Discover Viruses

RESEARCH Open Access

Virome diversity and novel virus discovery in Italian ticks through metatranscriptomics

Andrea Silverj[1,2,3,7], Niccolò Alfano[3,8], Fausta Rosso[3], Valentina Tagliapietra[3], Luciano Toma[4], Marco Di Luca[4], Alessandra Torina[5], Francesco La Russa[5], Nicola Segata[2], Omar Rota-Stabelli[1,2,3*†] and Annapaola Rizzoli[3,6*†]

†Omar Rota-Stabelli and Annapaola Rizzoli are co-senior authors.

*Correspondence:
Omar Rota-Stabelli
omar.rotastabelli@unitn.it
Annapaola Rizzoli
annapaola.rizzoli@fmach.it

Full list of author information is available at the end of the article

Abstract

Ticks are parasites of a wide range of animal species and the main vectors of many viruses of health concern. However, a genome-scale characterization of tick diversity in Italy remains lacking, although several areas of the country are known to be affected by tick-borne viral infections. To begin addressing this important knowledge gap, we sequenced 22 tick transcriptomes from three different locations along a north-south gradient in Italy. We found that viral richness, prevalence, and abundance varied among samples. In our dataset, northern populations showed a higher presence of viruses of potential health concern; however, given the unbalanced and temporally split sampling design, these observations should be interpreted with caution. Using a combination of metagenomics and phylogenomics, we recovered 99 viral contigs, which were assembled into 89 uncultivated viral genomes (UViGs), as some viral species possess segmented genomes composed of multiple contigs. These 89 UViGs correspond to 31 viral operational taxonomic units, including 10 new putative species (i.e., unclassified viruses) and 19 species reported in Italy for the first time. We present the updated phylogenies of viruses of known or new potential zoonotic concern including Haseki tick virus, members of the *Orthototiviridae* and *Phenuiviridae* families, Tick-borne encephalitis virus, and four different phleboviruses (Brown dog tick phlebovirus 2, Norway phlebovirus 1, Leuven phlebovirus and Tick phlebovirus). These results increase our general understanding of the evolution of tick-borne viruses, provide a comprehensive look at their diversity in Italy, and represent a source of information for implementing future surveillance activities.

1 Introduction

The incidence of tick-borne diseases is increasing worldwide [1] due to the combined effects of different drivers, such as climate change, shifts in land use, biodiversity loss, and increased circulation of goods and humans. These ecological factors are contributing to expanding the distribution of ticks and their associated pathogens, raising the risk of new epidemics and threatening the health of humans and other animals [2, 3]. For this reason, many countries are improving their surveillance activities to rapidly identify and track new tick-borne pathogens and diseases and to monitor the trend of the existing ones [4, 5].

Gene targeted assays, such as PCR, are useful tools to test for the presence of a single (or few) already characterized pathogens, but cannot be used for large screenings of new viruses and for the identification of new pathogens [6]. Shotgun non-targeted methods such as metagenomics and metatranscriptomics are quickly becoming cost effective ways [7] to study all microorganisms at the same time, even the ones that are still unknown, and are therefore powerful tools for the early detection of new emerging pathogens. With non-targeted methods it is possible to process samples from different sources to understand how microbes circulate among the environment, humans and other animals, in a One Health approach [8]. These methods have recently been applied to discover and track tick-borne viruses that have a high epidemic potential and are the main cause of tick-related diseases [9–12].

At least 1,477 viruses from 73 different families have been described in ticks worldwide [13–15], but the pathogenicity for humans and animals of a large part of them remains unclear. Tick-borne encephalitis virus (TBEV), belonging to the species *Orthoflavivirus encephaliditis* [15, 16], is one of the most dangerous viruses for human health in Eurasia. It is the most widely distributed neurotropic arbovirus, belonging to the family *Flaviviridae*, genus *Orthoflavivirus*, tick-borne encephalitis serocomplex [16, 17]. TBEV is transmitted mainly by hard ticks as those belonging to the *Ixodes ricinus* species complex, and its infection can lead to serious long term neurological sequelae [18]. The incidence of this virus in Eurasia has been increasing in the last decade, with the appearance of new foci; in Europe, it is currently endemic in 27 countries [19], with an expanding trend also in previously unaffected regions. The highest numbers of cases are found in countries located in central and northern Europe (Czech Republic, Lithuania, Poland, Germany and Sweden), but the virus is progressively expanding to other countries such as the Netherlands and UK [20]. In Italy, TBEV is distributed in small natural foci as observed in other regions, and the risk is currently higher in the northern part of the country [21].

Apart from TBEV, the diversity of other tick-associated viruses hasn't been explored extensively in Italy so far. To fill this knowledge gap, we sampled ticks of different species from multiple Italian locations on a latitudinal gradient and sequenced their associated viromes with a metatranscriptomic approach. We assembled 89 uncultivated virus genomes (UViGs), many of which belong to species that had never been described in Italy before or to putative new species, and reconstructed their evolutionary history using phylogenetic methods.

2 Results

2.1 Overview of generated metatranscriptomic samples

We analyzed the transcriptomes of 22 tick samples belonging to five species (*Dermacentor marginatus*, *Haemaphysalis punctata*, *Ixodes ricinus*, *Rhipicephalus bursa*, and *Rhipicephalus sanguineus* s.l., hereafter *Rhipicephalus sanguineus*) from three different regions of Italy (Trentino-Alto Adige, Lazio, Sicily), maximizing the latitude range between all samples (Figs. 1a, b and Supplementary Table 1). These regions were selected not only to capture a broad north–south gradient, but also to represent ecologically distinct and characteristic environments within Italy that are relevant to tick ecology. Specifically, Trentino-Alto Adige reflects alpine and subalpine habitats in northern Italy, Lazio represents central Mediterranean ecosystems with mixed rural and peri-urban

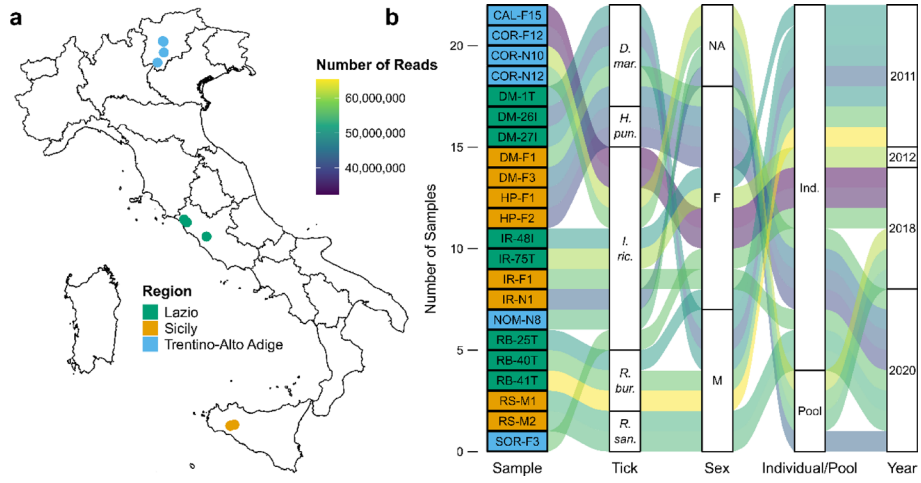

Fig. 1 Overview of the 22 metatranscriptomic samples generated in the study. **a** All sampling locations of the study are highlighted in the map, spanning northern (Trento, in light blue), central (Lazio, in green) and southern (Sicily, in orange) latitudes of Italy. **b** Alluvial diagram showing the main characteristics of the generated samples. The colors of the bar on the left indicate the sampling region, while the curves show how the samples are divided according to tick species, sex, type of sample (single tick or pool of more individuals), and year (four white bars on the right). The color gradient of the curve shows the number of reads for each sample

landscapes, and Sicily encompasses southern Mediterranean environments characterized by warmer climates and distinct vegetation types. Together, these locations provide a representative ecological framework for assessing tick-associated viral diversity across diverse climatic and environmental contexts in Italy.

2.2 Identification of 99 putative viral contigs

All 22 generated metatranscriptomes (Supplementary Table 2) were analyzed, filtering the host reads and removing all the remaining ribosomal RNA (Supplementary Table 3). To reduce methodological biases, we reconstructed viral contigs by employing two different assembly strategies to obtain a set of 180 contigs (see Methods and Supplementary Table 4). A series of filtering steps (Supplementary Tables 3, 4) resulted in a set of 99 contigs (Supplementary Table 5) that were selected as putatively viral and further annotated and checked for their quality (see Methods), including the number of reads mapping against them, calling and identifying their genes with two methods that returned similar results (see Methods and Supplementary Data). Furthermore, the depth and breadth of coverage for each contig was calculated to identify poorly assembled sequences (see Methods and Supplementary Fig. 1).

