TITLE: Molecular evidence for SARS-CoV-2 in samples collected from patients with morbilliform eruptions since late summer 2019 in Lombardy, Northern Italy

AUTHORS

Antonella Amendola, Ph.D.[1,2,*], Marta Canuti, Ph.D.[3,*], Silvia Bianchi, Ph.D.[2,4,*], Sudhir Kumar, Ph.D.[5,6,7], Clara Fappani, M.Sc[1], Maria Gori, Ph.D.[2], Daniela Colzani, M.Sc.[1], Sergei L. Kosakovsky Pond, Ph.D.[5,6], Sayaka Miura, Ph.D.[5,6], Melissa Baggieri, Ph.D.[8], Antonella Marchi MD[8], Elisa Borghi, Ph.D.[2,4], Gian Vincenzo Zuccotti, MD[9,10,#], Mario C. Raviglione, MD[11,#], Fabio Magurano, Ph.D.[8,#], Elisabetta Tanzi, Ph.D.[1,2,#]

*These authors contributed equally to this work

#These authors are joint senior authors

AFFILIATIONS

[1]Department of Biomedical Sciences for Health, University of Milan, Milan, Italy
[2]Coordinate Research Centre EpiSoMI, University of Milan, Milan, Italy
[3]Biology Department, Memorial University of Newfoundland, St John's, NL, Canada
[4]Department of Health Sciences, University of Milan, Milan, Italy
[5]Institute for Genomics and Evolutionary Medicine, Temple University, Philadelphia, PA, USA
[6]Department of Biology, Temple University, Philadelphia, PA, USA
[7]Center for Excellence in Genome Medicine and Research, King Abdulaziz University, Jeddah, Saudi Arabia
[8]Department of Infectious Diseases, Istituto Superiore di Sanità, Rome, Italy
[9]Department of Paediatrics, Children Hospital V. Buzzi, University of Milan, Milan, Italy
[10] Romeo and Enrica Invernizzi Pediatric Research Center, University of Milan, Milan, Italy
[11]Centre for Multidisciplinary Research in Health Science, University of Milan, Milan, Italy

CORRESPONDING AUTHOR

Elisabetta Tanzi,
Department of Biomedical Sciences for Health
University of Milan
Via Pascal 36, 20133 Milan, Italy
Tel: +39-0250315139 - Fax: +39-0250315120

ABSTRACT

BACKGROUND. Skin manifestations have been reported in patients with SARS-CoV-2 infection. As a reference laboratory for measles and rubella surveillance in Lombardy, we evaluated the association between SARS-CoV-2 infection and measles-like symptomatology while providing evidence for undetected early circulation of SARS-CoV-2.

METHODS. Overall, 435 samples from 156 cases were investigated. RNA isolated from oropharyngeal swabs (N=148) and urine (N=141) was screened by hemi-nested PCRs targeting key sites for viral typing. After Sanger sequencing, detected mutations were used to estimate the time of emergence of the progenitor SARS-CoV-2 using a previously calibrated molecular clock. Sera (N=146) were tested for anti-SARS-CoV-2 IgG, IgM, and IgA, and the titre of neutralizing antibodies was assessed.

FINDINGS. Molecular evidence for SARS-CoV-2 infection was found in 13 subjects. Two patients were from the pandemic period (2/12, 16·7%, March 2020-March 2021) and 11 were from the pre-pandemic period (11/44, 25%, August 2019-February 2020). Five of the positive individuals showed the simultaneous presence of anti-SARS-CoV-2 antibodies. No clear evidence of infection was found in 281 samples collected between August 2018 and July 2019 from 100 patients. The first positivity for SARS-CoV-2 RNA was found in a sample collected on September 12, 2019. Mutations typical of B.1 (PANGOLIN classification) strains, previously reported to have emerged in January 2020, had already been circulating in October 2019. Hence, we estimate SARS-CoV-2 progenitor of known human infections to have emerged in late June-late August 2019.

INTERPRETATIONS. We find evidence that SARS-CoV-2 was circulating in Lombardy during the late summer of 2019. This finding highlights the importance of retrospective surveillance studies to understand the early dynamics of COVID-19 spread and improve national-level preparedness.

FUNDING. Romeo and Enrica Invernizzi Pediatric Research Center; U.S. National Science Foundation; U.S. National Institutes of Health.

RESEARCH IN CONTEXT

EVIDENCE BEFORE THIS STUDY

We searched PubMed and Google Scholar for articles published up to May 2021 showing early evidence of SARS-CoV-2 circulation in non-Asian countries. The research showed that the earliest molecular and serological evidence reported SARS-CoV-2 circulation in Southern Europe in November-December 2019, including our previous report of a SARS-CoV-2 sequence in the oropharyngeal swab from a child from Milan (Lombardy) with suspected measles. Additionally, our literature research focused on reported cases of SARS-CoV-2 with skin manifestation. This was conducted without language restrictions using the search terms “COVID-19”, “SARS-CoV-2” plus “skin manifestations” and synonyms.

ADDED VALUE OF THIS STUDY

Our study provides strong evidence that SARS-CoV-2 was already circulating in Northern Italy by late summer of 2019 and shows a clear association between measles-like cutaneous manifestations and SARS-CoV-2.

