



OPEN

# European primary datasets of alien bacteria and viruses

DATA DESCRIPTOR

Chiara Magliozzi<sup>1</sup>✉, Marc Artois<sup>2</sup>, Assunta Bertaccini<sup>3</sup>, Thierry Candresse<sup>4</sup>, Konstantinos Tsiamis<sup>1</sup>, Fabio D'Amico<sup>1</sup>, Ivan Deriu<sup>1</sup>, Eugenio Gervasini<sup>1</sup> & Ana Cristina Cardoso<sup>1</sup>

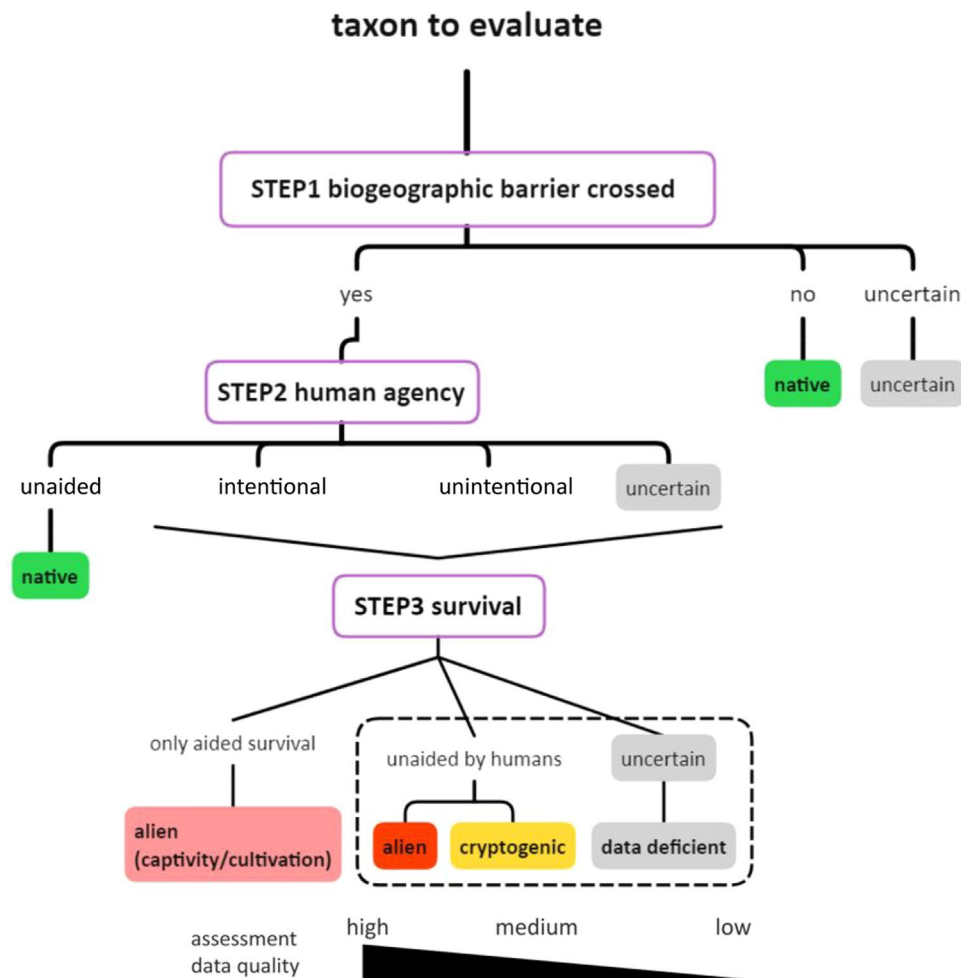
Bacteria and viruses are a natural component of Earth biodiversity and play an essential role in biochemical and geological cycles. They may also pose problems outside their native range, where they can negatively impact on natural resources, wildlife, and human health. To address these challenges and develop sustainable conservation strategies, a thorough understanding of their invasion related-factors is needed: origin, country and year of introduction, and pathways dynamics. Yet, alien bacteria and viruses are underrepresented in invasion ecology studies, which limits our ability to quantify their impacts and address future introductions. This study provides primary datasets of alien bacteria and viruses of plants and animals present in the European environment. The datasets contain expert-revised data on 446 taxa and their invasion related- factors across terrestrial and aquatic environments. Taxa information are complemented with spatial occurrences. The datasets provide a basis for collaborative initiatives to improve the collection of alien bacteria and viruses' data, and a starting point for data-driven conservation practices.

## Background & Summary

Bacteria and viruses are microscopic organisms inhabiting virtually every ecosystem on Earth<sup>1</sup>. They constitute the greatest biodiversity on the planet and play an essential role in global biochemical and geological cycles<sup>2,3</sup>. Despite being an indispensable part of life, bacteria and viruses may cause a range of problems, such as diseases affecting both the health and resources available to e.g., plants, animals, and humans<sup>4-6</sup>. To develop adequate biosecurity and management strategies, it is necessary to know bacteria and viruses' distributions and link their dispersal dynamics to their region of origin, vectors and impacts on the social-ecological system.