We identified 31 viral operational taxonomic units (vOTUs) (Fig. 2a). The abundance of each vOTU was measured using the logarithm of reads per kilobase per million (logRPKM). This gives a general picture of the composition of the viral communities identified in the study for each sample and for different hosts and locations. At this level, TBEV accounted for almost all *Flaviviridae* reads, and it was detected in five *I. ricinus* samples, four of which were from Trentino-Alto Adige, and in one *H. punctata* from Sicily (HP-F1, though this is most likely an index-hopping artifact in this sample, as discussed in the Results and Methods sections). PCR analyses for detecting TBEV obtained in a previous study on the same samples [21] were coherent with what we found using metatranscriptomics, representing a further and independent validation of our approach (see Methods). In particular, we found a high number of reads mapping against the

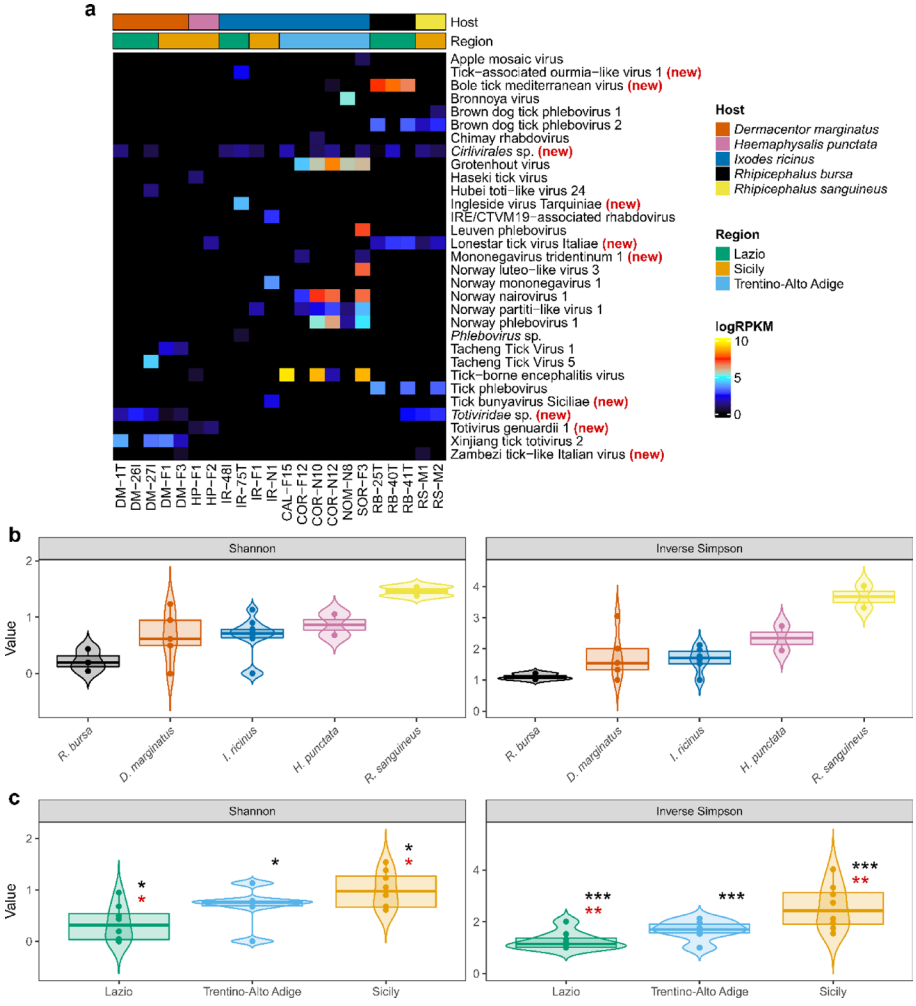

Fig. 2 Overview of the vOTUs found and viral diversity. **a** A heatmap showing a measure of viral abundance expressed as the logarithm of reads per kilobase per million (logRPKM) for all viral operational taxonomic units (vOTUs) identified for each sample, with higher values (in orange, red and yellow shades) indicating greater viral abundance. Color bars on top indicate the host from which the sample was extracted and its location. The differences among all values and between each single species were all not significant ($P > .05$ for both Kruskal-Wallis and Wilcoxon rank-sum test). **b** Alpha diversity values (at the viral species level) calculated for each tick species. **c** Alpha diversity values (at the viral species level) grouped by sampling region. Significance comparisons for alpha diversity are indicated with black asterisks (Kruskal-Wallis test) or red asterisks (Wilcoxon rank-sum test); *$P = .05$, **$P = .01$, ***$P = .005$. Only significant comparisons ($P = .05$) are shown

TBEV assembled genomes for the samples that tested positive for the virus with PCR (CAL-F15, n = 2478393; SOR-F3, n = 2216638; COR-N10, n = 2291288), and much lower reads for a sample that was weakly positive (COR-N12, n = 1232). Very few reads were retrieved for a sample from Sicily (HP-F1, n = 586), for which no PCR data were available.

The viruses identified belonged to thirteen known viral families, as well as to unknown or unclassified ones (Supplementary Fig. 2). Some viral families were host-specific, like the *Chuviridae*, that are present only in *R. bursa* ticks, and the *Botourmiaviridae*, *Flaviviridae*, Unclassified *Tolivirales* family (formerly *Luteoviridae* [22]), *Partitiviridae* and *Rhabdoviridae*, that are mainly (and in some cases only) present in *I. ricinus*. Vice versa, some families showed a broader host range, like the Unclassified *Cirlivirales* and the *Orthototiviridae*. Some viral families showed high logRPKM values in several samples

(in particular *Flaviviridae*, *Chuviridae* and *Nairoviridae*), while others were very low. The reads mapping to unclassified viruses were merged together and represented different viral groups that did not necessarily belong to the same taxonomic unit.

Alpha diversity was compared across regions and tick species (Figs. 2b, c); however, because sampling years differed among regions and tick species composition varied by site, these comparisons should be interpreted with caution. Given these sources of heterogeneity, we consider the analyses of geographic patterns in alpha diversity to be descriptive rather than inferential. Accordingly, the results are presented as an exploratory overview of observed variation among regions, without drawing formal conclusions about geographic effects.

The tick species with the highest alpha diversity (Fig. 2b) was *R. sanguineus*, followed by *H. punctata*, *I. ricinus*, *D. marginatus* and *R. bursa* for both Shannon and inverse Simpson indexes, even though the differences were not statistically significant ($P > .05$ for both Kruskal-Wallis test and pairwise Wilcoxon rank-sum test). Alpha diversity values varied between regions (Fig. 2c), with Sicily being the most diverse sampling area, followed by Trentino-Alto Adige and Lazio; differences among all samples were statistically significant ($P < .05$, Kruskal-Wallis test). Individual differences between areas were significant only for the Lazio-Sicily pair ($P < .05$, Wilcoxon rank-sum test).

2.3 Phylogenetics and genome comparisons reveal known and new viral species

We identified 17 vOTUs suitable for phylogenetic analysis (see Methods). We used genetic distances to determine whether our newly reconstructed genomes could be assigned to previously known (classified) viruses ("kv") or instead represented putative new (unclassified) viruses ("pnv") (see Table 1, Supplementary Tables 4 and 5, and Methods). Throughout the manuscript, the term "putative new virus" refers to a virus that could not be assigned to an existing classified viral taxon based on its genetic distance to its closest classified relative, rather than to a virus sequenced for the first time. Likewise, references to the "closest relative" indicate the closest classified relative available

Table 1 Putative new viruses identified in this study

Novel viruses (virus common name)	Closest classified relative	Viral family	Species demarcation criteria
Bole Tick Mediterranean virus	Bole Tick virus 3	*Chuviridae*	< 90% RdRp similarity
Cirlivirales sp.	Circoviridae sp.	Unclassified *Cirlivirales* family	< 80% whole-genome similarity
Ingleside virus tarquiniae	Ingleside virus	*Virgaviridae*	< 80% whole-genome similarity
Lonestar tick virus italiae	Lonestar tick totivirus	*Orthototiviridae*	< 90% RdRp similarity
Mononegavirus tridentinum 1	Norway mononegavirus 1	*Unclassified*	< 80% whole-genome similarity
Tick-associated ourmia-like virus	Aspergillus neoniger ourmia-like virus 1	*Botourmiaviridae*	< 90% RdRp similarity
Tick bunyavirus siciliae	Ixodes scapularis bunyavirus	*Peribunyaviridae*	< 90% RdRp similarity
Orthototiviridae sp.	Xinjiang tick totivirus 2	*Orthototiviridae*	< 90% RdRp similarity
Totivirus genuardii 1	Hubei toti-like virus 24	*Orthototiviridae*	< 90% RdRp similarity
Zambezi tick-like Italian virus	Zambezi tick virus 1	*Orthomyxoviridae*	< 80% whole-genome similarity

10 putatively novel viral species, belonging to 8 different viral families, were identified based on sequence identity, phylogenetics and comparisons to the closest known genomes

in current reference databases. For the ones with high identity to known viral species and no putative new viruses, we reconstructed species-level phylogenies (Supplementary Fig. 3).

