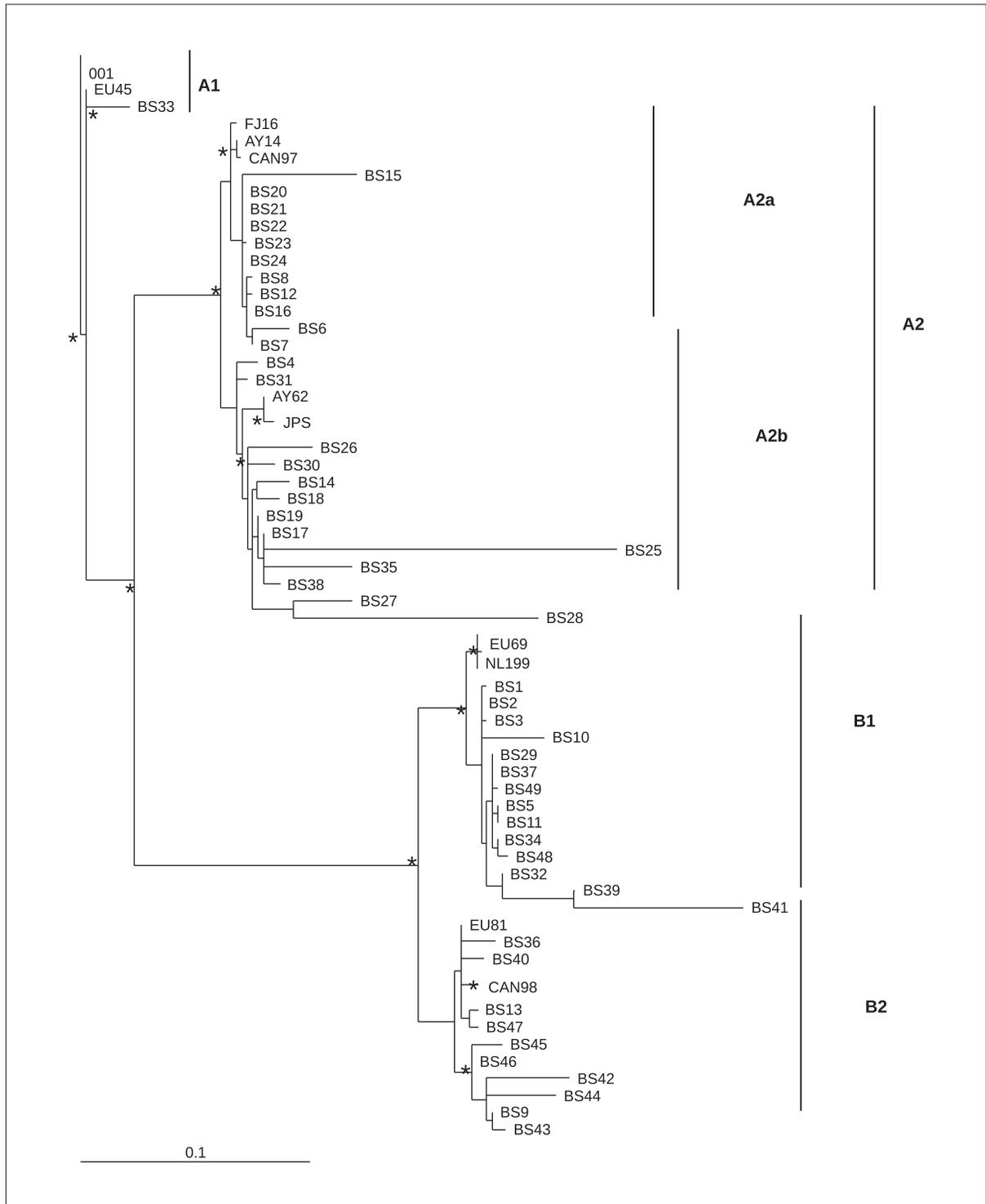


FIGURE 1 - Maximum likelihood phylogenetic analysis of hMPV F gene sequences. Analysis was carried out on a 435 nt fragment length. Branch lengths were estimated with the best fitting nucleotide substitution model according to a hierarchical likelihood ratio test [Posada and Crandall, 1998], and were drawn in scale with the bar at the bottom indicating 0.1 nucleotide substitutions per site. One \* along a branch represents significant statistical support for the clade subtending that branch ( $p < 0.001$  in the zero-branch-length test) and bootstrap support >75%.