In the last decades, hundreds of bacteria and viruses were introduced in Europe from their native range and called alien species. These introductions were the result of the increasing global transport of commodities, international trade of wildlife species or of their use as biological control agents<sup>7-10</sup>. The alien species (AS) which can establish self-sustaining populations, spread and impact negatively on biodiversity at different biological levels<sup>11-13</sup> (from single organism to population), and on sustainable exploitation of biodiversity<sup>13</sup> are called invasive alien species (IAS). In Europe, priority IAS that are introduced accidentally or deliberately into a natural environment are regulated by the Regulation (EU) 1143/2014/EC<sup>14</sup> (hereafter IAS Regulation), which establishes concerted actions at European Union level to prevent, minimise and mitigate their effects on biodiversity. Invasive alien bacteria and viruses are not considered in the scope of the IAS Regulation, either because they are listed in other legal acts (e.g. Regulation (EU) 2016/2031<sup>15</sup>) or targeted by the human health research<sup>16</sup>. Except for pathogenic (disease-causing) bacteria and viruses, which are the focus of public health epidemiology<sup>17</sup> (e.g. European Food Safety Authority (EFSA), European Centre for Disease Prevention and Control (ECDC)), the processes of introduction-establishment-invasion of alien (also non-pathogenic) microorganisms have not been yet well investigated, and as a result these groups are poorly represented in alien and invasive alien species databases<sup>18</sup>. In this context, there are three key issues for data paucity on bacteria and viruses. First, their identification is often complex. Unlike many plants and animals, bacteria and viruses are microorganisms difficult to detect without analytical approaches using culture-based or molecular tools<sup>19</sup>. In addition, viruses can be principally detected after recognition of a disease signs (morbidity and mortality)<sup>20</sup>. A second issue is the determination of the organism's alien status. In invasion science, the definition of alien is based on the evaluation of three main criteria: i) biogeographic barriers, ii) human agency (*i.e.* intentional,

<sup>1</sup>European Commission, Joint Research Centre (JRC), Ispra, Italy. <sup>2</sup>LISAE (Lorraine investigation in animal and environmental health), Lagny, France. <sup>3</sup>Alma Mater Studiorum, University of Bologna, Bologna, Italy. <sup>4</sup>Univ. Bordeaux, INRAE, UMR 1332 BFP, Villenave d'Ornon, France. ✉e-mail: [chiara.magliozzi@ec.europa.eu](mailto:chiara.magliozzi@ec.europa.eu)



**Fig. 1** Conceptual scheme to assess the alien status of a taxon. It considers three assessment criteria, *i.e.* biogeographic barriers, human agency, and survival, and the levels of uncertainties of data quality (modified from Essl *et al.*<sup>10</sup>).

Group	Alien	Cryptogenic	Data Deficient	Total
bacteria	55	17	21	93
viruses	48	19	286	353

**Table 1.** Number of bacteria and viruses of plant and animals reviewed by experts. The status is assigned according to Fig. 1.

unintentional introduction), and iii) survival after introduction<sup>10</sup>. The location and strength of biogeographic barriers is linked to the species or species' carrier dispersal potential (*e.g.* host, reservoir and vectors) and ecological requirements, and for bacteria and viruses baseline inventories are limited especially across taxonomic realms<sup>18,21,22</sup>. As a result, diagnostic tools such as molecular and phylogenetic analysis are essential to refine the hypothesis of the geographical origin of bacteria and viruses<sup>23</sup>. Finally, it is often a challenge to assign impact to alien bacteria and viruses. They have shorter generation times than plant or animals, and their demographic processes have been little studied in wildlife populations<sup>24</sup>. In this paper, we assembled expert-revised and harmonised datasets of alien bacteria and viruses filling the gap in European biodiversity resources. The datasets give unique insights into taxon traits, invasion-related factors and location characteristics<sup>25</sup> (*i.e.* taxonomy, year and country of first introduction, native distribution, pathways of introduction, and environment), and invasion processes (*i.e.* introduction, establishment and invasion), starting from the recognition that invaders' traits interact with invaded communities allowing or preventing invasions. The datasets are not comprehensive of all alien bacteria and viruses in Europe, but focus on those known to affect wild, free-living species of plant and animals. We envision the datasets as an on-going collaborative initiative for easily sharing invasion-related data about alien bacteria and viruses, avoiding duplication of efforts, and as a first step toward knowledge and control