To further confirm the species demarcation scenario obtained using sequence identity, we obtained phylogenetic trees at the family level based on the amino acid sequence of the RdRp (or Rep) protein for all the new putative viral species (Supplementary Figs. 4, 5, 6, 7, 8, 9, 10 and 11).

Viruses exhibit diverse substitution patterns, ranging from very low for the known species (e.g., Grotenhout virus, Norway nairovirus 1, Norway partity like virus 1, Norway phlebovirus 1; Supplementary Fig. 3), confirming their identification at this level, to very high in others (e.g., Ingleside virus, Bole tick virus 3, Norway mononegavirus, Xinjiang tick totivirus 2; Supplementary Figs. 4, 5, 6, 7, 8 and 10, 11), supporting their likely classification as new species. In some cases, such as Bole tick virus 3 (Supplementary Fig. 5), we included some previously obtained sequences into our newly proposed putative viral species (in this case, Bole Tick Mediterranean virus). We significantly expanded the number of available sequences for the species analyzed for some groups, doubling or more than doubling the number of genomes (e.g., Norway partiti like virus 1 and Grotenhout virus, Hubei toti-like virus 24; Supplementary Figs. 3 and 8, respectively). The new highly prevalent *Cirlivirales* sp. detected in 12 samples (Supplementary Fig. 6) revealed a closer match to the Replication and Capsid proteins that characterize the typical *Cirlivirales* genome structure [23]. To increase the phylogenetic resolution of this analysis, we also included genomes from all *Cirlivirales* families (*Circoviridae*, *Vilyaviridae*, and *Endolinaviridae*), and reconstructed a phylogeny using more distant relatives inferred trough protein structures (Supplementary Fig. 7).

Our phylogenies indicate the need to expand the number of viral species: for example, in the case of Xinjiang tick totivirus 2 (Supplementary Fig. 8) according to genetic distance, we may define at least 3 distinct viral species. To confirm the taxonomic assignment of our sequences, we used the genome annotation obtained for all genomes to carry out genome comparisons (Supplementary Figs. 4, 5 and 6 and 9 10, 11, 12 and 13). We built Clinker plots (see Methods) for each species data set to explore the gene content and synteny of each genome. This approach was essential to characterize those species for which there were not enough reference sequences to reconstruct a phylogeny (Supplementary Fig. 12). In most cases we retrieved the full *RdRp* gene of the virus, and for some viruses we could reconstruct the nearly complete genome (IRE/CTVM19-associated rhabdovirus, Tacheng tick virus 5). In other cases, we were only able to get partial genomes or single proteins besides RdRp, which hinders a reliable taxonomic assignment (such as for Norway mononegavirus 1 and Zambezi tick virus 1; Supplementary Fig. 13). Beyond RdRp, the annotated proteins included glycoproteins, nucleocapsid, nucleoproteins and putative functional ORFs. Cases for which viral species were particularly difficult to classify, either for the lack of reference sequences or for the high fragmentation of the genome, are presented in Supplementary Fig. 14. Three partial genomes were assigned to three different species of the *Phenuiviridae* family (Brown dog tick phlebovirus 1 and 2 and Phlebovirus sp.).

2.4 Widespread presence of TBEV in northeastern Italy

To better characterize the diversity and evolution of TBEV, a maximum likelihood phylogeny with the five newly generated sequences was reconstructed (Fig. 3). The phylogenetic tree in Fig. 3a indicates that TBEV genomes can be divided into five clusters based on their genetic divergence: Far Eastern, Baikalian, Siberian, Himalayan and European. All the five sequences from this study (four from Trentino-Alto Adige and one from Sicily) cluster together within the European group (Fig. 3a, inside the dotted red circle). These sequences are all part of a clade (depicted with a red circle in Fig. 3a and b) which also includes another Italian genome (OM084948.1, isolated in the Friuli-Venezia Giulia region, close to Trentino-Alto Adige, in 2021) and sequences from Germany (GQ266392.1, collected in 2005) and Finland (GU183381.1, sequenced from a sample collected in 1960). The slow substitution rate of TBEV is reflected by the very small scale-bar (Fig. 3b). The Sicilian sample, which has never been detected in this region before and is very similar to those from Trentino-Alto Adige, could be due to an index-hopping artifact, as discussed in detail in the following sections.

2.5 Non-viral microbial profiles reveal the presence of well-known tick symbionts

We screened our non-tick and non-viral reads for the presence of microbial taxa using a read-profiling approach (see Methods). Our results show the presence of well-known tick symbiotic bacteria, such as *Midichloria mitochondrii*, *Coxiella* spp., *Wolbachia pipientis*, *Delftia acidovorans*, *Rickettsia* spp. (Fig. 4). We detected *M. mitochondrii* and *Delftia* spp. only in *I. ricinus* ticks, *Rickettsia* spp. only in *D. marginatus*, and *Coxiella* spp. only in *Rhipicephalus* spp.

As metatranscriptomics reflects the intensity of transcriptional activity of a biological entity, the estimated values must be interpreted as relative abundances of the detected transcripts, which could not reflect the actual genomic abundances. Nonetheless, high relative abundances strongly suggest the presence of the corresponding species in the sample. Even considering this important limitation, it is worth noting that well-known tick symbionts (highly abundant in the host) tend to have the highest values in the heatmap.

3 Discussion

We explored the diversity and evolution of RNA viruses harbored by ticks from Italy, a European country which has been recently experiencing an increasing risk of tick-borne diseases. In particular, an increase in the number of cases of TBE has been registered in the last decades in the northern regions [21, 24]. Tick populations were sampled from three geographically different and distant regions (Trentino-Alto Adige, Lazio and Sicily). Our aim was to consider different tick species capable of parasitizing various vertebrates to characterize the overall diversity of RNA viruses associated with them. Many contigs were similar to unclassified viruses that were present in the databases, while others showed only low identity scores (< 80% on the whole genome or < 90% on the RdRp) for all known viruses present in the GenBank database. A taxonomic refinement was carried out by using genome comparisons and phylogenomics. Our analysis revealed the presence of viruses belonging to at least 13 known viral families. Some of them were detected in several tick species (like the Unclassified *Cirlivirales*, *Phenuiviridae* and *Orthototiviridae*), while others seemed to be more species-specific (like the