IMPLICATIONS OF ALL THE AVAILABLE EVIDENCE

The spread of SARS-CoV-2 in Southern Europe in the last quarter of 2019 suggests that a wider geographical area and a broader timespan should be considered during virus origin investigations.

Future retrospective studies are crucial to seek further early SARS-CoV-2 cases as these are essential to more accurately identify the time and location of viral emergence.

INTRODUCTION

The novel betacoronavirus Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) causes the coronavirus disease 2019 (COVID-19) and it was first recognized in China in late December 2019[1]. Molecular evidence suggests that SARS-CoV-2 emerged as a capable human pathogen, likely from a bat reservoir, although the existence of intermediate hosts that allowed the spillover cannot be ruled out. Although the timeline of SARS-CoV-2 emergence has not yet been resolved, evolutionary analyses predicted that the virus likely circulated in China for some time, even months, before the first recorded December outbreak in the Hubei province[1–4].

Shortly after the first identification of SARS-CoV-2, cases were recorded in non-Asian countries with early reports in mid-late January 2020 in the western regions of North America and Europe[5]. Italy was the first European country that reported sustained community transmission of SARS-CoV-2 following the first identification of a non-travel related case in Codogno, Lombardy, on February 20, 2020. Italy quickly became the epicenter of the European epidemic, with Lombardy being the most affected area[6–8]. The viral strain that dominated the Lombardy outbreak, and that subsequently spread across Europe and beyond, is distinguished from reference strains reported from China in 2019 by the simultaneous presence of several mutations, such as C3037T, C14408T, and A23403G (resulting in the Spike protein amino acid variant D614G)[9,10]. This strain is now classified as 20A in NextStrain[10] and B.1 in Pangolin[9], and it has an αβ mutational signature[4].

Several lines of evidence suggest that SARS-CoV-2 had already been spreading unnoticed within Europe for several weeks before being recognized. In fact, SARS-CoV-2 RNA was detected in wastewaters in Northern Italy in late December 2019[11], and there is indirect evidence for 2019 infections through antibody detection in France[12]. In addition, SARS-CoV-2 RNA was detected as early as December 2019 in a respiratory sample from a French patient hospitalized for haemoptysis[13] and in the oropharyngeal swab from a child from Milan (Lombardy) with suspected measles[14], while viral antigens and RNA were detected in paraffin-preserved skin biopsies of a woman with dermatosis prepared in November 2019 in Milan[15].

A variety of skin manifestations have been reported in patients with SARS-CoV-2 infection, especially by Lombard dermatologists[16]. These manifestations, which pose diagnostic difficulties in the context of COVID-19[17], appear at any disease stage, are sometimes characterized by late-onset, and have variable duration, severity, and prognosis[17–19]. Skin manifestations can also appear at a late stage of infection or even in the absence of respiratory symptoms[18], posing a diagnostic challenge since respiratory swabs from affected patients can be RNA-negative when cutaneous manifestations appear[19,20].

As a MoRoNet-accredited and WHO-accredited Subnational Reference Laboratory for measles and rubella surveillance operating in Lombardy, we receive every year oropharyngeal swabs and/or urine from patients that present with morbilliform rash to perform molecular and serological diagnostics[21]. We monitor Milan and its surroundings, a densely populated area with about 4 million inhabitants, as well as other Lombard provinces (i.e., Brescia, Varese, Como, and Monza-Brianza). Besides diagnostics, we evaluate every year several performance indicators, including the rate of discarded non-measles and non-rubella cases, expressed as the number of non-measles/non-rubella cases in a year divided by the average population in the study area. In 2019,

this rate was over two times higher than the average of the previous two years (3·00 per 100,000 vs. 1·25 per 100,000)[22] and, since the summer of 2019, the percentage of suspected measles and rubella cases that tested negative steadily increased from about 30% to 70% at the start of the first pandemic wave, and to 100% in 2021 (Figure 1).

These observations prompted us to further investigate whether SARS-CoV-2 infection could be involved in morbilliform clinical manifestations, thus explaining an epidemiological trend that began several months before the first known COVID-19 cases had been observed in Italy. This study expands on our previous report[14] and describes the results of the surveillance conducted on samples collected over two and a half years, beginning more than a year prior to the beginning of the COVID-19 pandemic. We simultaneously investigate the association between SARS-CoV-2 and measles-like symptomatology and provide further evidence for an undetected early circulation of SARS-CoV-2 in Lombardy. For this purpose, all samples from measles/rubella negative patients that our laboratory received within the measles and rubella surveillance framework since August 2018 were tested for evidence of SARS-CoV-2 infection. These samples were all collected when the first tentative diagnosis of measles or rubella was made, which is when the skin manifestations became evident.

MATERIALS AND METHODS

Specimens

Samples from cases with morbilliform rashes that received negative measles and rubella diagnostics (discarded cases) collected between the summer of 2019, when the increase in discarded measles-rubella cases was observed, and the beginning of the pandemic were the main focus of this survey. Additionally, samples from all discarded cases collected during the prior 12

months were included as control, and those received during the pandemic period were included for comparison.