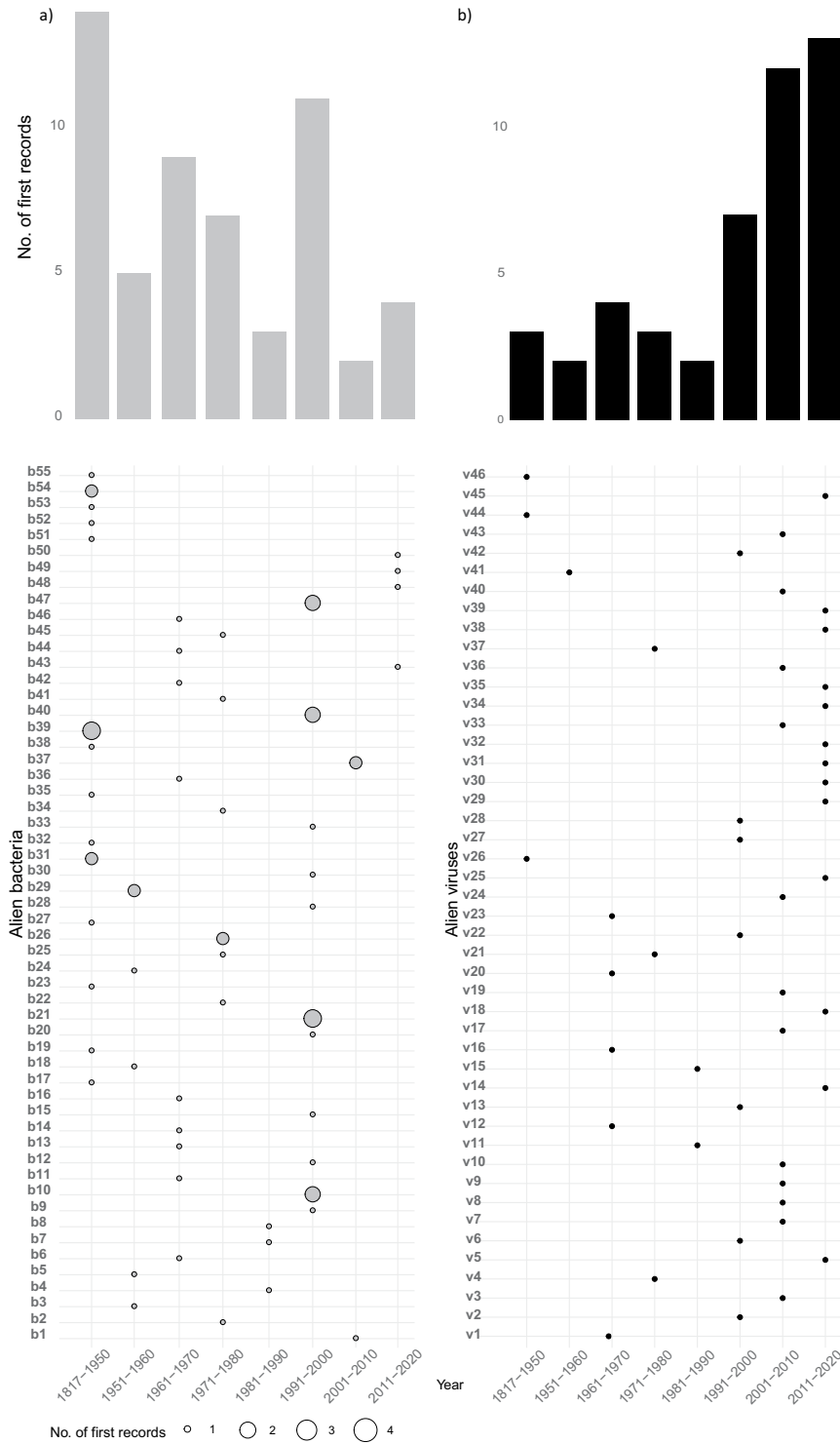
Database name	Description	Web source link
Centre for Agriculture and Bioscience International (CABI)- Invasive Species Compendium <sup>43</sup>	bacteria, virus	<a href="#">CABI Invasive Species Compendium</a>
European and Mediterranean Plant Protection Organization (EPPO)-EPPO-Q-bank database International Phytoplasma Working Group (IPWG)-Phytoplasma Collection	bacteria, plant pests, phytoplasmas	<a href="#">EPPO-Q-bank database</a> ; <a href="#">Phytoplasma Collection</a>
World Organisation for animal health (OIE)- WAHIS database	animal pathogens	<a href="#">WAHIS database</a>
Plant Viruses Database –DPV <sup>44,45</sup>	virus	<a href="#">DPV database</a>
Virus Identification Data Exchange (VIDE) project database	virus	<a href="#">VIDE database</a>
Genbank database <sup>46</sup>	virus	<a href="#">integrated through Pubmed</a>
International Committee for the Taxonomy of Viruses (ICTV)	virus	<a href="#">ICTV</a>
National Collection of Plant Pathogenic Bacteria (NCPBB)	bacteria	<a href="#">NCPBB</a>