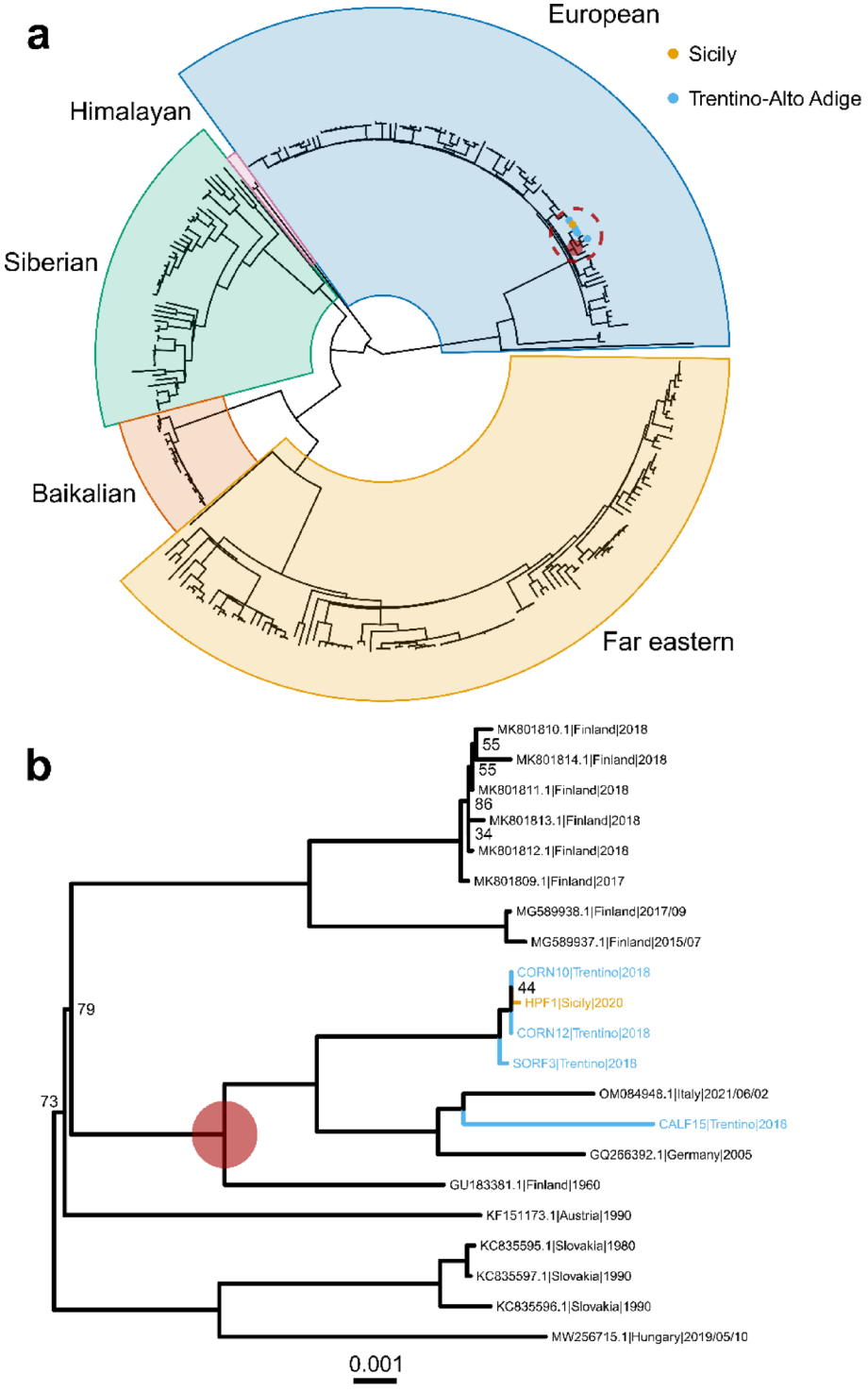

Fig. 3 Maximum likelihood phylogeny reconstructed for TBEV. **a** A maximum likelihood phylogenetic tree showing the evolutionary relationships among all publicly available ($n = 282$) and the newly generated TBEV genomes ($n = 5$). The dotted circle highlights the area of the tree which is enlarged and visualized in panel **b**, containing a clade (full red dot) composed of all the new genomes and 3 other similar sequences from Italy, Germany and Finland. A scale bar at the bottom indicates the number of substitutions per site. Bootstrap values below 90 are indicated in the tree in Fig. 1b with the exact number

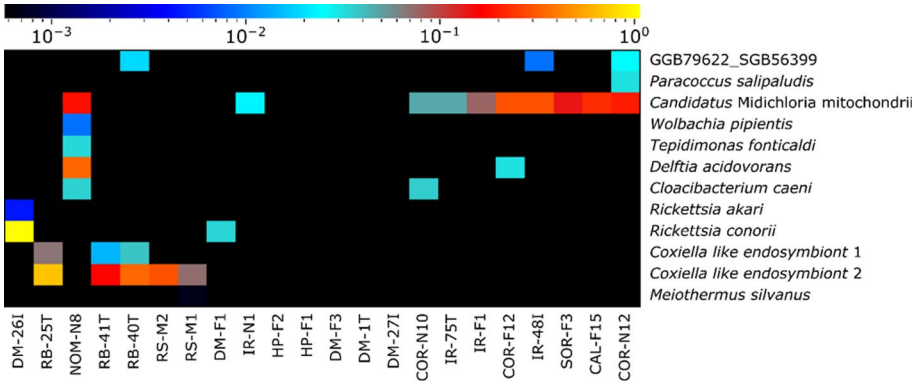

Fig. 4 Presence of bacterial taxa in the analyzed samples. Bacterial taxa (species names on the left) detected in the samples (bottom) are shown. Colors indicate the relative abundance of each species, estimated by MetaPhlAn based on read mapping to a database of clade-specific marker genes

Chuviridae and *Partitiviridae*, which were found only in *R. bursa* and *I. ricinus*, respectively). The differences observed in genome structure within the Unclassified *Cirlivirales* family most likely reflect the circular nature of their genomes. When circular genomes are linearized for sequencing, the cleavage site (and therefore the start position of the sequence) is often arbitrary and may differ between studies or database submissions, producing apparent mismatches in gene order that are not true biological rearrangements [25–27]. Phylogenetic analysis of this group, both using sequence (Supplementary Fig. 6) and structural data (Supplementary Fig. 7), which usually have more resolution at higher evolutionary distances [28], did not bring to a reliable placement of our contigs into an existing family within the *Cirlivirales* order. Given these results, further studies and new *Cirlivirales* genomes are required to better define the taxonomy of this group.

The detection of some other families, where only one virus was found in only one tick species and sample, seems to be accidental, like *Botourmiaviridae* and *Bromoviridae*, which are indeed viruses of fungi and plants. This may be due to the fact that the ticks were not washed using aggressive treatments (such as bleach-based solutions) for better preservation of the samples, potentially leaving some traces of non-tick viruses on their cuticles. The observed difference degree in specificity may also be influenced by other factors, including viral preferences in terms of tick species, localized viral circulation, or differences in the sampling strategy. Different tick species were sampled in different regions, with some regions including more species than others. A similar bias could influence the observed diversity, with better sampled regions showing higher values compared to the downsampled ones. For this reason, the diversity scores estimated in this study must be interpreted with care.

At the viral contig level, we found five different phleboviruses (Brown dog tick phlebovirus 2, Norway phlebovirus 1, Leuven phlebovirus and Tick phlebovirus). *Phlebovirus* is a genus of viruses included in the large order of *Hareavirales*, characterized by a high genetic diversity [29] and divided into two antigenic complexes, one transmitted by Diptera (mosquitoes, sandflies, etc.) and one transmitted by ticks [30]. Tick-borne phleboviruses were not considered a human health issue until the discovery of a severe fever with thrombocytopenia syndrome (SFTS), reported in China in 2010 for the first time [31]. A similar hemorrhagic fever was observed in the USA, caused by a tick phlebovirus named Heartland virus, genetically closely related to SFTSV [32]. Even if phleboviruses found in

Italy are still not clearly associated with human disease, their presence is noteworthy, as some of these viruses are characterized by gene mutation and genetic reassortment that could lead to new pathogenic strains [31].

Some of the viruses detected in our samples could be related to new and emerging tick-borne infections. This is the case of Haseki virus, a novel Flavi-like virus discovered recently in Russia in ticks (*Ixodes persulcatus*) and in patients with fever and respiratory problems after tick bites [33]. It must be noticed that Haseki virus often coinfected patients with other tick-borne pathogens, like *Anaplasma* and *Borrelia*, causing mixed infections and complicating the set of symptoms [33]. In our study, Haseki virus was detected in *H. punctata* from Sicily, and this is the first report of this virus in Italy to date. We also found a virus which showed similarities with Bole-tick virus 3 [34] in *I. ricinus* and *R. bursa*. This is the first report of this virus in *Ixodes* ticks, as in literature it has only been found as a component of the virome of *Rhipicephalus* and *Hyalomma* ticks [35].

Another virus potentially linked to human disease is Tacheng virus, whose serotype 5 was detected in our study. Recently, a variant of this virus was isolated in China, from a hospitalized patient with a syndrome involving different organs after a tick bite [36]. The same virus was identified in ticks (*Dermacentor* and *Hyalomma*) collected in the area where the patient lived, suggesting the existence of a pathogenic tick-borne virus.

Hubei toti-like virus, a double stranded RNA virus, was particularly abundant in our samples belonging to the species *D. marginatus.* This virus was recorded in *I. persulcatus* in North China [37] and the classification of the family was revisited in 2006; in fact, at the origin *Orthototiviridae* were considered mycoviruses, but the discovery in invertebrates and fish of viruses clustering phylogenetically with this family suggested the creation of a Toti-like group [38].

In some samples, we also detected the presence of Xinjiang tick totivirus 2. Other viruses detected in our analysis, like Grotenhout virus and Bronnoya virus, are considered tick symbionts, or part of the core virome of ticks. Both viruses are members of the order *Bunyavirales*, belonging to different genera [39, 40]. The presence of Trinbago virus [41] in *H. punctata* is noteworthy, as this virus seems able to infect different tick species and it is classified as *Orthoflavivirus*, but its pathogenicity is still to be investigated.

Interestingly, we found a plant virus contig in our results, even after all filtering steps. This is not infrequent, and it is probably due to ticks activities on plant surfaces [14], as well as to the fact that we did not use aggressive treatments to remove viruses present on the cuticles. Apple Mosaic Virus (ApMV; *Bromoviridae*) was present in *I. ricinus* collected in Val di Non, located in the north-west of the Trentino-Alto Adige region and well known in Italy for apple production. The presence of this virus in apple orchards has been very rare since the 90s, when it became mandatory for young plants used for restocking to be certified as ApMV free. It is true, however, that this virus can affect more than 60 species of plants (hazelnut, birch, etc.) [42], so it is still possible to detect it in the environment.