Overall, 435 samples (148 oropharyngeal swabs, 141 urine samples, and 146 serum samples) collected from 156 cases were investigated for SARS-CoV-2 infection. These included samples from 44 cases collected between August 2019 and February 2020 (pre-pandemic cases), 12 cases between March 2020 and March 2021 (pandemic cases), and 100 cases between August 2018 and July 2019 (control cohort). The paucity of samples received during the pandemic period was due to lockdown measures and the slowdown in measles surveillance activities. Since all samples were received from hospitals, patients from the pandemic period had already tested negative for SARS-CoV-2. Samples included oropharyngeal swabs (Copan, Copan Italia SPA, Italy) submerged in 3 ml universal transport medium (UTM), urine, and sera, all stored at -80°C in the biobank of the Laboratory at the University of Milan.

This work was carried out as part of the Integrated Measles-Rubella Surveillance, performed by law in accordance with the Prime Minister's Decree of 3 March 2017 (https://www.gazzettaufficiale.it/eli/id/2017/05/12/17A03142/sg). The retrospective study was conducted in the absence of consent following the declaration of state of emergency caused by the SARS-CoV-2 pandemic, in accordance with letter i, article 9 of the GDPR (https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=OJ:L:2016:119:FULL&from=FI).

Molecular testing and sequence analysis

Two hundred and eighty-nine samples (148 oropharyngeal swabs, 141 urine samples) were investigated for SARS-CoV-2 RNA. When not already available from previous surveillance activities, RNA was freshly isolated. RNA isolations were performed with the NucliSENS®

easyMAG™ automated system (bioMérieux bv, Lyon, France) from 1ml of UTM and/or using pellet obtained from 5-15 ml of urine as input, and cDNA was synthesized using SuperScript™ II Reverse Transcriptase (ThermoFisher, Waltham, Massachusetts) with the reverse primer (final concentration of 20 µM), according to the manufacturer's instructions. Samples were screened with four different hemi-nested PCRs designed to amplify genomic fragments within the non-structural protein 3 (NsP3), the RNA-dependent RNA-polymerase (RdRp), and the spike protein (two different PCRs) spanning regions that are known to contain some of the key mutations and the furin cleavage site (Table 1). PCRs were performed with the DreamTaq DNA polymerase (ThermoFisher, Waltham, Massachusetts) using 5 µl of cDNA with PCR conditions outlined in Table 1. Amplicons from positive samples were purified with the NucleoSpin Gel and PCR Clean-Up kit (Macherey-Nagel GmbH & Co. KG, Germany) and outsourced for Sanger sequencing. Isolated RNA was also tested for SARS-CoV-2 by Real-Time PCR according to the diagnostic protocol of the CDC[23].

Obtained sequences (accession numbers MZ223385-MZ223398) were compared to the reference strain Wuhan-Hu-1 (accession number NC_045512.2) to detect mutations. Identified mutations were evaluated using a recently developed mutation order analysis approach, and the time of emergence of the SARS-CoV-2 progenitor (proCoV2)[4] was calculated using a previously calibrated molecular clock[24].

Serological testing

For serological investigations, 146 sera (98 control cases, 38 pre-pandemic cases, and 10 pandemic cases) were tested for anti-SARS-CoV-2 IgG, IgM, and IgA using the semi-quantitative Anti-SARS-CoV-2 ELISA (Euroimmun, Lübeck, Germany) tests, according to manufacturer's instructions.

Positive sera were sent to the Italian National Institute of Health (ISS) to confirm the data and quantify the titer of neutralizing antibodies by SARS-CoV-2 plaque reduction neutralization test (serum dilution causing 80% plaque reduction, PRNT80) conducted, as previously described[25], under biosafety level 3 facilities at ISS (Rome, IT).

Statistical analyses

Differences between positivity rates in different groups were evaluated for statistical significance using the Mid-p exact test with OpenEpi (www.OpenEpi.com), with p-values ≤ 0·05 (two-tailed tests) considered significant.

Role of the funding source

Funders have played no role in the research.

RESULTS

SARS-CoV-2 infection detection

Molecular evidence for SARS-CoV-2 infection was found in 13 subjects (Table 2), with a positivity rate of 16·7% (2/12) for the pandemic cases and 25% (11/44) for the pre-pandemic cases (Table 3). Noteworthy, none of the 191 swab and urine samples collected before September 2019 showed any positivity. For six of the 13 patients (46·2%) viral RNA was found in urine, while for the remaining seven patients (53·8%) the virus was found in respiratory material. The virus was found only in the urine of patients from the pandemic period (2021), confirming the negative SARS-CoV-2 diagnosis performed at the hospital using respiratory samples. Samples from nine of the 11 patients of the pre-pandemic period were all collected in 2019. In the considered period, the very first sample that tested positive for SARS-CoV-2 RNA was a urine sample collected as

early as September 12th, 2019, from an 8-months old child whose serum was also IgG and IgM positive (Table 2). Of note, all samples from positive patients only became positive after nested PCRs and none of the samples tested positive with the Real-Time PCR diagnostic protocol, indicative of low viral load at the detection limit.