**Table 2.** Examples of bibliographic repositories on bacteria and viruses reviewed by experts.

Field	Description
SPECIES SCIENTIFIC NAME	Species scientific name
STATUS	Status of the species, A for Alien, C for Cryptogenic, DD for Data Deficient <sup>10</sup>
ORIGIN	Origin of the terrestrial/freshwater/marine species (Country or Continent). (see Supplementary File 1)
YEAR OF FIRST INTRODUCTION IN EUROPE	Must contain a numeric value (e.g. 2020). For BC dates a minus must be used (e.g. –1500). It is the year of the first observation of an alien species in Europe. It is used as the best available estimate of the year of its initial introduction when the latter could not be determined with certainty, based on a thorough review of the scientific and grey literature.
COUNTRY OF FIRST INTRODUCTION IN EUROPE	ISO 3166-1 Alpha-2 code of European or neighbouring Country of first introduction (e.g. IT for Italy) including the outermost regions (see Supplementary File 1)
REFERENCE	reference to the publication for the first introduction of the species
DOI	Digital Object Identifier of the publication
NOTES	note about the species introduction
TERRESTRIAL	Value 1 if the species is terrestrial.
FRESHWATER	Value 1 if the species is freshwater.
MARINE	Value 1 if the species is marine.
OLIGOHALINE	Value 1 if the species is estuarine.
PARASITE	Value 1 if the species is parasite
KINGDOM	Taxonomic kingdom of the species
PHYLUM	Taxonomic phylum of the species
CLASS	Taxonomic class of the species
GENOME	Genome composition (optional)
ORDER	Taxonomic order of the species
FAMILY	Taxonomic family of the species
GENUS	Taxonomic genus of the species
REFERENCE	reference to the nomenclature used (e.g. ITIS, EPPO)
PARTLY NATIVE	Value 1 if the species is partly native in Europe.
COUNTRY/Marine Strategy Framework Directive (MSFD) REGION	Value 1 if the species is partly native in such Country or MSFD Region. See Supplementary File 1
CBD 2014 PATHWAYS	Value P_HIGH: primary pathway of introduction with high level of uncertainty.
	Value P_MEDIUM: primary pathway of introduction with medium level of uncertainty.
	Value P_LOW: primary pathway of introduction with low level of uncertainty.
	Value P: primary pathway of introduction with unknown level of uncertainty.
	Value S: secondary pathway of introduction

**Table 3.** Invasion- related factors and location characteristics of the bacteria and viruses datasets. Fields and description of the collected information for each species of bacteria and viruses.

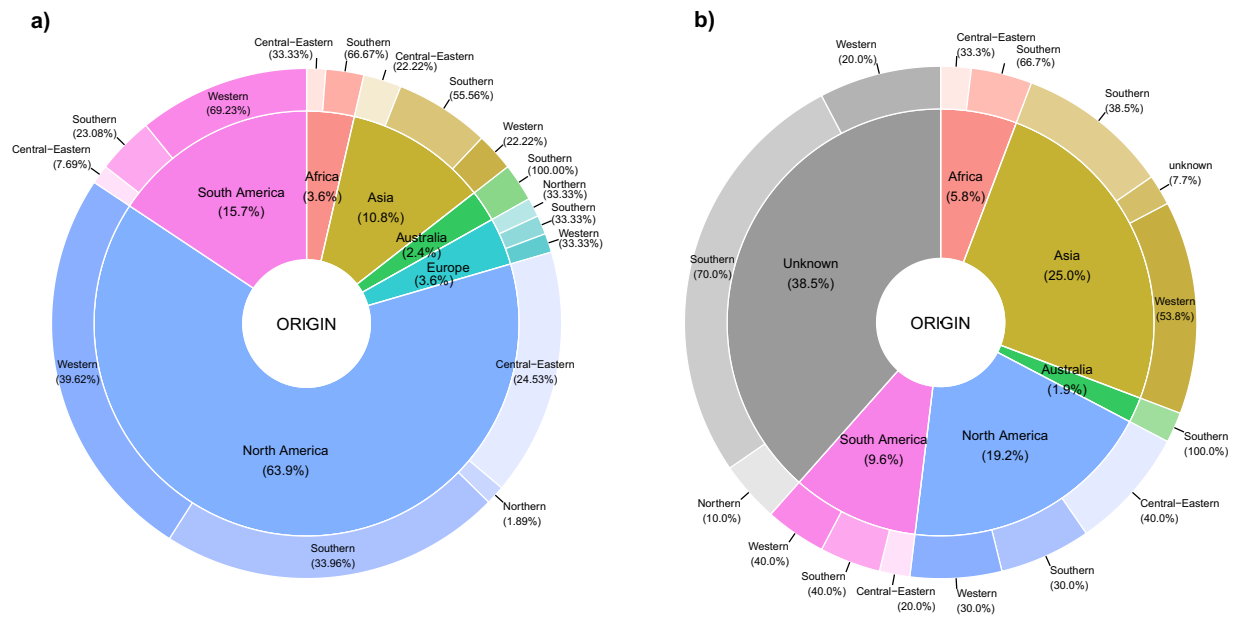
of alien pathogens in wild species. The datasets can also be of use in epidemiology, and particularly on zoonoses associated with wildlife, and more broadly by considering alien bacteria and viruses in the practice of invasion science to provide evidence in support of conservation objectives. Finally, we ensure their access openly through the European Alien Species Information Network (EASIN).



**Fig. 2** Temporal distribution of first records of alien bacteria (a) and viruses (b) in European countries (n = 28). Point sizes represent the number of records per species and time period. The Grapevine red blotch virus and the Apple scar skin viroid are not included in the list as the year or the country of introduction in Europe are unknown. See Table 3 for species scientific names and codes in Supplementary File 1.

