Novel variants of respiratory syncytial virus A ON1 associated with increased clinical severity of bronchiolitis

Fabio Midulla[1], Greta Di Mattia[1], Raffaella Nenna[1], Carolina Scagnolari[2], Agnese Viscido[2], Giuseppe Oliveto[2], Laura Petrarca[1], Antonella Frassanito[1], Serena Arima[3], Guido Antonelli[2], Alessandra Pierangeli[2].

[1]Department of Pediatrics, [2]Virology Laboratory, Department of Molecular Medicine, [3]Department of Methods and Models in Economics, the Territory and Finance, "Sapienza" University, Rome, Italy.

The results were partially presented at the ATS congress, Dallas (TX), May 17-22 2019 and at the ERS congress Madrid (Spain) September 28 – October 2 2019.

Summary: Over six bronchiolitis seasons (2012-2013/2017-2018), the only RSV-A genotype circulating in Rome was ON1; isolates showed increasing genetic divergence over time. Our results may suggest that the divergence of ON1 strains was associated with the increase of bronchiolitis clinical severity.

Reprint requests and correspondence to:

Prof. Fabio Midulla

Department of Paediatrics, "Sapienza" University of Rome, V.le Regina Elena 324, 00161, Rome – Italy Tel 00390649979363 – FAX 00390649977412 E-mail: midulla@uniroma1.it

Abstract

Objective: to study RSV-A genotype ON1 genetic variability and clinical severity in infants hospitalized with bronchiolitis, over six epidemic seasons (2012-2013 to 2017-2018).

Methods: From prospectively enrolled term infants hospitalized for bronchiolitis, samples positive to RSV-A ON1 (N=139) were sequenced in the second half of the G gene. Patients' clinical data were obtained from medical files and each infant was assigned a clinical severity score. ANOVA comparison and adjusted multinomial logistic regression were used to evaluate clinical severity score and clinical parameters.

Results: The phylogenetic analysis of 54 unique strains showed three distinct clades, in which sequences of the last two seasons set apart from the others. The most divergent and numerous cluster of 2017-2018 strains was characterized by a novel pattern of amino acid (aa) changes, some of which located in antigenic sites. Several aa changes altered predicted glycosylation sites; divergent lineages were characterized by the acquisition of around ten new O-glycosylation sites. Clinical severity of bronchiolitis increased in the 2016-2017 and 2017-2018 seasons and changed according to the epidemic seasons only.

Conclusion: The aa changes detected in the hypervariable part of the G protein may have altered functions and/or changed its immunogenicity determining an impact on disease severity.

Key words: respiratory syncytial virus, bronchiolitis, genetic variability, genotypes, clinical severity

Background

Bronchiolitis is the most frequent lower respiratory tract viral infection in infants younger than 12 months of age and the main cause of hospitalization in this age group [1-3]. Around 150 million new cases of bronchiolitis are diagnosed each year, with 2-3% of affected children requiring hospital admission [4]. The typical seasonal pattern of bronchiolitis that occurs between November and May, reflects the activity of respiratory syncytial virus (RSV) [3].

RSV is an enveloped, single-stranded, negative-sense RNA virus of the *Pneumoviridae* family. The genome encodes for 11 proteins, among which the external glycoproteins G and F are involved into attachment and entry in the host cells [5]. The G extracellular part is composed of a central conserved domain (CCD) and two hypervariable regions presenting many N- and O-glycosylation sites, which contribute to antigenicity [5, 6]. Reactions with monoclonal antibodies classify RSV into subtypes A and B and sequencing of the second high variable region of the G gene differentiates RSV A and RSV B into 13 and 20 genotypes, respectively [5-6]. The subtypes A and B usually co-circulate during epidemic seasons, with predominance of RSV A worldwide [5, 6]. Variability of RSV genome causes the emergence of new clades, which may co-circulate or replace the previous circulating viral strains. RSV A ON1, that is characterized by the duplication of 23 amino acids in the C-terminal region of the G protein, was firstly detected in Ontario, Canada, in 2010 [7] and rapidly spread replacing the previous circulating RSV A genotypes [8-12]. Many authors questioned whether the fitness advantage of ON1 was due to a differential replication ability and/or to a more rapid evolutionary rate than other RSV genotypes [7-13]. In its first years of circulation in Rome, we reported that ON1 bronchiolitis had a mild clinical course [8]. Moreover, we recently demonstrated in a large and homogeneous cohort of infants hospitalized with bronchiolitis in Rome – Italy over 12 epidemic seasons, that RSV A ON1

firstly detected in 2011-2012 season, replaced the previously circulating RSV A NA1 since 2012-2013 and we confirmed that it caused milder form of bronchiolitis with respect to the previous RSV A strain, NA1 [14]. Indeed, more clinical and virological data on ON1 infections over years are needed. Of outstanding interest is to understand whether the increased fitness of this virus is associated with increased severity and immune evasion influencing the vaccine strategy.

We aimed to explore genetic variability of RSV A ON1 strains infecting infants hospitalized with bronchiolitis, over six epidemic seasons. We also analyzed patients' data, to verify if ON1 variability was associated to change in disease severity.