The most abundant viruses, in terms of RPKM, were *Flaviviridae* (in particular TBEV; Supplementary Fig. 2). This probably reflects their capacity of reaching high titers in ticks, which favors the possibility of transmission to other species. The detection of TBEV in some *I. ricinus* samples from the Trentino-Alto Adige region confirms

its circulation in the area, which is also close to other countries with a similar scenario, like Austria and Germany. It is important to note that the *I. ricinus* ticks that were positive to TBEV in this study came from a previous molecular screening of 2500 samples [21], and that the frequency of infected ticks in nature is much lower. Surprisingly, we detected this pathogen in a *H. punctata* sample from Sicily, which would make this the first proof of the occurrence of this virus in ticks on the island, even though a suspected case of autochthonous TBEV was recorded in 2018 [43]. This virus has been detected even at southern latitudes in the past, such as in North Africa and Tunisia [44]. However, the very low number of reads of the virus in the sample (well below the index-hopping threshold, with read frequencies < 0.1%; see Methods), its close similarity and phylogenetic relatedness to the TBEV genome from the COR-N10 sample that was sequenced together, and its fragmentation (5 different contigs that were merged together) suggests the possibility of it being a false positive. Even though Sicily and Trentino-Alto Adige are geographically distant, we cannot rule out that the very strong similarity between the two genomes could be due to the role of migratory birds in the transportation of the pathogen. All this evidence calls for further studies in Sicily to verify the presence of TBEV.

The sampling process was not oriented to test differences between geographical areas or tick species, given that the number and species of ticks collected differed by location. This limitation prevented us from using beta-diversity indexes or to test the effect for a possible influence of sex, location, species, life stage, number of ticks collected per pool and other parameters on the viral profiles retrieved in the analysis.

To summarize, in this study we identified new genomes for recently identified tick viral species, expanding the available taxon sampling. Our study confirmed ticks as a rich source of viruses, other than bacteria and protozoa, but further and deeper ecological, molecular and epidemiological analysis will be necessary to improve knowledge on their pathogenicity and define potential public health risk.

4 Materials and methods

4.1 Tick collection

In order to provide a wide representation of Italy from North to South, ticks were collected in three different areas spanning a large latitudinal distance: Trentino-Alto Adige (Northern Italy), Lazio (Central Italy) and Sicily (Southern Italy). Ticks were collected in 2018 in Trentino-Alto Adige, in 2011–2012 in Lazio, and in 2020 in Sicily (the exact collection dates and other details about sampling can be found in Supplementary Table 1). At each location, questing ticks were collected by dragging or flagging the vegetation. All ticks were identified microscopically by species, sex and life stage using taxonomic keys [45]. Larvae were discarded. All individuals were washed once in 70% ethanol followed by deionized water for 5 min, and then stored at − 80 °C until nucleic acid extraction. Adults were treated individually, while nymphs were grouped in pools of 10. Samples from different regions in Italy were directly collected by institutions officially authorized to perform animal sampling and pathogen screening as part of an institutional research mission by the Italian National Ministry of Health (Department of Infectious Diseases, Istituto Superiore di Sanità; Istituto Zooprofilattico Sperimentale della Sicilia "A. Mirri") or by the provincial government (Research and Innovation Centre, Edmund Mach Foundation).

4.2 RNA extraction and sequencing

Ticks were homogenized in 350 µl of RLT buffer using a MM200 mixer mill (Retsch). Total RNA was extracted using the RNeasy Mini Kit (Qiagen) and quantified with Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Potential traces of genomic DNA were removed through the TURBO DNA-free kit (Thermo Fisher Scientific). RNA integrity was assessed using a 2200 Tapestation (Agilent Technologies). Host rRNA was depleted and sequencing libraries prepared using the Zymo-Seq RiboFree Total RNA Library Kit (Zymo). Library fragment size distribution was checked on a 2200 Tapestation (Agilent Technologies) and concentration quantified using the KAPA Library quantification kit (Roche) on a LightCycler 480 Real-Time PCR System (Roche). Sequencing libraries were pooled in equimolar amounts and sequenced on a NovaSeq 6000 platform (Illumina) at the Institute of Applied Genomics (IGA, Udine, Italy). All extraction and sequencing procedures involving ticks were conducted following BSL-3 biosafety protocols.

4.3 Reads processing and genome assembly

A total of 1,110,547,774 paired-end sequence reads 150-bp long were generated and then sorted by index sequences (all read stats can be found in Supplementary Tables 2–3). An in-house analysis pipeline was used to process the data. Briefly, adapter sequences were trimmed from the reads using Cutadapt v1.16 [46], and low-quality reads were removed by Trimmomatic v0.36 [47]. The host background was removed by aligning the filtered reads against the full genome of the closest related species available in GenBank at the time of the study (*Ixodes ricinus* assembly GCA_000973045.2; *Rhipicephalus microplus* assembly GCA_002176555.1; *Haemaphysalis longicornis* assembly GCA_008122185.1) using Bowtie2 v2.3.4.1 [48]. rRNA reads were filtered out using SortMeRNA v2.1 [49] with default settings (six SILVA SSU/LSU databases and two RFAM databases). The resulting reads were de novo assembled using both Spades v3.11.1 [50] and Trinity v2.6.6 [51] assemblers. The dataset generated and analyzed in this study is available in the NCBI Sequence Read Archive (SRA) repository under the accession number PRJNA1068039.

4.4 Identification of putative viral contigs

Contigs from the two assemblers were pooled and clustered to remove duplicated or highly similar sequences using USEARCH v11.0.667 [52] with a 90% threshold identity value. The contigs centroids of the clusters were subjected to BLASTn v2.11.0 [53] alignment against the NCBI RefSeq viral nucleotide database (release 200) [54] with an E-value cutoff of 1e-05. Sequences that showed low or no homology at the nucleotide level were compared against the NCBI RefSeq viral protein database using BLASTx v2.11.0 search with an E-value <1e-03. Sequences with a significant hit at either step were considered viral candidate sequences. In order to remove false-positive viral hits, these sequences were further subjected to a second series of similarity searches: the contigs were screened against the entire NCBI nucleotide database (nt) using BLASTn with an E-value cutoff of 1e-10, and those viral candidates without significant BLASTn hits were searched by BLASTx against the complete GenBank non-redundant protein database (nr) using an E-value cutoff of 1e-03. For each BLAST search, the top 5 BLAST hits for each sequence were retained and a query was considered viral if the best hit corresponded to a viral subject, for a total of 2197 putative viral contigs. The best hits were filtered further for contig length > = 1000 bp, E-value < = 1e-20, match length > = 300

and identity > = 60, obtaining a set of 173 contigs. To further detect complete genomes of distantly related viruses, avoiding false negatives, we retained for a later inspection all contigs longer than 7000 bp that didn't pass the filtering, resulting in another set of 7 contigs. The two sets were merged and each contig was aligned against the corresponding reference that was used as subject in the BLASTn search using MAFFT v7.490 [55]. Putative contigs with overlapping ends were merged together by using the results of the BLASTn table and by manually inspecting the alignments with BioEdit v7.2.5 [56]. When two or more contigs belonging to the same viral species were assembled from the same sample, the longest contig was selected if the region shared by the contigs showed > 95% sequence identity. Otherwise, both contigs were retained and considered as two possible strains or species. Contigs < = 1500 bp with short, matched regions in the BLAST searches (< 50% of breadth of coverage to the reference genome) and showing low identity (< 75%) to the whole genome or to a viral marker different from RdRp were not analyzed further, as the available information was considered insufficient to reliably assign them as viral. The complete table showing all data for the 180 contigs, together with detailed information about the manual curation step, can be found in Supplementary Table 4. Eventually, a set of 99 contigs was used for the following steps of the analysis. This set was further annotated, and all metadata (including the GenBank accession numbers obtained for each sequence) can be found in Supplementary Table 5.

4.5 Depth and breadth of coverage

Genome coverage for each of the 99 selected contigs was calculated by mapping the clean reads (the set which was quality checked and filtered for the host genome and rRNAs) against the selected contigs using Bowtie v2.3.5.1 [48] and obtaining sorted bam files with SAMtools v1.3.1 [57]. The number of reads that mapped against each contig and the total number of reads for each library were calculated using SAMtools v1.3.1. The depth and breadth of coverage for each contig were calculated with CMseq v1.0.4 (https://github.com/SegataLab/cmseq), using a set of different coverage thresholds per base ("--minqual" 3, 5 and 10) and a minimum base quality of 30 (--minqual 30), and can be found in Supplementary Table 6. All these data were used to validate the contigs that were merged in the previous steps. According to the breadth and depth of coverage, the contigs were classified into 3 different categories: low quality contigs (breadth < 0.9 and/or depth < 5), high quality contigs (breadth ≥ 0.99 and depth ≥ 50) and, in all other cases, they were defined as medium quality contigs.

4.6 Quality control and validation using standard viromic tools

We further checked the set of the 99 putative viral contigs and benchmarked our results by running CheckV v1.0.3 [58] and geNomad v1.8.0 [59].

4.7 Genome annotation

All genomes were annotated using both Prokka v1.14.6 [60], with the "--meta --viruses" options enabled, and TransDecoder v5.7.1 [61]. For the TransDecoder analysis, the program was first run with the option "TransDecoder.LongOrfs". The output (longest_orfs.pep) was then used for a BLASTp search against SwissProt database, with an E-value of 1e-5, and for running hmmsearch against Pfam (Pfam-A.hmm), with an E-value of 1e-10. Eventually, we obtained the final predictions by running "TransDecoder.Predict"

with the contigs and the results of the BLAST and hmm searches as input files. All annotations were manually checked and merged together.