**Methods**

**Alien status definition.** The definition of alien species is controversial, and a range of definitions have been put forward over the years<sup>26–29</sup>. For example, in public health epidemiology alien and pathogenic species are named emerging infectious diseases (EIDs<sup>30</sup>) and refer to several mechanisms that allow pathogens to be transmissible to humans, *i.e.* adaptive (genetic changes allowing host jump from animals to humans) and geographical (long-distance or localized translocation). In this paper, we followed the scheme proposed by Essl *et*

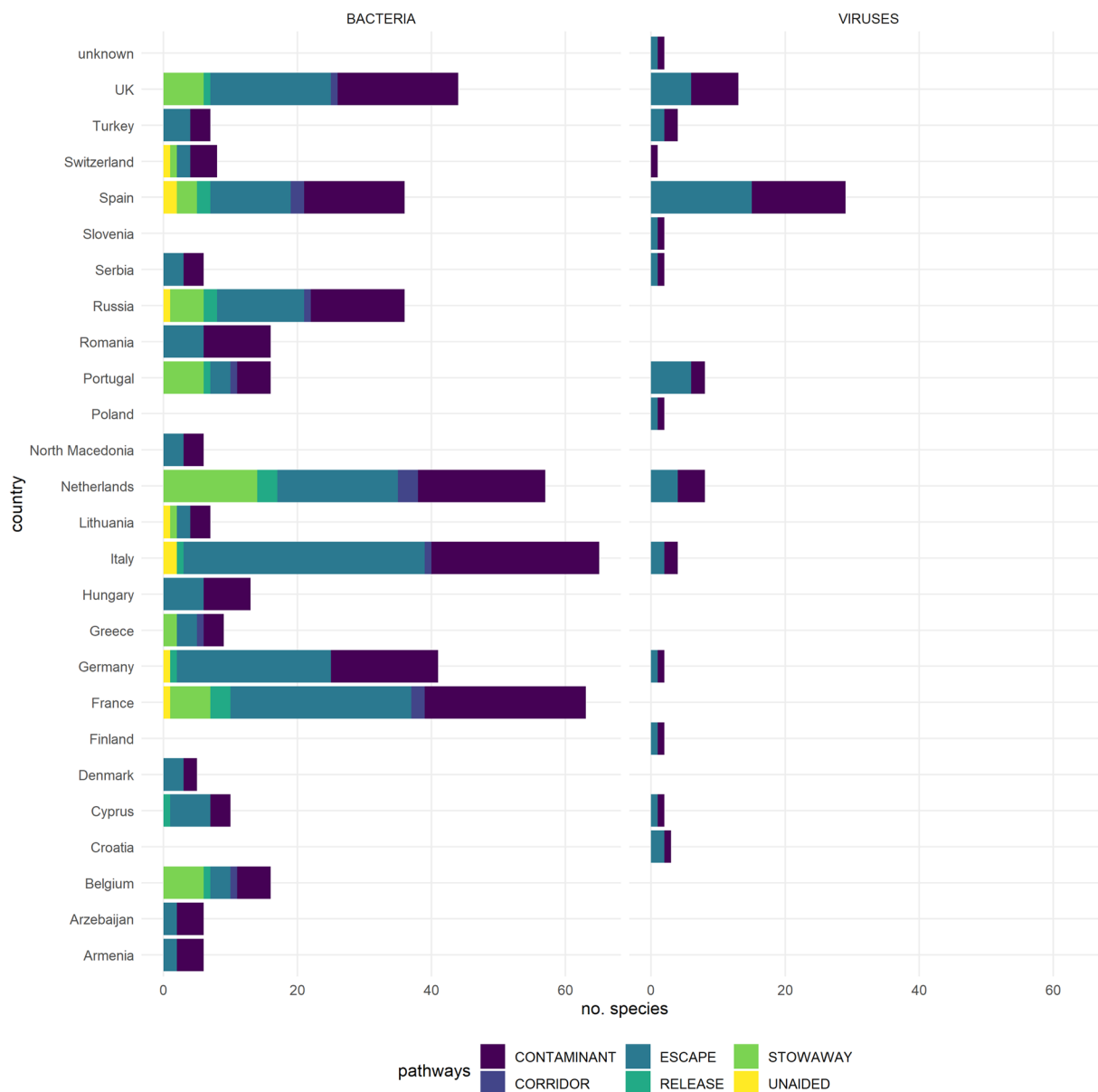


**Fig. 3** Distribution of alien bacteria (a) and viruses (b) in European regions according to their origin. The internal pie refers to the Continents of origin (Europe, North America, South America, Asia, Africa, Australia), while the outer circular crown to the European regions invaded. The European regions are grouped into Central and Eastern Europe, Northern Europe, Southern Europe, and Western Europe (EuroVoc). Unknown refers to taxa of unknown origin.