Methods

Study group

We prospectively enrolled infants hospitalized for bronchiolitis in the Pediatric Emergency Department, "Sapienza" University of Rome, during the epidemic seasons 2012-2013 to 2017-2018. Bronchiolitis was defined as the first lower respiratory tract infection in infants < 12 months, characterized by a history of upper respiratory tract infection followed by the onset of cough, respiratory distress and diffuse crackles on chest auscultation [1]. Infants with underlying chronic diseases (e.g., cystic fibrosis, congenital heart disease, immunodeficiency) or prematurity were excluded. Demographic and clinical data were obtained from parents with a structured questionnaire and from medical files. At admission to the hospital, physicians assigned to each infant a clinical severity score ranging from 0 to 8. The score consists of the sum of four parameters: 1-respiratory rate (0= <40, 1= 40-60, 2= >60 breaths/min); 2- oxygen saturation on room air (SaO2) (0= >95%, 1= 90%-95%, 2= <90%);

3- presence of retractions (0= absent, 1= mild, 2= severe); and 4- ability to tolerate oral feeding (0= normal, 1= <75% of the usual food intake, 2= intra venous rehydration), as previously described [1]. Bronchiolitis was classified according to the clinical severity score in mild (score 0-3), moderate (score 4-5), and severe (score 6-8).

In line with confidentiality requirements, the database was anonymized, and the ethic committees of Policlinico Umberto I approved the study (Prot. 107/12) after the informed consent was obtained from infants' parents.

RSV detection and phylogenetic analysis

Within 1 day after hospitalization, infants underwent a nasopharyngeal washing (NPW). Samples were delivered on ice within 1-2 hours to the laboratory of Virology, and divided into two aliquots: one was used for nucleic acid extraction and the other was stored at -80°C for further analysis. Fourteen respiratory viruses (RSV, influenza virus A/B, coronaviruses OC43, 229E, NL-63, HUK1, adenovirus, rhinovirus, parainfluenza 1-3, metapneumovirus, and bocavirus) were tested with reverse transcriptase polymerase chain reactions (RT-PCR) as previously described [15]. Samples positive to RSV were sequenced in part of the G conserved region and the second hypervariable region up to the stop codon (380 nt; amino acids 172-297) as previously described [8]. 139 were , RSV-A positive samples, ON1 genotype, and sequences were aligned with reference sequences using Bioedit v7.1.3, to identify sequencing errors that were removed and redundant sequences that were grouped. The final dataset includes 54 unique ON1 sequences isolated in Rome during six epidemic seasons from 2012-13 to 2017-18 and 4 reference strains, i.e. the first Ontario strain (ON1-1.1), the divergent genotype ON1-1.2, first detected in Ancona [8], the closest ON1 ancestor NA1 [7], and the reference genotype GA2 [8]. The best-fit evolutionary model and parameters were selected using Model test on Mega 6.06 [16]: accordingly, the phylogenetic

tree was constructed using the Maximum Likelihood method based on the Tamura-Nei model and a discrete Gamma distribution with 5 categories (+G) to model evolutionary rate differences among sites, with bootstrap values of 1,000. The average p-distance value (i.e. the proportion of nucleotide sites that differ between sequences) of the ON1 strains over the six epidemic seasons was calculated by pairwise comparison, in Mega 6.06.

The bioinformatic tools NetNGlyc and NetOGlyc (http://www.cbs.dtu.dk/services) were used to predict N- and O-linked glycosylation sites in ON1 amino acid sequences.

Statistical analysis

Binary variables such as sex, ethnicity, familiar atopy, passive smoking exposure, breastfeeding, and type of birth, were compared in the different epidemic seasons using univariate logistic regression model and post-hoc comparison of the coefficients of the epidemic season. For quantitative variables, such as age, gestational age and birth weight, in order to evaluate which of these variables significantly vary across epidemic seasons, we used univariate regression model and post-hoc comparison of the coefficients of the epidemic season. The clinical severity score has been treated as a 3-categorical variable. Multinomial logistic regression was used to evaluate whether the clinical severity score changed across epidemic seasons and to control the evolution of the clinical severity score in the different epidemic seasons for confounding variables (sex, age, ethnicity, gestational age, type of birth, birth weight, familiar atopy, passive smoking exposure, breastfeeding and type of birth). The first epidemic season has been set as corner point in the standard ANOVA parameterization. We used a step-wise selection step and biologically plausible approaches. P values less than 0.05 were considered statistically significant. Statistical data were analyzed using R Statistical Software [17].

Results

RSV detection and subtype distribution

Infants hospitalized with bronchiolitis were prospectively tested for respiratory viruses over 6 epidemic seasons (October–May 2012-2013 to October–May 2017-2018). After excluding patients with no available data because of refusals of consent or co-infections with other respiratory viruses, of the 283 NPW samples positive for RSV, 218 were successfully sequenced resulting in 144 RSV-A ON1 and 74 RSV-B. RSV-A dominated in all seasons with the exception of 2014-2015 when the number of RSV-B cases were slightly higher (supplementary Figure 1). Of the 144 cases with RSV-A ON1, 139 were suitable to be analyzed in this study.

Sequence analysis

Sequences of the 2nd half of the G gene obtained from NPW samples of the 139 enrolled patients were aligned with BioEdit and identified as ON1 strains for the presence of the insertion. Nineteen sequences were not further analyzed because of poor chromatograms and/or of mixed signals due to co-infections with different ON1 strains. One hundred and twenty study sequences were examined and identical sequences of the same epidemic season were grouped; phylogenetic analysis was performed on a dataset 54 sequences and 4 reference strains.