4.8 Viral taxonomy

We used results from the BLAST searches, read mapping and from the genome annotation to taxonomically classify the identified viral contigs. To assign taxonomy, we referred to the current ICTV guidelines specific to each taxon (see Supplementary Table 7). When specific criteria were not available, species showing > = 80% identity on the whole genome or > = 90% identity for the RdRp gene relative to a known species were assigned to the closest viral species (and marked as known viruses, "kv"), whereas the others were considered putative new viruses (referred to here as "pnv"), as previously described in other studies [62–65]. Further information about taxonomy was obtained from the output of geNomad (see the paragraph "Quality control and validation using standard viromic tools" above).

4.9 Genome comparisons

The annotated genomes that were taxonomically assigned in the previous step were compared with the genomes of the closest reference using clinker v0.0.28 [66], to check for gene presence, order and sequence similarity. The annotations obtained with Prokka and TransDecoder were compared to the clinker plots, enriching and completing the comparison between the new contigs and the reference.

4.10 Phylogenetic analysis and validation of putative new viral species

The selected hits were filtered excluding low-quality sequences (ambiguous characters > = 10%). We attempted to reconstruct the phylogeny for each identified UViG. In particular, for known viral groups, we downloaded all available reference genomes with the associated metadata from NCBI, while for the unknown UViGs we did the same using the closest reference (access date to NCBI: 31/07/2023), generating a database containing all genomes from this study and all the available references for each UViG. Datasets were cleaned and filtered using in-house Python scripts (available at https://github.com/andrea-silverj/BioinfoToolkit). For example, for TBEV we included all available complete (or nearly complete) TBEV genomes in GenBank at the time of this study (n = 282) and the new assembled genomes, for a total of 287 sequences. All datasets with less than 4 sequences were discarded, as that represents the minimum number of units to perform phylogenetic inference. Sequences of each set were aligned using MAFFT and trimmed with trimAl [67] (--gappyout). Phylogeny was reconstructed using RAxML v8.1.15 [68], with the options "raxmlHPC-PTHREADS-SSE3 -p 1989 -m GTRGAMMA -x 2483 -# 100 -f a -T 20", specifying the identified model for each database. The resulting trees were inspected with FigTree (http://tree.bio.ed.ac.uk/software/figtree/) and annotated and plotted using the retrieved metadata and the ggtree package implemented in Rstudio build 382 [69]. Phylogenetic trees at the family level were built for the 10 putative identified viral species. The RdRp or Rep proteins of each new putative virus were used, and we built a database of RdRp or Rep proteins for each viral family that was associated with the putatively new species (i.e., *Botourmiaviridae*, *Chuviridae*, Unclassified *Cirlivirales*, *Peribunyaviridae*, *Orthototiviridae*, *Virgaviridae*) by downloading a curated RdRp virus family database from the website of the International committee on Taxonomy

of Viruses (https://ictv.global/), which in the case of the *Virgaviridae* family was also expanded with strains coming from another study [70]. For the Unclassified *Cirlivirales* family, we retrieved representative genomes for all families belonging to this group (*Circoviridae*, *Vilyaviridae*, and *Endolinaviridae*), using the most updated ICTV classification. Alignments were obtained as above, and trees were reconstructed using RAxML v8.1.15, with the options "raxmlHPC-PTHREADS-SSE3 -p 1989 -m PROTGAMMAAUTO -x 2483 -# 100 -f a -T 20". We folded the Rep proteins of six representatives of our *Cirlivirales* phylogeny (based on the distances obtained from the maximum likelihood phylogenetic tree; see Supplementary Fig. 7), including a representative from our *Cirlivirales* sp. (which all shared 100% identity among each other for the Rep protein), using the AlphaFold3 web server [71]. The representative structure of the *Cirlivirales* sp. contigs was used as query to search the BFVD, CATH50, and PDB100 databases [72–74] through the Foldseek [75] web server (search results were stored in Supplementary Table 8), identifying distant relatives. The top five hits against each database were aligned and included in a phylogeny based on 3Di characters with the Foldmason web server [76]. All folded structure, as well as the structural multiple sequence alignments and the resulting trees, can be accessed at https://github.com/andrea-silverj/TickVirome_IT/tree/main/analyses/cirlivirales_structural_analysis.

4.11 Quantification of viral abundance

We estimated viral abundance in each library using reads per million (RPM), reads per kilobase per million (RPKM), and transcript per million (TPM) using custom R scripts.

4.12 PCR confirmation of TBEV positive samples

We used the results of previous study [21], which used PCR methods [77] to test the presence of TBEV in a wide range of tick samples from Trentino (some of which were included in the present study) to validate our findings deriving from metagenomics analyses. In all cases results were coherent with what we found using metatranscriptomics, representing a further and independent validation of our approach.

4.13 Contamination control

When sequencing multiple libraries simultaneously on the same flow cell, a phenomenon known as index-hopping can result in a small proportion of reads being misassigned to the wrong sample. This phenomenon is known to affect several sequencing platforms [78], including the one used for this study (Illumina NovaSeq 6000). To minimize the impact of this technical artifact on downstream analyses, a stringent filtering threshold was applied. Reads for a virus present at a frequency of less than 0.1% compared to the most abundant reads of that virus in another library were considered likely to represent contamination caused by index-hopping [78].

4.14 Identification of non-viral microbes

We analyzed the unfiltered reads for bacterial rRNA to check for the presence of bacterial species. Microbial profiles were obtained for each sample using MetaPhlAn-4 [79], which was run using the mpa_vOct22_CHOCOPhlAnSGB_202212 database. Relative abundances were calculated using the output table of MetaPhlAn-4, which is set on the calculation on DNA data.

4.15 Alpha diversity and statistical tests

Alpha diversity was calculated using a custom R script in which homemade functions to calculate the Shannon index and the Inverse Simpson index were developed. The functions were applied on the TPM values calculated at the species level. Overall differences among species and regions were tested for significance using Kruskal-Wallis test, while individual comparisons among the same categorical variables were assessed with a pairwise Wilcoxon rank-sum test (applying the Bonferroni correction).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s44370-026-00047-y.

Supplementary Material 1.

Supplementary Material 2.

Supplementary Material 3.

Supplementary Material 4.

Supplementary Material 5.

Supplementary Material 6.

Supplementary Material 7.

Supplementary Material 8.

Supplementary Material 9.

Acknowledgements
We would like to thank the Department of Infectious Diseases, Istituto Superiore di Sanità (Rome, Italy) and the Istituto Zooprofilattico Sperimentale della Sicilia "A. Mirri" (Palermo, Italy) for providing the samples and metadata for Lazio and Sicily, respectively. We also thank Prof. Mang Shi for his precious comments and suggestions.

Author contributions
A.S. wrote the original draft, performed data curation, formal analysis, visualization, and contributed to methodology, investigation, and software development. N.A. wrote the original draft, contributed to data curation, formal analysis, conceptualization, and software development, and participated in review and editing. F.R., V.T., L.T., M.D.L., A.T., and F.L.R. contributed resources and participated in writing—review and editing. N.S. contributed to methodology, supervision, resources, validation, and participated in writing—review and editing. O.R.S. contributed to supervision, methodology, conceptualization, validation, and participated in writing—review and editing. A.R. contributed to project administration, supervision, resources, methodology, investigation, funding acquisition, and conceptualization, and participated in writing—review and editing. All authors reviewed and approved the final manuscript.

Funding
This project has been partially funded by the Autonomous Province of Trento to Edmund Mach foundation under the agreement for the implementation of epidemiological surveillance activities in the field of public health with a One Health perspective (DGP n.2244 on 23.12.2024).

Data availability
Raw reads generated in this study can be found in the NCBI Sequence Read Archive under accession number PRJNA1068039. All supplementary figures, tables, data, sequence files in FASTA and GenBank format, and code used for this analysis (including the equations to calculate the abundance and diversity values) can be found at https://github.com/andrea-silverj/TickVirome_IT/tree/main and at https://drive.google.com/drive/folders/1f2aAg48uilMPqzI_Q_WL7d7WFUfU3aeL?usp=sharing. Viral sequences and gene annotations have been submitted to GenBank under accession numbers PZ383697-PZ383795. Viral sequences and gene annotations are also available on Zenodo (https://zenodo.org/records/20076790).

Code availability
Raw reads generated in this study can be found in the NCBI Sequence Read Archive under accession number PRJNA1068039. All supplementary figures, tables, data, sequence files in FASTA and GenBank format, and code used for this analysis (including the equations to calculate the abundance and diversity values) can be found at https://github.com/andrea-silverj/TickVirome_IT/tree/main and at https://drive.google.com/drive/folders/1f2aAg48uilMPqzI_Q_WL7d7WFUfU3aeL?usp=sharing. Viral sequences and gene annotations have been submitted to GenBank under accession numbers PZ383697-PZ383795. Viral sequences and gene annotations are also available on Zenodo (https://zenodo.org/records/20076790).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Accordance statement
All computational methods and analyses were performed with relevant guidelines and standards. Publicly available data were used in compliance with the respective database terms of use.