Four of the nine patients from the pre-pandemic period, tested positive for anti-SARS-CoV-2 antibodies, with IgM being the most frequently detected antibody class. However, only one of these sera revealed partially neutralizing antibodies (causing 62% plaque reduction) (Table 2). Additionally, serological positivity in the absence of viral detection was observed for four subjects that were sick during the pandemic period and presented IgA in combination with at least another class of antibody (Table 2). Interestingly, independently from the date of sample collection, we detected antibodies in all six positive children aged one year or younger, while only three out of nine older patients resulted positive in ELISA (Table 2).

One serum sample collected at the end of August 2019, and five serum samples collected in July 2019 showed IgA positivity alone. Since no molecular evidence for SARS-CoV-2 was found, and the neutralization assay could not confirm the results (only partial neutralization was observed for a few samples: 50-59% plaque reduction), we concluded that the evidence of infection was too weak and considered the diagnosis for these patients inconclusive. A similar determination was made for two other patients that presented symptoms during the pre-pandemic period. None of the 98 sera from the control group tested positive for IgG or IgM (Table 3). Remarkably, results obtained for pre-pandemic and pandemic cases were comparable, and each patient presented a unique pattern of positivity, with none showing a positive result for every test performed (Table 2).

None of the positive patients reported any history of travel in the prior two weeks the onset of rash. The first pre-pandemic cases were mainly localized East of Milan and Brescia (September-October, 2019), while later cases were identified in the area North-West of Milan (November-December, 2019). Interestingly, no cases were reported from Como, Monza-Brianza, and Varese, none of which were particularly affected by COVID-19 during the first wave[7,8]. Five out of 11 pre-pandemic cases occurred in the period October 12–

23, 2019. Pandemic cases were mostly localized in the province of Milan and were all detected in the first months of 2021.

Sequence analysis

In total, we obtained 15 sequences, including 12 from pre-pandemic cases (Table 2). All the three major variants (C3037T, C14408T, and A23403G), which were first detected weeks after the outbreak in China, were observed whenever these regions were sequenced, indicating that sequences from October 2019 already carried variants that were absent in the first sampled strains (e.g., Wuhan HU-1) reported from China (Supplementary Table S1, Figure 2). Six out of the seven partial S sequences (fragment B) were 100% identical to the reference sequence Wuhan HU-1.

We found no mutations within the furin cleavage site, and one sequence (case ID #5) contained two additional non-synonymous mutations in NsP3, unique in our dataset: T2987C (F908L) and T3012C (L916S), whose significance is unclear. Interestingly, evidence for a possible co-infection with multiple strains was noted in one patient (case ID #10) as two double peaks were observable in the electropherograms of the S sequences: 23000C/T and 23222G/A (Supplementary Figure S1).

The three common mutations (C3037T, C14408T, and A23403G) belong to the β group of mutations[4] previously not observed in samples collected before late January 2020. All globally circulating genomes with β mutations also carry three α mutations (T18060C, T8782C, and C28144T), inferred to precede them temporally; however, these mutations lie outside the fragments sequenced in this study. Nonetheless, the mutation tree predicts that these mutations are likely to be present in the pre-pandemic strains, implying that these strains are at least six mutations away from the inferred SARS-CoV-2 progenitor strain and belong to the αβ lineage, which has produced almost all the major offspring strains circulating today (Figure 2).

Since all three β mutations were already detected by October 22, 2019, we project that the progenitor of SARS-CoV-2 already existed 11·6 – 16·2 weeks earlier than October 22, i.e., in late June 2019 to late August 2019, using a simple extrapolation based on six mutations and the current range of mutation rate estimates[4,24].

DISCUSSION

There is a clear association between morbilliform eruptions and SARS-CoV-2 infection[18], although diagnostics remains challenging given the non-specific clinical spectrum, the inconsistencies in laboratory results, and the lack of an optimal time for testing. While antibody testing might not provide a definite diagnosis, especially in cases of sustained viral spread, viral detection is complicated by lower viral loads. In fact, all samples that we identified as SARS-CoV-2-positive (pre-pandemic and pandemic cases) were only positive after two rounds of amplification and were consistently negative in standard Real-Time PCR-based diagnostic protocols. There was concordance between our data and the diagnostics performed at the hospital as none of the respiratory samples obtained during the pandemic period tested positive, while SARS-CoV-2 RNA

was identified only in urine samples collected from two patients during this period. Additionally, each patient showed a unique pattern of positivity.

The identification of viral RNA in urine indicates that a systemic infection occurs concurrently with the development of the morbilliform skin rash and the low detectability in respiratory samples, likely a consequence of low RNA load, could reflect a shifted appearance of the skin rash in respect to respiratory symptoms. Furthermore, our analysis of the pandemic samples might be biased towards cases that are "difficult to detect" as in this period we only received samples from patients whose respiratory sample tested negative for SARS-CoV-2 RNA using RT-PCR, which is the standard diagnostic method, and were, therefore, already characterized by a viral load below the limit of detection. Finally, since some of the cases identified during the pandemic period declared close contacts with COVID-19 confirmed cases during the days and weeks before symptom onset, epidemiological investigations remain crucial in helping to diagnose these cases, and future studies should further elucidate the relationship between SARS-CoV-2 and morbilliform skin rash.