The nucleotide alignment showed complete identity in the CCD, containing the CX3C motif and the heparin binding site (residues 184 to 198); as expected, many nucleotide substitutions were found in the hypervariable region. As a first measure of ON1 genomic variability, we calculated the average p-distance in each epidemic season and between each epidemic season and the following one (Table 1). Both analysis suggested an increasing genetic divergence of

ON1 strains over the study period. Moreover, using a linear regression model, we proved that the within group p-distances show a significant increasing trend across the epidemic seasons (p-value <0.001).

Phylogenetic reconstruction of study sequences and reference strains in Figure 1 shows that numerous variants of the ON1 genotype appeared since 2012 and clustered into at least three distinct clades. The first clade of strains is derived from sequences identical to the ON1 reference strain [7], now named ON1-1.1 [13]. A sub genotype, named ON1-1.2, early differentiated [8, 13]; study strains similar to ON1-1.2 present in Rome since the 2012-13 epidemic season constitute a second clade (Figure 1). A third clade of sequences originated from ON1-1.2 starting from the 2015-16 epidemic season and continued to diversify (Figure 1). Interestingly, strains of the 2016–17 and 2017-18 seasons, distributed in all clades in the phylogenetic tree, clustered apart from those of the previous epidemic seasons (Figure 1), consistently with their higher genetic divergence.

Analysis of amino acid substitutions and predicted glycosylation sites

In order to appreciate changes in the G protein eventually associated to different phenotypes, ON1 amino acid (aa) sequences of the hypervariable tract (aa 200-321) stratified by clade were aligned with Bioedit v7.1.3 (Figure 2). The changes L274P in the positively selected site 274, and L298P at the corresponding position in the insertion (Figure 2), were distinctive of the ON1-1.2 sub-genotype [8]. The likely third sub-genotype ON1-1.3 was characterized by the distinctive changes I243S and E262K (Figure 2), while retaining L274P and L298P in most isolates. Other changes appeared in the known hypervariable sites [18] or were unique mutations found in a single epidemic season; overall, aa changes were more abundant in the more recent strains. We could identify signature amino acid residues in the most divergent cluster derived from ON1-1.1 (circled in red in Figure 1), that is composed of four unique

sequences found in 11 clinical isolates of the 2017-18 season; the six signature changes (*T200P, P215L, N255D, S275N, V279I, E295V*) are boxed in red in figure 2. T200P is a substitution, never reported before, located in a well characterized antigenic region [19, 20], where the replacement of the conserved polar aa threonine with the hydrophobic proline might have altered the epitope. Substitutions S275N, V279I and E295V that are placed either in the duplicated tract (*S275N,* V279I) or in the insertion (E295V), contribute to make the homologous tracts different from each other. Similarly, in ON1-1.2 and ON1-1.3 recent isolates, several other substitutions render the homologous tracts different, perhaps contributing to antigenic escape [21].

Next, amino acid sequences were submitted to online bioinformatics tools (NetNGlyc and NetOGlyc), in order to predict N- and O-linked glycosylation sites. Substitutions T320I (in 2014-15) and T320A (in 2016-17 and 2017-18) caused the loss of one predicted NetNGlyc site, with respect to all other ON1 strains (dotted boxes in Figure 2). ON1-1.1 strains with higher genetic distances (mostly of the epidemic seasons 2016-17 and 2017-18, Figure 2) acquired around five NetOGlyc sites. Furthermore, strains belonging to the ON1 divergent lineage ON1-1.2 were characterized by the acquisition of around ten NetOGlyc for a total of 40-46 sites with respect to 31-35 predicted sites in the first ON1-1.1 strains.

Analysis of patients' demographic and clinical data

To appreciate any variation in pathogenicity of ON1 isolates, we analyzed demographic and familial data of the 139 enrolled ON1-infected infants (median age of 2.1 months, age range: 0.53-9.2 months, 74 males) hospitalized with bronchiolitis.

Over the epidemic seasons 2012-2013 to 2017-2018, no differences were found when comparing age, sex, ethnicity, gestational age, type of birth, birth weight, familiar atopy and passive smoking exposure among the seasons, except for a significant increase in the

percentage of breastfeed infants during the last three seasons (Table 2). Nonetheless, the number and percentage of patients with severe bronchiolitis differed significantly among seasons, particularly in the last two seasons (Figure 3). The multinomial logistic regression model, accounting for all potentially confounding factors (sex, age, ethnicity, gestational age, type of birth, birth weight, atopic predisposition, smoke exposure, and breastfeeding), showed that the clinical severity score changed according to the epidemic seasons and none of the other factors affected its increase (Table 3).

To further test the effect of ON1 variability, we included in the multivariate analysis the within group p-distances as predictors instead of the epidemic season and obtained overlapping results (data not shown). Moreover, we grouped sequences of the ON1 RM cases in mild (score 0-3), moderate (score 4-5), and severe (score 6-8), and, to each case, we attributed the median p-distance within group value of the proper season (values shown in table 1). Comparing the values of severe cases to those of mild and moderate cases, it resulted that severe cases are associated with significantly higher gene p-distances with respect to mild and moderate cases (p-value= 0.0002, Kruskal-Wallis test). This comparison is illustrated in supplementary Figure 2.