*al.*<sup>10</sup> and applied it to Europe (27 EU Member States, 5 Candidate countries to the European Union, and 32 other neighbouring countries) (Fig. 1). Alien taxa are species or infra-species assessed through a set of criteria, *i.e.* biogeographic barriers, human agency, and survival, considering uncertainties. Alien species as addressed here have crossed biogeographic barriers by direct (intentional) or indirect (unintentional) human agency and survived in the wild (Fig. 1). Following the methodology of Essl *et al.* (2018), the scientific evidence in support of these criteria must be reliable and robust (low uncertainty). Low uncertainty is achieved when the data are of good quality and the assessor estimates a probability of correct classification greater than 90% (Fig. 1). We operationalized the alien species definition by including partly alien species in Europe, a sub-group of the alien taxa with a native distribution in at least one European country (administrative boundaries) but alien in other(s). Our datasets include two additional categories: cryptogenic, *i.e.* species of uncertain alien status, and data deficient, *i.e.* species for which an assessment of the alien status is currently unfeasible because of the lack of data<sup>10</sup> (Fig. 1, Table 1). Finally, the terms taxon traits, invasion-related factors and location characteristics were adopted from the terminology of Pysek *et al.*<sup>25</sup> to refer to the species taxonomy, year and country of first introduction, native distribution, introduction pathways, and environment for alien, cryptogenic and data-deficient taxa.

**Data collection.** Data were retrieved by plants and animal pathologists with expertise on epidemiological surveillance or diagnostics and characterization, including molecular plant/animal-virus interaction and metagenomics. Sources encompassed international scientific and grey literature, disease surveillance databases, repositories, and search engines relevant to epizootiology or epidemiology, ecology and invasion science (Table 2). Additional sources of information are available in the static version of the datasets as CSV format.

For each species of plant and animal bacteria and viruses the following information is provided: i) valid/accepted nomenclature and taxonomy, ii) the year and country of first introduction, iii) native distribution, iv) primary introduction pathways, iv) environment, and v) spatial reference (Table 3). The taxonomic nomenclature follows the approaches of the Integrated Taxonomic Information System (ITIS) and the European and Mediterranean Plant Protection Organization (EPPO) for bacteria, and the International Committee for the Taxonomy of Viruses' (ICTV) for viruses. When assigning taxonomy to invasive pathogens, it was not always possible or relevant to address the species level and the 'strains' (infra-species taxon) were considered. For phytopathogenic bacteria known as phytoplasmas and liberibacter, the provisional category '*Candidatus* Phytoplasma/*Candidatus* Liberibacter'<sup>31</sup> was used. The year and country of first introduction refers to the first observation of the alien species, and it is often used the name of the first author and the year of publication associated with the source's primary citation (e.g. bibliographic reference) (Fig. 2). The native distribution data ranges from continent to sub-continent level (administrative boundaries) (Fig. 3), and species with part of their native range in a European country are tagged as partly alien. The first arrival of an alien species in Europe was documented according to the CBD (2014)<sup>32</sup> pathway categorization scheme: contaminant, corridor, escape, release, stowaway, unaided (Fig. 4, Table 3, Table 1 in Supplementary File 1). The pathways refer to both the host and to the pathogen without the host assigned. A certainty score was given to each assigned refer pathway for every species. The high certainty (P HIGH) score indicates direct evidence of a pathway. The species is clearly