00-1 (accession number AF371337) was used as a reference sequence to trace the exact position of the amino acids found under positive selection.

## RESULTS

### Phylogenetic and evolutionary analysis

The phylogenetic analysis showed that the hMPV sequences clustered into five main clades (A1, A2a, A2b, B1, and B2), statistically supported by bootstrap values > 75%, Figure 1.

Of the 49 strains, 1 (2.04%) grouped in the A1 lineage, 11 (22.4%) grouped in the A2a lineage, 13 (26.5%) grouped in the A2b lineage, 14 (28.5%) grouped in the B1 lineage and 10 (20.4%) grouped in the B2 lineage.

The genetic distance was measured grouping the sequences of A2, B1, B2 clades. The A1 clade was excluded from the analysis because it contained only one sequence (BS33) of the 49 identified in our samples. The nucleotide genetic distance between and within clades showed that the B1 clade is more distant from A2 clade than B2 (0.242 vs 0.088). This finding was also confirmed by the amino acidic genetic distance (0.029 vs 0.022). The within nucleotide genetic distance showed that clade A2 is more heterogeneous than others (0.029), while the within amino acidic genetic distance showed clade B2 as more heterogeneous than others (0.019). Analyzing hMPV F gene sequences, it was found that the  $\alpha$  parameter of the gamma distribution for this gene was 1.78. The distribution has a characteristic bell-shape and models weak rate heterogeneity over sites.

Selection pressure analysis through computation of the ratio of non-synonymous (dN) to synony-

mous (dS) substitutions per site dN/dS ( $\omega$ ) revealed limited positive selection and abundant negative selection for the F proteins. Indeed, only two positively selected sites were found in the F protein (aa 294 and 304 of hMPV isolate 00-1, accession number AF371337) confirmed by both REL and FEL analysis. The  $\omega$  values were 1.408 for the aa position 294 and 1.429 for aa 304 position (Table 2), respectively. These positively selected sites were located in the extracellular domain of the F protein. Twenty-seven sites were negatively selected as confirmed by both REL and FEL analysis.

## DISCUSSION

Many factors may influence the genetic evolution of hMPV strains, including virus infectivity, pre-existing immunity in the population, the low variability at DNA level and antigenic variability. In particular, antigenic variation may play an important role in the ability of these viruses to escape the pre-existing immune response. In our study, a molecular evolutionary approach was used to investigate the presence of sites under positive and negative selection in the genome of the hMPV using a selective pressure analysis method.

We analyzed 49 partial hMPV F gene sequences obtained from Italian samples (Brescia hospital) collected over a 5-year period, from 2005 to 2009. The classification and naming of serotypes/serogroups, subgroups, strains, variants and isolates of hMPV is still evolving. Genomic sequence and phylogenetic analysis confirmed that there are two (A, B) major genotypes of hMPV.

TABLE 2 - Selection analysis for the partial fusion protein (F) gene, using REV as evolutionary model. Different amino acid present at each site under positive selection are given between brackets.

Positively selected sites* ( $\omega$ for sites >1) HYPHY software	294 (G, K, E); 304 (R, K, Q)
Negatively selected sites** ( $\omega$ for sites <1) HYPHY software	E226; R229; A230; S232; Q240; L243; L245; R248; V251; R252; R253; K254; F256; L259; S266; G277; V278; T281; K287; A288; C292; K295; N298; C301; L303; D331; K362

\*Positively selected sites are numbered according to amino acid position of HMPV isolate 00-1 accession number AF371337. The different amino acid present at each site under positive selection are given between brackets. \*\*Negatively selected sites are numbered according to amino acid position of HMPV isolate 00-1 accession number AF371337.

Our data confirm five distinct genetic lineages of hMPV, designated as A1, A2a, A2b, B1, B2.