Indeed, individuals in a season with a large within group p-distances are more likely to be classified as severe than those with smaller p-distances.

Discussion

This cohort is the first, to our knowledge, describing RSV-A ON1 variability and analyzing demographic and clinical data of infants hospitalized with bronchiolitis over six epidemic seasons. We identified new ON1 strains and found that the clinical severity of RSV A ON1

bronchiolitis progressively increased from epidemic season 2012-2013 to 2017-2018. In particular, we demonstrated a significant higher clinical severity score in the last two analyzed seasons (2016-2017, 2017-2018), with a higher prevalence of severe forms of bronchiolitis. The hypothesis that these differences had been aff ected by potential clinical confounders was rejected by the demonstration that sex, ethnicity, familiar atopy, passive smoking exposure, breastfeeding, and type of birth did not diff er over time. Taken together our results may suggest that the divergence of ON1 strains was associated with the increase of bronchiolitis clinical severity.

Over the study period, the only RSV-A genotype circulating in Rome was ON1; isolates distributed in three distinct clades and showed increasing genetic divergence from 2012-2013 to 2017-2018 seasons. Interestingly, the third and more recent clade may represent the evolution of a local variant because no identical sequences are present in the NCBI database. Diversification at the local level has been reported and is thought to help ON1 to be maintained in a population over different seasons [10, 12, 13, 21]. The high frequency of amino acid substitution, reported in different RSV genotypes, is typical of the second hypervariable region of the G protein [5]. Nonetheless, over the six epidemic seasons of this study, an increasing number of amino acid substitutions was documented in this tract, particularly in the duplicated tracts and in the C-terminal portion of the G protein that could contribute to immune escape and/or to increased virulence [21]. Recently, an ON1 variant circulating in the Netherlands characterized by a set of eight novel amino acid substitutions and changes in the N-glycosylation sites of the G protein, was associated with more severe infections during the 2016-2017 season [22]. That strain disappeared in the subsequent season and was not detected in Rome, nor by other researchers, up to now; similarly, in this study, divergent strains associated with increased severity were not detected in Rome in the 2018-2019 winter season (data not shown).

In addition to amino acid substitutions in divergent isolates, we showed the acquisition of about ten predicted NetO-Glyc sites that would make recent ON1 strains more similar to the ancestors GA2 and NA1 [7,10,11]. O-linked glycosylation sites are important determinant in the infectious cycle of the enveloped, negative-sense single-stranded RNA viruses and for antigenic recognition [23]; particularly for RSV, positive selection of certain O-glycans in the second hypervariable region of the G protein has been shown in response to immune pressure [19]. In particular, during infections of mucosal layers, viral O-linked glycans have been shown to trigger an innate antiviral immune response, characterized by secretion of the chemokine CXCL10 and subsequent recruitment of neutrophils [24]. CXCL10 is abundant in bronchoalveolar lavages from infants with RSV severe bronchiolitis [25].

Bronchiolitis severity is determined by a complex interaction among viral factors and host immune response [1-3, 6, 26]. Well-known clinical factors affecting severity of bronchiolitis are: passive smoking exposure, lower birth weight, young age and lower weight, lower gestational age, male sex, birth by cesarean section [27-31]. Our analysis demonstrated that none of these risk factors for severe bronchiolitis seemed to influence clinical severity in our patients. In fact, during the epidemic seasons 2012-2013 to 2017-2018, no differences were observed for sex, age at admission, ethnicity, type of birth, birth weight, gestational age, exposure to smoke and family history of atopy. The only relevant difference was an increase in the rate of breastfed infants starting from the epidemic season 2015-2016, but it was not implicated in the variation of clinical severity according to the multinomial logistic regression.

With our data, we were able to describe an increase in the severity of bronchiolitis in the last two epidemic seasons together with the spread of new strains of RSV ON1 with some critical genetic modifications. There were no changes in the hospital setting (nurses and doctors of the emergency department, devices, admission criteria for hospitalization and transfer to

intensive care unit) during the study period and none of the analyzed demographic and clinical characteristics of the study population could explain the increased severity of bronchiolitis. According to these results, we are enough confident to correlate the severity of the disease to the novel ON1 strains. Of course, we could not attribute the increased severity only to a single G protein modification, because differences in disease are not necessarily related to variation in the second half of the G gene (sequence used to classify RSV genotypes) but may well be associated with other genomic tracts. Indeed, other unstudied viral factors may influence the disease course, together with the infants' immune response and other host factors. In this regard, De Vincenzo et al. showed that some adults with low titers of anti-RSV neutralizing antibodies were protected from an intranasal challenge with RSV, while others with high antibody levels were not [32]. Hence, we cannot exclude that an exaggerated immunological protection to RSV, amplified by increasing antibodies production over time in the mothers that were transferred to the babies, may have influenced the clinical course of our patients. Significant progresses in understanding the relationship between strain variability and disease severity will hopefully come from sequencing RSV whole-genome and characterizing the patients' specific immune response.

The strength of our study is that we prospectively enrolled a well-characterized cohort of hospitalized infants younger than 12 months, with a clinical diagnosis of bronchiolitis, without underlying chronic diseases, and with ON1 single infections, thus narrowing the impact of other variables. In fact, most previous studies described ON1 genetic modifications without any relation with clinical parameters [11-13, 21] or examined different acute respiratory illness and patients' age together or for a more limited time period [8-10, 33-35].