Author details
[1]Center Agriculture Food Environment (C3A), University of Trento, Trento, Italy
[2]CIBIO Department, University of Trento, Trento, Italy
[3]Research and Innovation Center, Fondazione Edmund Mach, San Michele all'Adige, Italy
[4]Department of Infectious Diseases, Istituto Superiore di Sanità, Rome, Italy
[5]Istituto Zooprofilattico Sperimentale della Sicilia "A. Mirri", Palermo, Italy
[6]NBFC, National Biodiversity Future Center, Palermo, Italy
[7]Present address: Department of Biological Sciences, University of Calgary, Calgary, Canada
[8]Present address: Human Technopole, Fondazione Human Technopole, Milan, Italy

Received: 31 October 2025 / Accepted: 1 June 2026
Published online: 16 June 2026

References

1. Wikel SK. Ticks and Tick-Borne Infections: Complex Ecology, Agents, and Host Interactions. Vet Sci China. 2018;5(2):60.
2. Dantas-Torres F. Climate change, biodiversity, ticks and tick-borne diseases: The butterfly effect. Int J Parasitol Parasites Wildl. 2015;4:452–61.
3. Gilbert L. The Impacts of Climate Change on Ticks and Tick-Borne Disease Risk. Annu Rev Entomol. 2021;66:373–88.
4. Abdullah S, Helps C, Tasker S, Newbury H, Wall R. Ticks infesting domestic dogs in the UK: a large-scale surveillance programme. Parasit Vectors. 2016;9:391.
5. Wisely SM, Glass GE. Advancing the Science of Tick and Tick-Borne Disease Surveillance in the United States. Insects. 2019;10(10):361.
6. Watts GS, Hurwitz BL. Metagenomic Next-Generation Sequencing in Clinical Microbiology. Clin Microbiol Newsl. 2020;42:53–9.
7. Giani AM, Gallo GR, Gianfranceschi L, Formenti G. Long walk to genomics: History and current approaches to genome sequencing and assembly. Comput Struct Biotechnol J. 2020;18:9–19.
8. Rabinowitz PM, Kock R, Kachani M, Kunkel R, Thomas J, Gilbert J, Wallace R, Blackmore C, Wong D, Karesh W, Natterson B, Dugas R, Rubin C, Stone Mountain One Health Proof of Concept Working Group. Toward proof of concept of a one health approach to disease prediction and control. Emerg Infect Dis. 2013;19(12):e130265.
9. Kazimírová M, Thangamani S, Bartíková P, Hermance M, Holíková V, Štibrániová I, Nuttall PA. Tick-Borne Viruses and Biological Processes at the Tick-Host-Virus Interface. Front Cell Infect Microbiol. 2017;7:339.
10. Ni X-B, Cui X-M, Liu J-Y, Ye R-Z, Wu Y-Q, Jiang J-F, Sun Y, Wang Q, Shum MH-H, Chang Q-C, Zhao L, Han X-H, Ma K, Shen S-J, Zhang M-Z, Guo W-B, Zhu J-G, Zhan L, Li L-J, Ding S-J, Zhu D-Y, Zhang J, Xia L-Y, Oong X-Y, Ruan X-D, Shao H-Z, Que T-C, Liu G-Y, Du C-H, Huang E-J, Wang X, Du L-F, Wang C-C, Shi W-Q, Pan Y-S, Zhou Y-H, Qu J-L, Ma J, Gong C-W, Chen Q-Q, Qin Q, Tick Genome and Microbiome Consortium (TIGMIC), Lam TT-Y, Jia N, Cao W-C. 2023. Metavirome of 31 tick species provides a compendium of 1,801 RNA virus genomes. Nat Microbiol 8:162–173.
11. Kong Y, Zhang G, Jiang L, Wang P, Zhang S, Zheng X, Li Y. Metatranscriptomics Reveals the Diversity of the Tick Virome in Northwest China. Microbiol Spectr. 2022;10:e0111522.
12. Gofton AW, Blasdell KR, Taylor C, Banks PB, Michie M, Roy-Dufresne E, Poldy J, Wang J, Dunn M, Tachedjian M, Smith I. Metatranscriptomic profiling reveals diverse tick-borne bacteria, protozoans and viruses in ticks and wildlife from Australia. Transbound Emerg Dis. 2022;69:e2389–407.
13. Brackney DE, Armstrong PM. Transmission and evolution of tick-borne viruses. Curr Opin Virol. 2016;21:67–74.
14. Cai X, Cai X, Xu Y, Shao Y, Fu L, Men X, Zhu Y. Virome analysis of ticks and tick-borne viruses in Heilongjiang and Jilin Provinces, China. Virus Res. 2023;323:199006.
15. Ye R-Z, Li Y-Y, Wang X-Y, Cao W-C. Tick-borne viruses: discovery, clinical aspects, influencing factors and control. Nat Rev Microbiol. 2026. https://doi.org/10.1038/s41579-026-01277-z.
16. Postler TS, Beer M, Blitvich BJ, Bukh J, de Lamballerie X, Drexler JF, Imrie A, Kapoor A, Karganova GG, Lemey P, Lohmann V, Simmonds P, Smith DB, Stapleton JT, Kuhn JH. Renaming of the genus Flavivirus to Orthoflavivirus and extension of binomial species names within the family Flaviviridae. Arch Virol. 2023;168:224.
17. Dobler G, Gniel D. A history of tick-borne encephalitis and its virus. History of Arbovirology: Memories from the Field. Cham: Springer Nature Switzerland; 2023. pp. 453–67.
18. Phipps LP, Johnson N. Tick-borne encephalitis virus. J Med Microbiol. 2022. https://doi.org/10.1099/jmm.0.001492.
19. Riccardi N, Antonello RM, Luzzati R, Zajkowska J, Di Bella S, Giacobbe DR. Tick-borne encephalitis in Europe: a brief update on epidemiology, diagnosis, prevention, and treatment. Eur J Intern Med. 2019;62:1–6.
20. Satapathy P, Kumar P, Chand K, Gahtori P, Rustagi S, Sah R, Neyazi A. 2023. The Rising Tide of Tick-borne Encephalitis across European Nations. QJM. https://doi.org/10.1093/qjmed/hcad226
21. Alfano N, Tagliapietra V, Rosso F, Ziegler U, Arnoldi D, Rizzoli A. Tick-borne encephalitis foci in northeast Italy revealed by combined virus detection in ticks, serosurvey on goats and human cases. Emerg Microbes Infect. 2020;9:474–84.
22. Kidanemariam D, Abraham A. Luteoviruses. Plant RNA Viruses. Elsevier; 2023. pp. 57–77.
23. Rosario K, Breitbart M, Harrach B, Segalés J, Delwart E, Biagini P, Varsani A. Revisiting the taxonomy of the family Circoviridae: establishment of the genus Cyclovirus and removal of the genus Gyrovirus. Arch Virol. 2017;162:1447–63.