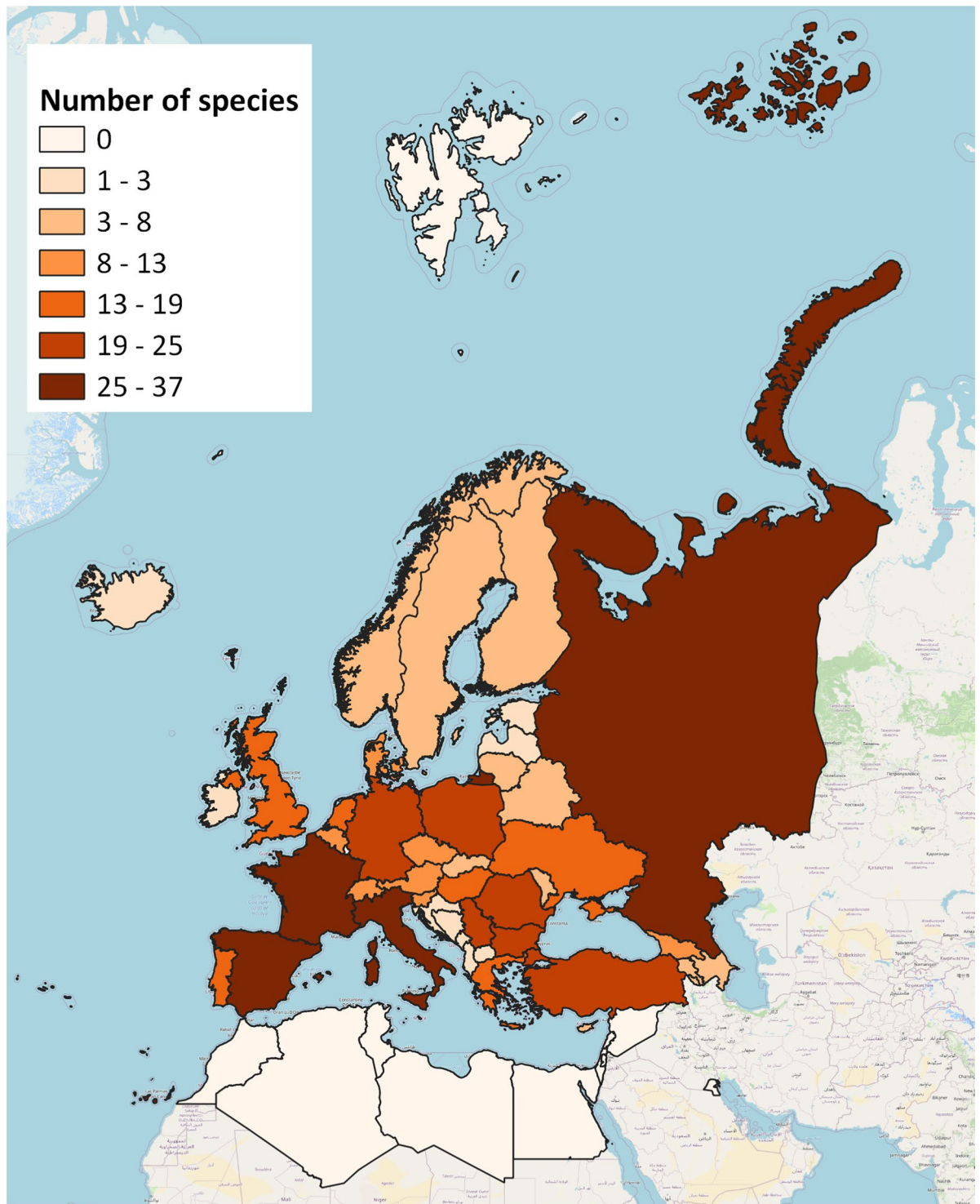


**Fig. 4** Pathways of first introduction in Europe. Number of species of alien bacteria and viruses and relevant introduction pathway, i.e., contaminant, corridor, escape, release, stowaway, and unaided (Supplementary Table 1), across European countries.

associated to a pathway at the time of introduction to a given locality, and it is often the case of intentional introductions. The medium certainty (P MEDIUM) denotes pathways that can be likely inferred, e.g., the taxon appears for the first time in a locality where a pathway is known to operate. In many cases inference is based on known examples of introductions elsewhere for the same or similar species, the biology and ecology of the species, the habitats and locales it occupies in both the native and introduced range, and its pattern of dispersal (if known). The low certainty (P LOW) is used when not reliable pathways can be assigned to the alien taxon. Inference is based on the activities in the locality where the species was found and may include evidence on similarly behaving species reported elsewhere (Table 3). In cases of unknown level of uncertainty, only the pathway was indicated (P). Finally, taxa spatial records are provided to foster examining the relationship between host and pathogens at different scale, the so-called ‘epidemiological surveillance’<sup>33</sup>.

### Data Records

The static version of the bacteria and viruses’ datasets (Tables 3, 4), as presented in this data descriptor, can be downloaded as CSV format from *figshare*<sup>34</sup>. Two types of data are available for download: a dataset containing the information on the species taxonomy, origin, date of first introduction and pathways, and a dataset with species’ records at country scale. These primary datasets of alien bacteria and viruses were also integrated into the EASIN databases and services<sup>35</sup>, which ensure long-term persistence and preservation of data to support



**Fig. 5** Example of coverage of geo-referenced records of bacteria with alien status across Europe (EPSG: 3857-WGS84/Pseudo Mercator, Contains Basemap: © OpenStreetMap).

research communities on AS and research for policy. EASIN collects, manages and stores alien species data across Europe and makes them available through web tools for data analysis, open sharing and reporting. The quality of data inputs in EASIN is assured by thematic specialists<sup>36</sup> (i.e. Editorial Board members and experts). Before being stored in the EASIN databases, the data is cleansed and standardised (spatial data accordingly to the Darwin Core Standard<sup>37,38</sup>). Data is released under the CC-BY licence enabling reuse with attribution to the original sources. The information on the bacteria and viruses' invasion-related factors and location characteristics is stored in the [EASIN Catalogue database](#) (version 9.1), which currently hosts approximately 14,000 alien taxa in Europe. As an evolving data product, updates and corrections to bacteria and viruses' data will

Field	Description
SPECIES SCIENTIFIC NAME	Species scientific name
COUNTRY	ISO 3166-2 standard for occurrences at country scale
GRID	EEA reference code for 10 km × 10 km grid
YEAR	Year of the observation of the species
REFERENCE	reference to the publication for the species observation

**Table 4.** Spatial occurrences of the bacteria and viruses datasets. Fields and description of the collected information for each species of bacteria and viruses.

occur with the following versions of the EASIN Catalogue. Versions are labelled using semantic versioning to indicate the introduced changes. The spatial data comprises polygon features at country scale (International Organization for Standardization (ISO) 3166-2) and 10 km × 10 km grid cells (EEA reference system) (Table 4, Fig. 5, and Supplementary Tables 3 and 4).