The single center nature of this cohort represents a limitation of the study. Nevertheless, we were uniquely able to control the hospital setting and the strict adherence to the inclusion criteria for bronchiolitis diagnosis. Another obvious limitation is the lack of the F gene

sequencing, due to the limited availability of NPW specimens, that would have completed the picture of ON1 variability.

In conclusion, this large study provides important insights into ON1 strains causing hospitalization for bronchiolitis over six epidemic seasons and highlights potential association with clinical severity. These data on RSV variability strengthens the concept that RSV variants may represent a threat and should be monitored at a local level. RSV virological surveillance may contribute to understanding its pathogenicity and to the definition of control strategies in human infections better than previously accomplished in animal or in cell culture models.

References

1. Midulla F, Scagnolari C, Bonci E, et al. Respiratory syncytial virus, human bocavirus and rhinovirus bronchiolitis in infants. Arch Dis Child 2010 ; 95:35-41.
2. Piedimonte G, Perez MK. Respiratory syncytial virus infection and bronchiolitis. Pediatr Rev 2014 ; 35:519-30.
3. Cangiano G, Nenna R, Frassanito A, et al. Bronchiolitis: analysis of 10 consecutive epidemic seasons. Pediatr Pulmonol 2016 ; 51:1330-5.
4. Shi T, McAllister DA, O'Brien KL et al Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. Lancet 2017 ; 390:946-58.
5. Pangesti KNA, Abd El Ghany M, Walsh MG, Kesson AM, Hill-Cawthorne GA. Molecular epidemiology of respiratory syncytial virus. Rev Med Virol 2018 ; 28:e1968.
6. Griffiths C, Drews SJ, Marchant DJ. Respiratory syncytial virus: infection, detection, and new options for prevention and treatment. Clin Microbiol Rev 2016 ; 30:277-319.
7. Eshaghi A, Duvvuri VR, Lai R, et al. Genetic variability of human respiratory syncytial virus A strains circulating in Ontario: a novel genotype with a 72 nucleotide G gene duplication. PLoS One 2012 ; 7:e32807.
8. Pierangeli A, Trotta D, Scagnolari C, et al. Rapid spread of the novel respiratory syncytial virus A ON1 genotype, central Italy, 2011 to 2013. Euro Surveill 2014 ; 19:pii=20843.
9. Prifert C, Streng A, Krempl CD, Liese J, Weissbrich B. Novel respiratory syncytial virus A genotype, Germany, 2011-2012. Emerg Infect Dis 2013 ; 19:1029-30.
10. Otieno JR, Kamau EM, Agoti CN, et al. Spread and evolution of respiratory syncytial virus A genotype ON1, coastal Kenya, 2010-2015. Emerg Infect Dis 2017 ; 23:264-71.

11. Hirano E, Kobayashi M, Tsukagoshi H, et al. Molecular evolution of human respiratory syncytial virus attachment glycoprotein (G) gene of new genotype ON1 and ancestor NA1. Infect Genet Evol 2014 ; 28:183-91.

12. Otieno JR, Kamau EM, Oketch JW, et al. Whole genome analysis of local Kenyan and global sequences unravels the epidemiological and molecular evolutionary dynamics of RSV genotype ON1 strains. Virus Evol 2018 ; 4:vey036.

13. Duvvuri VR, Granados A, Rosenfeld P, Bahl J, Eshaghi A, Gubbay JB. Genetic diversity and evolutionary insights of respiratory syncytial virus A ON1 genotype: global and local transmission dynamics. Sci Rep 2015 ; 5:14268.

14. Midulla F, Nenna R, Scagnolari C, et al. How respiratory syncytial virus genotypes influence the clinical course in infants hospitalized for bronchiolitis. J Infect Dis 2019 ; 219:526-34.

15. Pierangeli A, Gentile M, Di Marco P, et al. Detection and typing by molecular techniques of respiratory viruses in children hospitalized for acute respiratory infection in Rome, Italy. J Med Virol 2007 ; 79:463-8.

16. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 2013 ; 30:2725-9.

17. A Language and Environment for Statistical Computing, R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2017, https://www.R-project.org.

18. Zlateva KT, Lemey P, Vandamme AM, Van Ranst M. Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup A: positively selected sites in the attachment g glycoprotein. J Virol 2004 ; 78:4675–83.

19. Fuentes S, Coyle EM, Beeler J, Golding H, Khurana S. Antigenic Fingerprinting following Primary RSV Infection in Young Children Identifies Novel Antigenic Sites and Reveals Unlinked Evolution of Human Antibody Repertoires to Fusion and Attachment Glycoproteins. PLoS Pathog 2016 ; 12:e1005554.

20. Lee J, Klenow L, Coyle EM, Golding H, Khurana S. Protective antigenic sites in respiratory syncytial virus G attachment protein outside the central conserved and cysteine noose domains. PLoS Pathog 2018 ; 14:e1007262.

21. Bin Lu, Liu H, Tabor DE, et al. Emergence of new antigenic epitopes in the glycoproteins of human respiratory syncytial virus collected from a US surveillance study, 2015-17. Sci Rep 2019 ; 9:3898.