24. Rezza G, Farchi F, Pezzotti P, Ruscio M, Lo Presti A, Ciccozzi M, Mondardini V, Paternoster C, Bassetti M, Merelli M, Scotton PG, Luzzati R, Simeoni J, Mian P, Mel R, Carraro V, Zanin A, Ferretto R, Francavilla E, TBE Virology Group. 2015. Tick-borne encephalitis in north-east Italy: a 14-year retrospective study, January 2000 to December 2013. Euro Surveill. 20.
25. Durbin R, De Sanctis B, Blumer M, Rotate. A command-line program to rotate circular DNA sequences to start at a given position or string. Wellcome Open Res. 2023;8:401.
26. Ayad LAK, Pissis SP. MARS: improving multiple circular sequence alignment using refined sequences. BMC Genomics. 2017;18(1):86.
27. Grossi R, Iliopoulos CS, Mercas R, Pisanti N, Pissis SP, Retha A, et al. Circular sequence comparison: algorithms and applications. Algorithms Mol Biol. 2016;11(1):12.
28. Moi D, Bernard C, Steinegger M, Nevers Y, Langleib M, Dessimoz C. Structural phylogenetics unravels the evolutionary diversification of communication systems in gram-positive bacteria and their viruses. Nat Struct Mol Biol. 2025;32(12):2492–502.
29. Ohlendorf V, Marklewitz M, Kopp A, Yordanov S, Drosten C, Junglen S. Huge diversity of phleboviruses in ticks from Strandja Nature Park, Bulgaria. Ticks Tick Borne Dis. 2019;10:697–703.
30. Klimentov AS, Belova OA, Kholodilov IS, Butenko AM, Bespyatova LA, Bugmyrin SV, Chernetsov N, Ivannikova AY, Kovalchuk IV, Nafeev AA, Oorzhak ND, Pilikova OM, Polienko AE, Purmak KA, Romanenko EN, Romanova LI, Saryglar AA, Solomashchenko NI, Shamsutdinov AF, Vakalova EV, Lukashev AN, Karganova GG, Gmyl AP. Phlebovirus sequences detected in ticks collected in Russia: Novel phleboviruses, distinguishing criteria and high tick specificity. Infect Genet Evol. 2020;85:104524.
31. Liu Q, He B, Huang S-Y, Wei F, Zhu X-Q. Severe fever with thrombocytopenia syndrome, an emerging tick-borne zoonosis. Lancet Infect Dis. 2014;14:763–72.
32. McMullan LK, Folk SM, Kelly AJ, MacNeil A, Goldsmith CS, Metcalfe MG, Batten BC, Albariño CG, Zaki SR, Rollin PE, Nicholson WL, Nichol ST. A new phlebovirus associated with severe febrile illness in Missouri. N Engl J Med. 2012;367:834–41.
33. Kartashov MY, Gladysheva AV, Shvalov AN, Tupota NL, Chernikova AA, Ternovoi VA, Loktev VB. Novel Flavi-like virus in ixodid ticks and patients in Russia. Ticks Tick Borne Dis. 2023;14:102101.
34. Li C-X, Shi M, Tian J-H, Lin X-D, Kang Y-J, Chen L-J, Qin X-C, Xu J, Holmes EC, Zhang Y-Z. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. Elife. 2015;29:4e05378.
35. Zhao T, Gong H, Shen X, Zhang W, Shan T, Yu X, Wang SJ, Cui L. Comparison of Viromes in Ticks from Different Domestic Animals in China. Virol Sin. 2020;35:398–406.
36. Dong Z, Yang M, Wang Z, Zhao S, Xie S, Yang Y, Liu G, Zhao S, Xie J, Liu Q, Wang Y. Human Tacheng Tick Virus 2 Infection, China, 2019. Emerg Infect Dis. 2021;27:594–8.
37. Qin T, Shi M, Zhang M, Liu Z, Feng H, Sun Y. Diversity of RNA viruses of three dominant tick species in North China. Front Vet Sci. 2022;9:1057977.
38. Tighe AJ, Ruane NM, Carlsson J. Potential origins of fish toti-like viruses in invertebrates. J Gen Virol. 2022. https://doi.org/10.1099/jgv.0.001775.
39. Vanmechelen B, Laenen L, Vergote V, Maes P. Grotenhout Virus, a Novel Nairovirus Found in Ixodes ricinus in Belgium. Genome Announc. 2017;5(21):e00288–17.
40. Pettersson JH-O, Shi M, Bohlin J, Eldholm V, Brynildsrud OB, Paulsen KM, Andreassen Å, Holmes EC. Characterizing the virome of Ixodes ricinus ticks from northern Europe. Sci Rep. 2017;7:10870.
41. Sameroff S, Tokarz R, Charles RA, Jain K, Oleynik A, Che X, Georges K, Carrington CV, Lipkin WI, Oura C. Viral Diversity of Tick Species Parasitizing Cattle and Dogs in Trinidad and Tobago. Sci Rep. 2019;9:10421.
42. Petrzik K, Lenz O. CHAPTER 6: Apple mosaic virus in Pome Fruits. Virus and Virus-Like Diseases of Pome and Stone Fruits. The American Phytopathological Society; 2011. pp. 25–8.
43. Ventura Spagnolo E, Mondello C, Roccuzzo S, Cardia L, Raffino C. A lethal Tick-Borne Encephalitis (TBE) due to TBE Virus in Sicily (Italy): a case of IgG+/IgM-. response? Clin Ter. 2018;169:e145–8.
44. Khamassi Khbou M, Romdhane R, Foughali AA, Sassi L, Suin V, Rekik M, et al. Presence of antibodies against tick-borne encephalitis virus in sheep in Tunisia, North Africa. BMC Vet Res. 2020;16(1):441.
45. Manilla G. Acari Ixodida. Fauna d'Italia. Bologna Calderini; 1998.
46. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 2011;17:10–2.
47. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
48. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.
49. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28:3211–7.
50. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, Prjibelsky A, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, McLean J, Lasken R, Clingenpeel SR, Woyke T, Tesler G, Alekseyev MA, Pevzner PA. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. Research in Computational Molecular Biology. Berlin Heidelberg: Springer; 2013. pp. 158–70.
51. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29:644–52.
52. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.
53. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
54. Brister JR, Ako-Adjei D, Bao Y, Blinkova O. NCBI viral genomes resource. Nucleic Acids Res. 2015;43:D571–7.
55. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
56. Hall TA, Others. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, pp. 95–98. In Nucleic acids symposium series. Oxford.
57. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. Gigascience. 10(2):giab008.
58. Nayfach S, Camargo AP, Schulz F, Eloe-Fadrosh E, Roux S, Kyrpides NC. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol. 2020;39:578–85.

59. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, Chain PSG, Nayfach S, Kyrpides NC. Identification of mobile genetic elements with geNomad. Nat Biotechnol. 2024;42:1303–12.
60. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
61. Haas B, Papanicolaou A, Yassour M, TransDecoder. 2017. https://github.com/TransDecoder/TransDecoder.
62. Wille M, Shi M, Hurt AC, Klaassen M, Holmes EC. RNA virome abundance and diversity is associated with host age in a bird species. Virology. 2021;561:98–106.
63. Wang J, Pan Y, Yang L-F, Yang W-H, Lv K, Luo C, Wang J, Kuang G, Wu W-C, Gou Q-Y, Xin G-Y, Li B, Luo H-L, Chen S, Shu Y, Guo D, Gao Z, Liang G, Li J, Chen Y-Q, Holmes E, Feng Y, Shi M. Individual bat virome analysis reveals co-infection and spillover among bats and virus zoonotic potential. Nat Commun. 2023;14:4079.
64. Wille M, Harvey E, Shi M, Gonzalez-Acuña D, Holmes EC, Hurt AC. Sustained RNA virome diversity in Antarctic penguins and their ticks. ISME J. 2020;14:1768–82.
65. Zhou H, Liu H, Wang Y-X, Zhao S, Guo M, Cui M, Chen J, Li J, Wang S, Zhang X, Qin J, Lv Z, Ma Z, Zhang R, Zhang B, Zheng W, Ji J, Wang T, Holmes EC, Wang Y-Z, Shi W. The risk of human- and mammal-infecting tick-borne viruses in northwest China and adjacent countries. Nat Commun. 2025;17:175.
66. Gilchrist CLM, Chooi Y-H. clinker & clustermap.js: automatic generation of gene cluster comparison figures. Bioinformatics. 2021;37:2473–5.
67. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3.
68. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3.
69. RStudio T. RStudio: Integrated Development for R. RStudio. Boston, MA: PBC; 2020. http://www.rstudio.com/.
70. Harvey E, Rose K, Eden J-S, Lo N, Abeyasuriya T, Shi M, Doggett SL, Holmes EC. Extensive Diversity of RNA Viruses in Australian Ticks. J Virol. 2019;93(3):10–128.
71. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493–500.
72. Kim RS, Levy Karin E, Mirdita M, Chikhi R, Steinegger M. BFVD-a large repository of predicted viral protein structures. Nucleic Acids Res. 2025;53(D1):D340–7.
73. Sillitoe I, Lewis TE, Cuff A, Das S, Ashford P, Dawson NL, et al. CATH: comprehensive structural and functional annotations for genome sequences. Nucleic Acids Res. 2015;43(Database issue):D376–81.
74. Vallat B, Rose Y, Piehl DW, Duarte JM, Bittrich S, Bi C, et al. RCSB Protein Data Bank: Delivering integrative structures alongside experimental structures and computed structure models. Nucleic Acids Res. 2026;54(D1):D489–98.
75. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, et al. Fast and accurate protein structure search with Foldseek. Nat Biotechnol. 2024;42(2):243–6.
76. Gilchrist CLM, Mirdita M, Steinegger M. Multiple protein structure alignment at scale with FoldMason. Science. 2026;391(6784):485–8.
77. Schwaiger M, Cassinotti P. Development of a quantitative real-time RT-PCR assay with internal control for the laboratory detection of tick borne encephalitis virus (TBEV) RNA. J Clin Virol. 2003;27:136–45.
78. Farouni R, Djambazian H, Ferri LE, Ragoussis J, Najafabadi HS. Model-based analysis of sample index hopping reveals its widespread artifacts in multiplexed single-cell RNA-sequencing. Nat Commun. 2020;11:2704.
79. Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, Manghi P, Dubois L, Huang KD, Thomas AM, Nickols WA, Piccinno G, Piperni E, Punčochář M, Valles-Colomer M, Tett A, Giordano F, Davies R, Wolf J, Berry SE, Spector TD, Franzosa EA, Pasolli E, Asnicar F, Huttenhower C, Segata N. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nat Biotechnol. 2023;41:1633–44.

Publisher's Note