Following the first demonstration of early SARS-CoV-2 circulation in Northern Italy[14] and the detection of viral RNA in urine samples from patients with morbilliform eruptions in 2021, we decided to screen all available urine and oropharyngeal swab samples collected from measles-negative patients that were submitted to our laboratory since August 2018. Our results provide strong evidence that SARS-CoV-2 was already circulating in Northern Italy by late summer of 2019, with the first molecular evidence of infection dating September 12th, 2019, and no PCR-positive result for any of the 191 samples collected before this date. Importantly, all identified cases, including the nine from 2019, were not travel-related.

Results obtained with samples from the pre-pandemic period (August 2019-February 2020) were analogous to those obtained with samples collected in 2021, further corroborating our hypothesis of the involvement of SARS-CoV-2 in morbilliform eruptions, which can, at least partially, explain the increase in the rate of discarded non-measles and non-rubella cases observed since late 2019. Additionally, we observed sequence variability and detected multiple variants circulating during the pre-pandemic period, as reported in China between December 2019 and January 2020[26]. These results confirm recent computational findings that several SARS-CoV-2 lineages had been spreading worldwide at least for several weeks before the first reported COVID-19 cases[4].

Partial sequence analysis showed that β mutations (diagnostic of lineage 20A and B.1) were already present in strains from the last quarter of 2019, implying that a lineage of the coronavirus substantially different from the putative progenitor was already circulating in Northern Italy at that time, pushing back the predicted date of the progenitor SARS-CoV-2 to the range between late June 2019 and late August 2019. This finding might seem surprising as these mutations have so far been only identified in strains from 2020. However, the pool of sequences from 2019 is very small, and genomes carrying these mutations may have simply gone undetected thus far. A recent analysis of recovered sequences from thirteen early Chinese isolates showed how the addition of even a few early isolates can significantly shift the relatively likelihoods for the most likely location of the viral outbreak[27]. Finally, although it is possible that viruses carrying β mutations were imported into Europe, it is also conceivable that these mutations have evolved in parallel outside China as the virus was circulating in other geographical areas, as also already previously hypothesized[4].

Since strains with an αβ mutational signature have produced all major lineages and no other high-frequency mutations have so far been observed, we could hypothesize that a strain capable of

efficient human-to-human transmission had already been circulating in Northern Italy by September 2019. Epidemiological stochasticity and the non-linear relationship between incidence and mortality and the time of community circulation commencement[28] could explain the delay between virus introduction in Lombardy (no later than September 2019) and the observed increase in mortality and hospitalizations (March 2020) (www.istat.it, report of the 5th of March, 2021). However, an alternative hypothesis of circulation of a somewhat genetically different virus with reduced transmissibility and/or virulence could not be excluded. This could also explain why we observed a lower degree of neutralization for sera from 2019 than those collected in 2021, possibly indicating a weaker virus-antibody bound[29]. However, it has to be stressed that immunological data suggest that the adaptive immune response against SARS-CoV-2 could be weak, especially if the patient experienced mild symptomatology, that antibodies can quickly disappear following recovery, and that a not insignificant proportion of patients does not develop neutralizing antibodies or develops them only for a short time frame[30]. Indeed, serological results alone could not be used as proof of early viral circulation as serological evidence for infection for pre-pandemic cases was weak. In our study, immunological markers were used to integrate the molecular investigation that, combined with sequencing, allowed us to draw much stronger conclusions.

Our study retrospectively tested samples that were collected for measles/rubella surveillance and not for COVID-19 diagnostics. Thus, these samples may have been suboptimal for SARS-CoV-2 detection as they were collected when the skin rash manifested, i.e., when the viral load in the respiratory tract might have been low. Although our study demonstrates a previously undetected early SARS-CoV-2 circulation, stronger and more conclusive answers could be obtained by retrospectively investigating samples collected during acute respiratory infections, such as those collected within the framework of surveillance systems for respiratory viruses. Additionally, to

definitively explain why the observed clinical cases predate the currently predicted emergence date[24], based on molecular clock analyses, of both SARS-CoV-2 and clade 20A strains, and to definitively clarify whether previously undetected mutations were present in initial strains, complete genome sequencing of early strains is pivotal.

Our results have relevant implications for the global effort to clarify the chain of events that lead to the emergence of SARS-CoV-2 in the human population and suggest that a wider geographical area and a broader timespan should be considered during virus origin investigations. The results presented here bear no relevance to the lab leak versus natural origin debate. Finally, we would like to encourage other groups to perform additional retrospective studies using samples stored in biobanks to seek further early SARS-CoV-2 cases as these are essential to more accurately identify the time and location of viral emergence. Furthermore, since we showed that Real-Time PCR might not be the preferred method for screening archived samples or detecting the virus in sub-optimal samples and cases of atypical clinical patterns, we suggest using a combination of multiple nested-PCRs targeting several different fragments since this was the approach that allowed us to identify a larger number of positives. Our study highlights the crucial role of surveillance systems in managing epidemics at their origin and as a tool to investigate early stages of pathogen transmission retrospectively.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Marino Faccini and the staff of the local health authorities of Milan and Brescia involved in the outbreak investigations and Mrs. Paola Bucci of the Istituto Superiore di Sanità (ISS) for technical support.