22. Vos LM, Oosterheert JJ, Kuil SD, et al. High epidemic burden of RSV disease coinciding with genetic alterations causing amino acid substitutions in the RSV G-protein during the 2016/2017 season in the Netherlands. J Clin Virol 2019 ; 112:20-6.

23. Ortega V, Stone JA, Contreras EM, Iorio RM, Aguilar HC. Addicted to sugar: roles of glycans in the order Mononegavirales. Glycobiology 2019 ; 29:2-21.

24. Iversen MB, Reinert LS, Thomsen MK, et al. An innate antiviral pathway acting before interferons at epithelial surfaces. Nat Immunol 2016 ; 17:150-8.

25. McNamara PS, Flanagan BF, Hart CA, Smyth RL. Production of chemokines in the lungs of infants with severe respiratory syncytial virus bronchiolitis. J Infect Dis 2005 ; 191:1225-32.

26. Openshaw PJM, Chiu C, Culley FJ, Johansson C. Protective and harmful immunity to RSV infection. Annu Rev Immunol 2017 ; 35:501-32.

27. Jones LL, Hashim A, McKeever T, Cook DG, Britton J, Leonardi-Bee J. Parental and household smoking and the increased risk of bronchitis, bronchiolitis and other lower respiratory infections in infancy: systematic review and meta-analysis. Respir Res 2011 ; 12:5.

28. Semple MG, Taylor-Robinson DC, Lane S, Smyth RL. Household tobacco smoke and admission weight predict severe bronchiolitis in infants independent of deprivation: prospective cohort study. PLoS One 2011 ; 6:e22425.

29. Papoff P, Moretti C, Cangiano G, et al. Incidence and predisposing factors for severe disease in previously healthy term infants experiencing their first episode of bronchiolitis. Acta Paediatr 2011 ; 100:e17-23.

30. Fonseca W, Lukacs NW, Ptaschinski C. Factors affecting the immunity to respiratory syncytial virus: from epigenetics to microbiome. Front Immunol 2018 ; 9:226.

31. Nenna R, Cutrera R, Frassanito A, et al. Modifiable risk factors associated with bronchiolitis. Ther Adv Respir Dis 2017 ; 11:393-401.

32. De Vincenzo JP, Wilkinson T, Vaishnaw A, et al. Viral load drives disease in humans experimentally infected with respiratory syncytial virus. Am J Respir Crit Care Med 2010 ; 182:1305–14.

33. Yoshihara K, Le MN, Okamoto M, et al. Association of RSV-A ON1 genotype with increased pediatric acute lower respiratory tract infection in Vietnam. Sci Rep 2016 ; 6:27856.

34. Streng A, Goettler D, Haerlein M, et al. Spread and clinical severity of respiratory syncytial virus A genotype ON1 in Germany, 2011-2017. BMC Infect Dis 2019 ; 19:613.

35. Calderón A, Pozo F, Calvo C, et al. Genetic variability of respiratory syncytial virus A in hospitalized children in the last five consecutive winter seasons in Central Spain. J Med Virol 2017 ; 89:767-74.

Figure legends

Figure 1. Phylogenetic analysis of the 2nd half of the G gene of the ON1 strains circulating in Rome (November 2012–April 2018).

The dataset included 54 unique ON1 sequences isolated in Rome during the six epidemic seasons and 4 reference strains. RM sequences are identified by year (2 digits), month (2 digits), sample number (2 or 3 digits). If more than three identical strain were found, their total number is indicated after the dot following the strain id. The phylogenetic tree is drawn to scale, and below the tree, scale bar shows the number of substitutions per site. Branches colored in green identify sequences from 2016-2017 and 2017-2018 seasons. The divergent clade derived from ON1-1.1, found in 2017-2018 season, is circled in red. Name and GenBank accession numbers of RSV reference strains are: ON67-1210A (JN257693) for ON1-1.1; 1302-319AN (KC858211) for ON1-1.2; WI/629-Q0284/10 (JF920053) for NA1 and CH57 (AF065258) for GA2.

Figure 2. Alignments of the second hypervariable tract of the G protein sequences of RSVA ON1 with the amino acid (aa) substitutions and the N- and O-linked glycosylation sites.

Alignments of 54 unique RM sequences, from aa position 200 to 321, are shown relative to the ON67-1210A (JN257693) used as the reference strain for ON1-1.1. Dots indicate aa identical to ON1-1.1; aa are colored according to Bioedit v7.1.3 color codes for ease in visual identification of aa polymorphisms. The duplicated tracts (23 of the 24 aa insertion) are evidenced in rectangles. The NTTK N-glyc predicted sites are boxed in light blue: the dotted box is for the site not fully conserved among study strains. Vertical red boxes indicate the six signature substitutions found in the divergent 2017-2018 clade derived from ON1-1.1.

Figure 3. Percentage of patients for each group of clinical severity scores during the epidemic seasons 2012-2013 / 2017-2018. The severity scores (range 0-8) were stratified in mild (score 0-3), moderate (score 4-5), and severe (score 6-8). The p-values for comparison among the epidemic seasons were >0.05 for mild and moderate groups and <0.001 for the severe group (by ANOVA comparison).