Most of the Brescia strains belong to A2 (49%) and to B1 (28.5%) lineages. In particular, of the 49 strains, only 1 (2.04%) grouped in the A1 lineage, 11 (22.4%) grouped in the A2a lineage, 13 (26.5%) grouped in the A2b lineage, 14 (28.5%) grouped in the B1 lineage and 10 (20.4%) grouped in the B2 lineage. Analyzing the mean genetic distances we found that clade A2 was more heterogeneous than the others as reported (Huck *et al.*, 2006). Further investigation on clade A2 would be necessary, since a study by Vicente *et al.* (2006) suggested that genotype A might be more pathogenic than genotype B, causing more severe clinical symptoms in children.

In our evolutionary analysis, since the average of the non synonymous-synonymous ratio ( $\omega$ ) is usually not sensitive enough to detect Darwinian selection at the molecular level, codon substitution models were used to detect sites under positive and negative selection.

In our F glycoprotein dataset we observed only two positively selected sites (confirmed both by REL and FEL analysis) with an  $\omega$  value of 1.408 and 1.429, respectively, both located in the extracellular domain of the F<sub>1</sub> chain in the cystein-rich region. These values fell over the threshold of  $\omega > 1$ , and therefore they were indicative of positive selection. Indeed, negatively selected sites were abundant for the F gene (27 negatively selected sites, Table 2), reflecting a high degree of conservation of the F protein, probably necessary for viral infection.

The hMPV F gene sequences analysis showed a value of  $\alpha$  parameter of the gamma distribution for this gene higher than 1, meaning a weak rate heterogeneity over sites. These results could be linked to the short length of F sequences available and to the closer years of isolation; the evolutionary analysis could be partially affected by these factors. However, the two positively selected sites could be considered evolutionary "hot spots" because they were under positive selection and/or relaxed selective constraints. Therefore, although a high proportion of amino acids are largely invariable probably because of structural and functional constraints, adaptive evolution may occur at certain sites of the genome.

Regarding the seasonal distribution the Italian samples collected over a 5 year period appear to

be intermixed in the phylogenetic tree indicating no difference in the strain (isolates) collected in the different years.

Unlike influenza virus that exhibits rapid genetic drift associated with antigenic variation resulting in immune evasion, hMPV does not appear to exhibit a progressive genetic evolution. In this respect, hMPV appears to be similar to other paramyxoviruses. RNA viruses mutate frequently due to the lack of proofreading activity of RNA-dependent-RNA polymerases. The reason for the lack of directional antigenic drift in paramyxoviruses is not clear. There could be functional constraints on paramyxovirus fusion proteins to prevent such drastic amino acid changes. In contrast to paramyxoviruses, the analogous influenza virus hemagglutinin and human immunodeficiency virus gp120 fusion proteins are capable of substantial mutation to escape neutralizing antibodies without loss of function. The mechanism of the functional constraints on paramyxovirus fusion protein diversity warrants further investigation.

A final consideration arising from our data and confirmed by other recent works is that the high degree of stability of the F protein could be important for the development of monoclonal antibodies and vaccines which will not need to be continuously updated.

## REFERENCES

- BASTIEN N., WARD D., VAN CAESELE P., BRANDT K., LEE S.H., McNABB G., KLISKO B., CHAN E., LI Y. (2003). Human metapneumovirus infection in the Canadian population. *J. Clin. Microbiol.* **41**, 4642-4646.
- BIACCHESI S., SKIADOPOULOS M.H., YANG L., LAMIRANDE E.W., TRAN K.C., MURPHY B.R., COLLINS P. L., BUCHHOLZ U.J. (2004). Recombinant human metapneumovirus lacking the small hydrophobic SH and/or attachment G glycoprotein: deletion of G yields a promising vaccine candidate. *J. Virol.* **78**, 12877-12887.
- BIACCHESI S., SKIADOPOULOS M.H., BOIVIN G., HANSON C.T. MURPHY B.R., COLLINS P.L., BUCHHOLZ U.J. (2003). Genetic diversity between human metapneumovirus subgroups. *Virology*. **315**, 1-9.
- BOIVIN G., DE SERRES G., CÔTÉ S., GILCA R., ABED Y., ROCHETTE L., BERGERON M.G., DÉRY P. (2003). Human metapneumovirus infections in hospitalized children. *Emerg. Infect. Dis.* **9**, 634-640.