This study is dedicated to the memory of Dr. Raffaele Gianotti, to whom we will be forever grateful for improving our knowledge of COVID-19-related skin manifestations, for appreciating our work, and for encouraging us to continue our research.

FUNDING

Research support from: Romeo and Enrica Invernizzi Pediatric Research Center, University of Milan, Milan, Italy; U.S. National Science Foundation to S.K. and S.M. (DEB-2034228) and S.P. (DBI-2027196); U.S. National Institutes of Health to S.K. (GM-139504) and S.P. (AI-134384).

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

AUTHOR CONTRIBUTION

Conceptualization: AA, MC, SB, ET; data curation: AA, SB, ET; formal analysis: MC, SB, SK, CF, MG, SLKP,SM; funding acquisition: AA, ET, GVZ; investigation: CF, MG, DC; methodology: SB; project administration: AA, ET, MCR; resources: MB, AM, EB, FM; supervision: AA, ET; visualization: MC, SB, SK, SLKP, CF, SM; writing – original draft: AA, MC, SB, SK, ET and writing – review & editing: all the authors.

DATA SHARING

Data collected for the study (de-identified patient and clinical data) will be made available to others upon written request to be sent by email to elisabetta.tanzi@unimi.it. These data will be shared with researchers who provide a methodologically sound proposal after approval by the institutional review board. Data will be available beginning 1 month after the publication of the study.

References

1 Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. *Nat Med* 2020; **26**: 450–2.

2 Leitner T, Kumar S. Where Did SARS-CoV-2 Come From? *Mol Biol Evol* 2020; **37**: 2463–4.

3 MacLean OA, Lytras S, Weaver S, *et al.* Natural selection in the evolution of SARS-CoV-2 in bats created a generalist virus and highly capable human pathogen. *PLOS Biology* 2021; **19**: e3001115.

4 Kumar S, Tao Q, Weaver S, *et al.* An evolutionary portrait of the progenitor SARS-CoV-2 and its dominant offshoots in COVID-19 pandemic. *Mol Biol Evol* 2021; msab118

5 Worobey M, Pekar J, Larsen BB, *et al.* The emergence of SARS-CoV-2 in Europe and North America. *Science* 2020; **370**: 564–70.

6 Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. *JAMA* 2020; **323**: 1545.

7 Carletti M, Pancrazi R. Geographic Negative Correlation of Estimated Incidence between First and Second Waves of Coronavirus Disease 2019 (COVID-19) in Italy. *Mathematics* 2021; **9**: 133.

8 Tosi D, Chiappa M. Understanding the Geographical Spread of COVID-19 in Relation with Goods Regional Routes and Governmental Decrees: The Lombardy Region Case Study. *SN Comput Sci* 2021; **2**: 203.

9 Rambaut A, Holmes EC, O'Toole Á, *et al.* A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 2020; **5**: 1403–7.

10 Hadfield J, Megill C, Bell SM, *et al.* Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* 2018; **34**: 4121–3.

11 La Rosa G, Mancini P, Bonanno Ferraro G, *et al.* SARS-CoV-2 has been circulating in northern Italy since December 2019: Evidence from environmental monitoring. *Sci Total Environ* 2021; **750**: 141711.

12 Carrat F, Figoni J, Henny J, *et al.* Evidence of early circulation of SARS-CoV-2 in France: findings from the population-based 'CONSTANCES' cohort. *Eur J Epidemiol* 2021; **36**: 219-22.

13 Deslandes A, Berti V, Tandjaoui-Lambotte Y, *et al.* SARS-CoV-2 was already spreading in France in late December 2019. *Int J Antimicrob Agents* 2020; **55**: 106006.

14 Amendola A, Bianchi S, Gori M, *et al.* Evidence of SARS-CoV-2 RNA in an Oropharyngeal Swab Specimen, Milan, Italy, Early December 2019. *Emerg Infect Dis* 2021; **27**: 648–50.

15 Gianotti R, Barberis M, Fellegara G, Galván-Casas C, Gianotti E. COVID-19-related dermatosis in November 2019: could this case be Italy's patient zero? *Br J Dermatol* 2021; **184**: 970-1.

16 Recalcati S. Cutaneous manifestations in COVID-19: a first perspective. *J Eur Acad Dermatol Venereol* 2020; **34**: e212–3.

17 Freeman EE, McMahon DE, Hruza GJ, *et al.* Timing of PCR and antibody testing in patients with COVID-19-associated dermatologic manifestations. *J Am Acad Dermatol* 2021; **84**: 505–7.

18 Visconti A, Bataille V, Rossi N, *et al.* Diagnostic value of cutaneous manifestation of SARS-CoV-2 infection. *Br J Dermatol* 2021; **184**: 880–7.

19 Giavedoni P, Podlipnik S, Pericàs JM, *et al.* Skin Manifestations in COVID-19: Prevalence and Relationship with Disease Severity. *J Clin Med* 2020; **9**: 3261.