Table 1. Between and within group average distance estimation. The average p-distance among nucleotide sequences of the six epidemic seasons was calculated by pairwise comparison using Tamura-Nei model, Gamma Distributed including transitions and transversions, in MEGA 6.

Epidemic seasons	**p-distance within group**	**Epidemic seasons**	**p-distance between groups**
2012-2013	0.012	2012-2013 vs 2013-2014	0.013
2013-2014	0.009	2013-2014 vs 2014-2015	0.014
2014-2015	0.016	2014-2015 vs 2015-2016	0.026
2015-2016	0.014	2015-2016 vs 2016-2017	0.023
2016-2017	0.025	2016-2017 vs 2017-2018	0.035
2017-2018	0.039		

Table 2. Demographic and familial data of 139 RSV A ON1 positive patients hospitalized with bronchiolitis during the epidemic seasons from 2012-2013 to 2017-18 (by ANOVA comparison).

	2012-2013 (N=24)	**2013-2014 (N=13)**	**2014-2015 (N=17)**	**2015-2016 (N=30)**	**2016-2017 (N=22)**	**2017-2018 (N=33)**	**p-value**
Age in months (mean ± sd)	2.44 ± 1.51	1.85 ± 1.21	2.13 ± 1.03	2.37 ± 0.85	2.64 ± 1.87	2.88 ± 1.68	0.08
Male sex N (%)	16/24 (66.7%)	8/13 (61.5%)	5/17 (29.4%)	17/30 (56.7%)	9/22 (40.9%)	19/33 (57.6%)	0.46
Caucasian N (%)	22/24 (91.7%)	11/13 (84.6%)	16/17 (94.1%)	25/30 (83.3%)	20/22 (90.9%)	26/33 (78.8%)	0.23
Gestational age in weeks (mean ± sd)	38.96 ± 1.23	39.15 ± 1.28	39.18 ± 1.24	39.1 ± 1.18	38.64 ± 0.95	38.81 ± 1.26	0.30
Cesarian section N (%)	7/24 (29.2%)	7/13 (53.8%)	9/17 (52.9%)	18/30 (60%)	10/22 (45.5%)	14/30 (46.7%)	0.34
Birthweight in Kg (mean ± sd)	3.229 ± 0.43	3.340 ± 0.42	3.303 ± 0.36	3.233 ± 0.45	3.083 ± 0.37	3.297 ± 0.43	0.75
Familial atopy N (%)	13/24 (54.2%)	5/13 (38.5%)	10/16 (62.5%)	16/28 (57.1%)	16/21 (76.2%)	23/32 (71.9%)	0.25
Smoke exposure N (%)	12/24 (50%)	8/13 (61.5%)	9/17 (52.9%)	17/27 (63%)	10/21 (47.6%)	9/32 (28.1%)	0.08
Breast feeding N (%)	9/24 (37.5%)	5/13 (38.5%)	7/17 (41.2%)	20/30 (66.7%)	15/22 (68.2%)	22/30 (73.3%)	0.001

Table 3. Results of the multinomial logistic model adjusted for the potentially confounding factors (sex, age, ethnicity, gestational age, type of birth, birth weight, familial atopy, smoke exposure, and breastfeeding) evaluating the clinical severity score of RSV A ON1 positive patients hospitalized with bronchiolitis during the epidemic seasons from 2012-13 to 2017-2018. The first epidemic season, 2012-2013 is set as corner point in the ANOVA parameterization.

Epidemic seasons	**Estimate**	**OR**	**Confidence Interval**	**P-value**
2013-2014	-0.452	0.636	0.154 – 2.535	0.522
2014-2015	0.653	1.922	0.565 – 6.657	0.297
2015-2016	0.363	1.438	0.469 – 4.457	0.526
2016-2017	2.071	7.932	2.522 – 26.185	0.0005
2017-2018	1.899	6.672	2.097 – 22.161	0.0001