- BOIVIN G., ABED Y., PELLETIER G., RUEL L., MOISAN D., CÔTÉ S., PERET T.C., ERDMAN D.D., ANDERSON L.J. (2002). Virological Features and Clinical Manifestations Associated with Human Metapneumovirus: A New Paramyxovirus Responsible for Acute Respiratory-Tract Infections in All Age Groups. *J. Infect. Dis.* **186**, 1330-1334.
- CARACCILO S., MININI C., COLOMBRITA D., ROSSI D., MIGLIETTI N., VETTORE E., CARUSO A., FIORENTINI S. (2008). Human metapneumovirus infection in young children hospitalized with acute respiratory tract disease: virologic and clinical features. *Pediatr. Infect. Dis. J.* **27**, 406-412.
- DE GRAAF M., OSTERHAUS A.D.M.E., FOUCHIER R.A.M., HOLMES E.C. (2008). Evolutionary dynamics of human and avian metapneumoviruses. *J. Gen. Virol.* **89**, 2933-2942.
- FALSEY A.R., ERDMAN D., ANDERSON L.J., WALSH E.E. (2003). Human metapneumovirus infections in young and elderly adults. *J. Infect. Dis.* **187**, 785-790.
- GOVINDARAJAN D., SAMAL S.K. (2005). Analysis of the complete genome sequence of avian metapneumovirus subgroup C indicates that it possesses the longest genome among metapneumoviruses. *Virus Genes.* **30**, 331-333.
- GOVINDARAJAN D., YUNUS A.S., SAMAL S.K. (2004). Complete sequence of the G glycoprotein gene of avian metapneumovirus subgroup C and identification of a divergent domain in the predicted protein. *J. Gen. Virol.* **85**, 3671-3675.
- HALL T.A. (1999). Bioedit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98 NT. *Nucl. Acids. Symp. Ser.* **41**, 95-98.
- HERFST S., MAS V., VER L.S., WIERDA R.J., OSTERHAUS A.D., FOUCHIER R.A., MELERO J.A. (2008). Low pH induced membrane fusion mediated by human metapneumovirus F protein is a rare, strain dependent phenomenon. *J. Virol.* **82**, 8891-8895.
- HERFST S., DE GRAAF M., SCHRAUWEN E.J., ÜLBRANDT N.D., BARNES A.S., SENTHIL K., OSTERHAUS A.D., FOUCHIER R.A., VAN DEN HOOGEN B.G. (2007). Immunization of Syrian golden hamsters with F subunit vaccine of human metapneumovirus induces protection against challenge with homologous or heterologous strains. *J. Gen. Virol.* **88**, 2702-2709.
- HUCK B., SCHARF G., NEUMANN-HEIFELIN D., PUPPE W., WEIGL J., FALCONE V. (2006). Novel human metapneumovirus sublineage. *Emerg. Infect. Dis.* **12**, 147-150.
- KASHIWA H., SHIMOZONO H., TAKAO S. (2004). Clinical pictures of children with human metapneumovirus infection: comparison with respiratory syncytial virus infection. *Jpn. J. Infect. Dis.* **57**, 80-82.
- KOSAKOVSKY POND S.L., FROST S.D.W. (2005). Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol. Biol. Evol.* **22**, 1208-1222.
- LAMB R.A., PATERSON R.G., JARDETZKY T.S. (2006). Paramyxovirus membrane fusion: lessons from the F and HN atomic structures. *Virology.* **344**, 30-37.
- NIELSEN R., YANG Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics.* **148**, 929-936.
- NAYLOR C.J., BROWN P.A., EDWORTHY N., LING R., JONES R.C., SAVAGE C.E., EASTON A.J. (2004). Development of a reverse-genetics system for Avian pneumovirus demonstrates that the small hydrophobic (SH) and attachment (G) genes are not essential for virus viability. *J. Gen. Virol.* **85**, 3219-3227.
- PELLETIER G., DÉRY P., ABED Y., BOIVIN G. (2002). Respiratory tract reinfections by the new human Metapneumovirus in an immunocompromised child. *Emerg. Infect. Dis.* **9**, 976-978.
- POND S.L., FROST S.D. (2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics.* **21**, 2531-2533.
- POND S.K., MUSE S.V. (2005). Site-to-site variation of synonymous substitution rates. *Mol. Biol. Evol.* **22**, 2375-2385.
- POSADA D., BUCKLEY T.R. (2004). Model selection and model averaging in phylogenetics: Advantages of the AIC and Bayesian approaches over the likelihood ratio tests. *Syst. Biol.* **53**, 793-808.
- POSADA D., CRANDALL K.A. (1998). MODEL TEST: Testing the model of DNA substitution. *Bioinformatics.* **14**, 817-818.
- RUSSELL C.J., JARDETZKY T.S., LAMB R.A. (2001). Membrane fusion machines of paramyxoviruses: capture of intermediates of fusion. *EMBO J.* **20**, 4024-4034.
- SCHLENDER J., ZIMMER G., HERRLER G., CONZELMANN K.K. (2003). Respiratory syncytial virus (RSV) fusion protein subunit F2, not attachment protein G, determines the specificity of RSV infection. *J. Virol.* **77**, 4609-16.
- SCHWALTER R.M., SMITH S.E., DUTCH R.E. (2006). Characterization of human metapneumovirus F protein-promoted membrane fusion: critical roles for proteolytic processing and low pH. *J. Virol.* **80**, 10931-10941.
- SKIADOPOULOS M.H., BIACCHESI S., BUCHHOLZ U.J., AMARO-CARAMBOT E., SURMAN S.R., COLLINS P.L., MURPHY B.R. (2006). Individual contributions of the human metapneumovirus F, G, and SH surface glycoproteins to the induction of neutralizing antibodies and protective immunity. *Virology.* **345**, 492-501.
- SWOFFORD D.L. (2002). PAUP. Phylogenetic analysis using parsimony (\*and other methods). Version 4.0. Sinauer Associates, Sunderland, Massachusetts.
- TAMURA K., DUDLEY J., NEI M., KUMAR S. (2007).

- MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596-1599.
- TAMURA K., NEI M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512-526.
- THOMPSON J.D., GIBSON T.J., PLEWNIAK F., JEANMOUGIN F., HIGGINS D.G. (1994). CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.
- TOQUIN D., DE BOISSESON C., BEVEN V., SENNE D.A., ETERRADOSSI N. (2003). Subgroup C avian metapneumovirus (MPV) and the recently isolated human MPV exhibit a common organization but have extensive sequence divergence in their putative SH and G genes. *J. Gen. Virol.* **84**, 2169-2178.
- VAN DEN HOOGEN B.G., VAN DOORNUM G.J.J., FOCKENS J.C., CORNELISSEN J.J., BEYER W.E.P., DE GROOT R., OSTERHAUS ADME, FOUCHIER R.A.M. (2003). Prevalence and clinical symptoms of human metapneumovirus infection in hospitalized patients. *J. Infect. Dis.* **188**, 1571-1577.
- VAN DEN HOOGEN B.G., DE JONG J.C., GROEN J., KUIKEN T., DE GROOT R., FOUCHIER R.A., OSTERHAUS A.D. (2001). A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat. Med.* **7**, 719-724.
- VICENTE D., MONTES M., CILLA G., PEREZ-YARZA E.G. AND PEREZ-TRALLERO E. (2006). Differences in clinical severity between genotype A and genotype B human metapneumovirus infection in children; *Clin. Infect. Dis.* **15**, e111-e113.
- WILLIAMS J.V., CHEN Z., CSEKE G., WRIGHT D.W., KEEFER C.J., TOLLEFSON S.J. HESSELL A., PODSIAD A., SHEPHERD B.E., SANNA P.P., BURTON D.R., CROWE J.E. JR, WILLIAMSON R.A. (2007). A recombinant human monoclonal antibody to human metapneumovirus fusion protein that neutralizes virus in vitro and is effective therapeutically in vivo. *J. Virol.* **81**, 8315-8324.
- WILLIAMS J.V., HARRIS P.A., TOLLEFSON S.J., HALBURNT-RUSH L.L., PINGSTERHAUS J.M., EDWARDS K.M., WRIGHT P.F., CROWE J.E JR. (2004). Human metapneumovirus and lower respiratory tract disease in otherwise healthy infants and children. *N. Engl. J. Med.* **350**, 443-450.
- YANG Z. (1994). Maximum-Likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306-314.