20 Le Cleach L, Dousset L, Assier H, *et al.* Most chilblains observed during the COVID-19 outbreak occur in patients who are negative for COVID-19 on polymerase chain reaction and serology testing. *Br J Dermatol* 2020; **183**: 866–74.

21 Magurano F, Baggieri M, Bucci P, *et al.* MoRoNet a network to strengthen the quality of measles and rubella surveillance in Italy. *Eur J Public Health* 2020; **30**: ckaa166.1336.

22 Bianchi S, Faccini M, Lamberti A, *et al.* Measles surveillance activities in the Metropolitan Area of Milan during 2017-2018. *J Prev Med Hyg* 2019; **60**: E286–92.

23 CDC. Research Use Only 2019-Novel Coronavirus (2019-nCoV) Real-time RT-PCR Primers and Probes. https://www.cdc.gov/coronavirus/2019-ncov/lab/rt-pcr-panel-primer-probes.html (accessed July 5, 2021)

24 Pekar J, Worobey M, Moshiri N, Scheffler K, Wertheim JO. Timing the SARS-CoV-2 index case in Hubei province. *Science* 2021; **372**: 412–7.

25 Magurano F, Baggieri M, Marchi A, Rezza G, Nicoletti L, COVID-19 Study Group. SARS-CoV-2 infection: the environmental endurance of the virus can be influenced by the increase of temperature. *Clin Microbiol Infect* 2021; **27**: 289.e5-289.e7.

26 WHO-convened global study of origins of SARS-CoV-2: China Part. https://www.who.int/publications-detail-redirect/who-convened-global-study-of-origins-of-sars-cov-2-china-part (accessed May 13, 2021).

27 Bloom JD. Recovery of deleted deep sequencing data sheds more light on the early Wuhan SARS-CoV-2 epidemic. *bioRxiv* 2021; 2021.06.18.449051.

28 Ylli A, Wu YY, Burazeri G, Pirkle C, Sentell T. The lower COVID-19 related mortality and incidence rates in Eastern European countries are associated with delayed start of community circulation. *PLOS ONE* 2020; **15**: e0243411.

29 Chen RE, Zhang X, Case JB, *et al.* Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat Med* 2021; **27**: 717–26.

30 Chia WN, Zhu F, Ong SWX, *et al.* Dynamics of SARS-CoV-2 neutralising antibody responses and duration of immunity: a longitudinal study. *Lancet Microbe* 2021; **2**: e240–9.

Table 1. Primers used for the molecular detection of SARS-CoV-2 RNA.

Target Gene	Primer Name (orientation)	Nucleotide sequence 5'→3'	Ta[1] (°C)	No. of cycles	Amplicon size	Position[2] nucleotide (amino acid)	Mutation or Motif[3]	Reference
NsP3	C3037T $F_{1/2}$ (+) C3037T R_1 (-) C3037T R_2 (-)	TTGATTTAGATGAGTGGAGTATGGCTAC GTCTGAACAACTGGTGTAAGTTCC CATCATCTAACCAATCTTCTTCTTGCT	60	40	263 bp	2949-3212 (76-164)	C3037T (F106F)	This study
RdRp	C14408T $F_{1/2}$ (+) C14408T R_1 (-) C14408T R_2 (-)	TTTGGATGACAGATGCATTCTGC GATAGTAGTCATAATCGCTGATAGCAG CCGGGTTTGACAGTTTGAAAAGC	58	40	335 bp	14343-14678 (300-412)	C14408T (P323L)	This study
Spike A	A23403G F_1 (+) A23403G F_2 (+) A23403G $R_{1/2}$ (-)	TTCAACTTCAATGGTTTAACAGGCA GTCCGTGATCCACAGACACTTG GTGGGTATGGCAATAGAGTTATTAGAGT	60	40	421 bp	23288-23709 (575-715)	A23403G (D614G) FCS[4]	This study
Spike B	OUT-F (+) MaSi_AR (-) SiMa_BF (+)	AGGCTGCGTTATAGCTTGGA ACACTGACACCACCAAAAGAAC TCTTGATTCTAAGGTTGGTGGT	55	35	470 bp	22882-23352 (440-596)	T22917G (L452R), G23012A (E484K), G23012C (E484Q), A23063T (N501Y), C23271A (A570D)	Ref.[14]

[1] Annealing temperature; [2] Referred to Wuhan HU-1 (NC_045512.2); [3]Amino acid mutations are indicated in parenthesis; [4] FCS: furin cleavage site (amino acids 680-689)

Table 2. Details of identified positive cases[$].