Figure 1

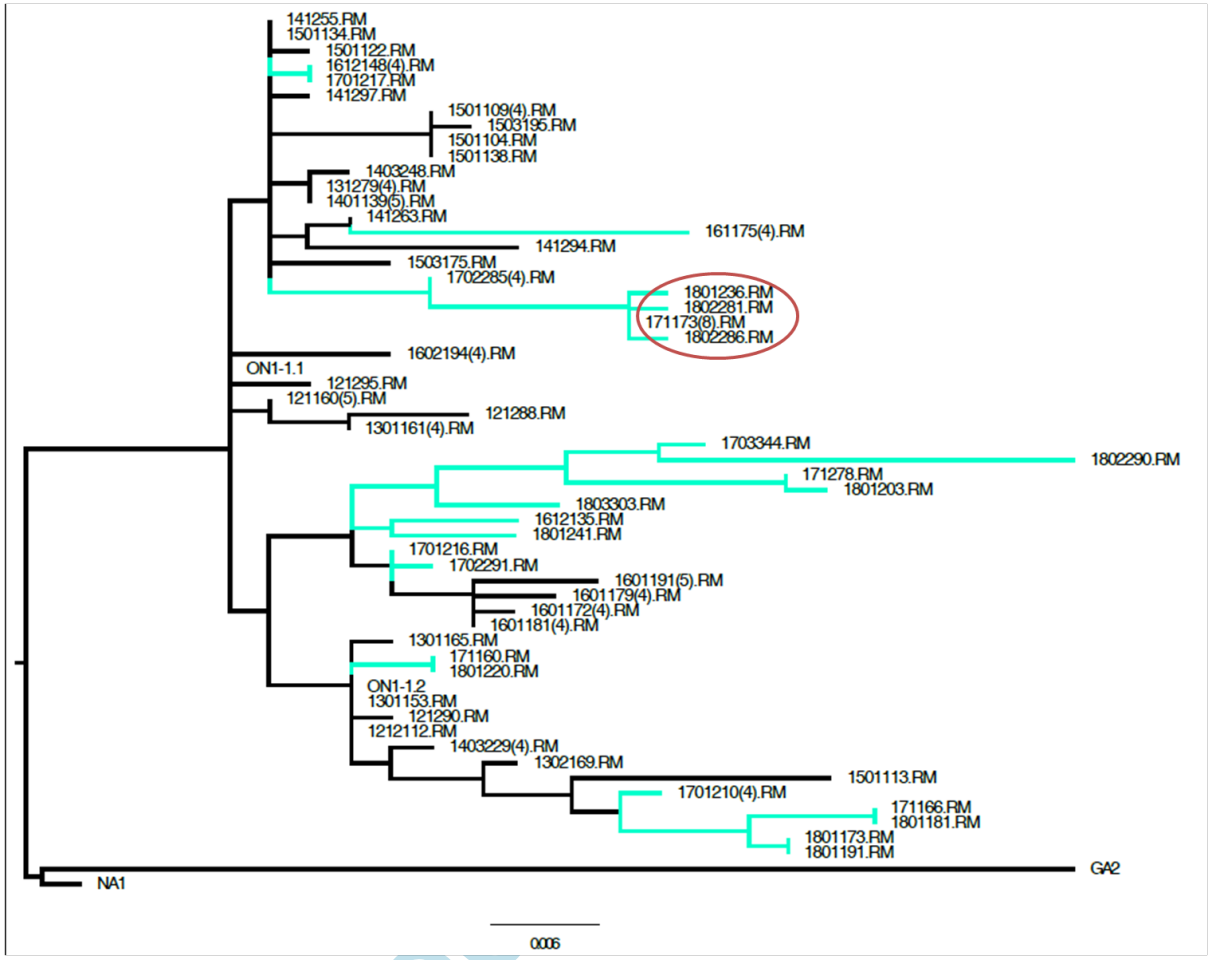

Figure 2

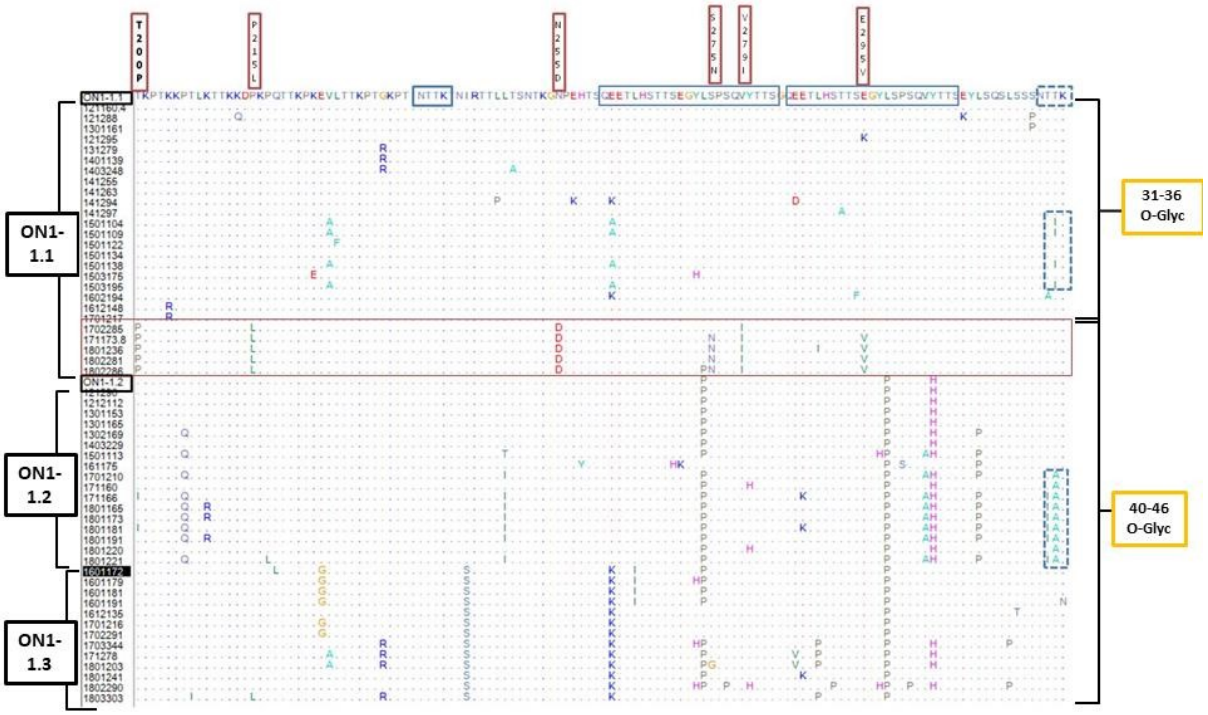

Figure 3