Patient ID	Area	Collection date	Delay[1]	Age (yr)	Sex	RT PCR	NsP3[2]	RdRp[2]	Spike A[2]	Spike B[2]	IgM[3]	IgA[3]	IgG[3]	PRNT 80%[4]
Pre-pandemic cases														
1	Milan	12/09/2019	3	0·67	M	-	-	-	-	+ (U)	+	-	+	-
2	East Milan	12/10/2019	0	18	F	-	-	-	-	+ (S)	-	-	-	-
3	Brescia	17/10/2019	1	9	F	-	-	+ (U)	-	-	-	-	-	-
4	North-West Milan	19/10/2019	1	1	M	-	-	+ (S)	-	-	-	+	+	-
5	Brescia	22/10/2019	1	1	F	-	+ (S)	-	-	-	+	-	-	-[5]
6	East Milan	23/10/2019	NA	1	F	-	-	+ (S)	-	-	+	-	-	-
7	North-West Milan	22/11/2019	7	2	M	-	+ (S)	-	-	-	NA			NA
8	West Milan	05/12/2019	4	4	M	-	-	-	-	+ (S)	NA			NA
9	South-East Milan	15/12/2019	2	25	F	-	-	+ (U)	+ (U)	+ (U)	-	-	-	-
10	Milan	09/01/2020	1	53	M	-	-	-	-	+ (U)	-	-	-	-
11	Brescia	14/01/2020	4	40	M	-	-	-	-	+ (S)	-	-	-	NA
Pandemic cases														

12	North-West Milan	15/01/2021	1	65	F	-	-	-	-	-	-	+	+	1:60
13	Lodi	16/01/2021	9	32	F	-	+ (U)	-	+ (U)	-	+	-	-	-
14	North Milan	25/01/2021	4	1	F	-	-	-	-	-	+	+	+	1:160
15	Milan	02/02/2021	1	1	F	-	-	-	-	-	+	+	-	-
16	Milan	26/02/2021	5	47	M	-	-	-	-	+ (U)	-	-	-	-
17	North Milan	24/03/2021	6	32	M	-	-	-	-	-	-	+	+	NA

[$]Patients who showed positivity for IgA alone are not reported as their diagnosis was considered inconclusive.

[1]Days from exanthema onset to sample collection. [2] U: urine, S: oropharyngeal swab. [3]NA: serum/data not available. [4]PRNT: plaque-reduction neutralization test.

[5]This serum caused 62% plaque reduction.

Table 3. Results of molecular and serological screening for SARS-CoV-2 performed on samples collected from measles- and rubella-negative patients between August 2018 and March 2021[$].

Group	Control	Pre-pandemic	Pandemic
Sampling period	August 2018 - July 2019	August 2019 - February 2020	March 2020 - March 2021
N. cases (N. samples)	100 (281)	44 (126)	12 (34)
PCR positive swabs (%)	0/93 (0)	7/43 (16·3)*	0/12 (0)
PCR positive urines (%)	0/90 (0)	4/39 (10·3)**	2/12 (16·7)***
Total PCR positive patients (%)	0/100 (0)	11/44 (25)*	2/12 (16·7)***
IgG positive sera (%)	0/98 (0)	2/38 (5·3)	3/10 (30)
IgM positive sera (%)	0/98 (0)	3/38 (7·9)***	3/10 (30)
IgA positive sera (%)	0/98 (0)	1/38 (2·6)	4/10 (40)
Ig positive patients (%)	0/98 (0)	4/38 (10·5)§	5/10 (50)*,#
Overall positives (%)	0 (0)	11 (25)*	6 (50)*

[$]Patients who showed positivity for IgA alone are not reported as their diagnosis was considered inconclusive.

* $p<0·001$ with respect to the control group (Mid-P exact, two-tail test)

** $p<0·01$ with respect to the control group (Mid-P exact, two-tail test)

*** $p<0·05$ with respect to the control group (Mid-P exact, two-tail test)

§ 1 subject positive for IgM and IgG, 1 subject positive for IgA and IgG

#1 subject positive for IgM and IgG, 1 subject positive for IgA, IgM, and IgG, and 2 subjects positive for IgA and IgG

Figure legends

Figure 1. Measles and rubella epidemiological trends observed in our laboratory from 2017 (start of surveillance activity) to 2021. The number of measles/rubella suspected cases investigated is indicated by grey bars (right axis), and the percentage of discarded measles/rubella suspected cases (suspected cases with negative diagnostic) is shown by a black line (left axis). The asterisk marks the period of interruption of surveillance activities due to lockdown measures.

Figure 2. Schematic display of mutations identified in this study mapped on to the global mutational history of SARS-CoV-2. The backbone mutational tree was reconstructed from an analysis of >68,000 genomes estimates in which mutations are denoted by Greek letters [4]. Variants observed in the Italian samples analyzed in this study are marked by the patient identifiers (ID, see Table 2) along with the earliest sample acquisition date for each variant in our samples. Two patients (#9 and #13) appear twice because they contained two mutations each (β_2 and β_3, and β_1 and β_2, respectively). Nucleotide mutations are indicated in black while the amino acid mutation in the spike protein (β_2) is indicated in red. Also shown are the points of attachment of 20A and B.1 coronavirus lineages as well as some strains of high concern: B.1.617 from India, B.1.351 from South Africa, P.1 from Brazil, and B.1.1.7 from the UK. A dotted line is connecting β mutations because their relative order in the cluster of three mutations could not be established with a high statistical confidence [4]. ProCoV2 is the most recent common ancestor (progenitor) of all known SARS-CoV-2 genomes sequenced to date.