**Data coverage.** The number of bacteria and viruses (as of December 2021) is 446 including alien, cryptogenic and data deficient taxa. The main pathway of introduction is contaminant and escape (Fig. 4) for both bacteria (e.g. around 50% in Italy, Germany, France) and viruses (i.e. all taxa). Bacteria were also found to be introduced as stowaway (e.g. 25% in Netherlands), corridor (2–6%), release (2–6%) and unaided (2–5%). The rate of first records of bacteria in Europe decreased on average from 7.6 new records/year over 40 year period (1950–1990) to 5.6 records/year in 1991–2020 (Fig. 2a). For viruses, the rate of first records of introduction increased on average from 2.8 records/year (1950–1990) to 10.6 records/year in 1991–2020 (Fig. 2b). Finally, bacteria alien to Europe mainly originated from North and South America (around 80% taxa) and Americas and Asia for viruses. Information is still lacking for about 38% of viruses taxa.

### Technical Validation

The high-quality data on invasion-related factors and location characteristics is ensured by three strategies. First, invasion data were collected and overseen by experts from peer-reviewed scientific literature and intergovernmental organisations (e.g. OIE<sup>30</sup>, EFSA). Second, species occurrences were benchmarked against country-level information to ensure that the year of first introduction and later spread corresponded to governmental notification systems for pests and diseases (e.g. EPPO) and scientific findings (i.e. geographical records added to the taxa invasion-related factors).

Finally, a channel for user feedback is provided, where reporting of issues on the quality of single species information and occurrences is encouraged by using the EASIN Editorial Board<sup>36</sup>. This is a forum-like web interface allowing taxonomic experts to analyse registered users' enquiries on species data and highlight records for revisions.

### Usage Notes

The datasets are available at *figshare*<sup>34</sup> for download as compressed folder containing the static.CSV files for each group, i.e. bacteria and viruses. Species data can be accessed through the EASIN Catalogue and the 'EASIN Species Search' function which allows filtering single or groups of species by taxonomy, pathway, alien status and impact. Species occurrences can be mapped and downloaded from the EASIN Geodatabase using the 'EASIN Species Mapper' function which enables mapping species occurrences as polygons at country scale or at 10 km × 10 km cells. This tool allows also selecting occurrences for time ranges, and exporting the map in.PDF format or download it in WKT (EPSG:3035, ETRS89) or GeoJSON (EPSG:4326, WGS84) formats.

### Code availability

No code is used in this study. Figures were produced using R statistical software<sup>39</sup> and packages<sup>40–42</sup>.

Received: 20 January 2022; Accepted: 16 June 2022;

Published online: 13 July 2022

### References

- Brandes, N. & Linial, M. Giant viruses—big surprises. *Viruses* **11**, 404 (2019).
- Jover, L. F., Efler, T. C., Buchan, A., Wilhelm, S. W. & Weitz, J. S. The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nat. Rev. Microbiol.* **12**, 519–528 (2014).
- Madsen, E. L. Microorganisms and their roles in fundamental biogeochemical cycles. *Curr. opinion biotechnology* **22**, 456–464 (2011).
- Gummow, B. Challenges posed by new and re-emerging infectious diseases in livestock production, wildlife and humans. *Livest. Sci.* **130**, 41–46 (2010).
- Becker, D. J., Streicker, D. G. & Altizer, S. Linking anthropogenic resources to wildlife–pathogen dynamics: a review and meta-analysis. *Ecol. letters* **18**, 483–495 (2015).
- Woolhouse, M. E. & Gowtage-Sequeria, S. Host range and emerging and reemerging pathogens. *Emerg. infectious diseases* **11**, 1842 (2005).
- Foster, R. *et al.* Pathogens co-transported with invasive non-native aquatic species: implications for risk analysis and legislation. *NeoBiota* **69**, 79–102 (2021).
- Brasier, C. The biosecurity threat to the uk and global environment from international trade in plants. *Plant Pathol.* **57**, 792–808 (2008).
- Ruiz, G. M. *et al.* Global spread of microorganisms by ships. *Nature* **408**, 49–50 (2000).



10. Essl, F. *et al.* Which taxa are alien? criteria, applications, and uncertainties. *BioScience* **68**, 496–509 (2018).
11. Blackburn, T. M., Bellard, C. & Ricciardi, A. Alien versus native species as drivers of recent extinctions. *Front. Ecol. Environ.* **17**, 203–207 (2019).
12. Hawkins, C. L. *et al.* Framework and guidelines for implementing the proposed iucn environmental impact classification for alien taxa (eicat). *Divers. Distributions* **21**, 1360–1363 (2015).
13. Corrales, X. *et al.* Advances and challenges in modelling the impacts of invasive alien species on aquatic ecosystems. *Biol. Invasions* **22**, 907–934 (2020).
14. Regulation, E. Regulation (eu) no 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species. *Off. J. Eur. Union* **57**, 35–55 (2014).
15. EU. Regulation (eu) 2016/2031 of the European Parliament of the Council of 26 October 2016 on protective measures against pests of plants, amending regulations (eu) 228/2013, (eu) 652/2014 and (eu) 1143/2014 and repealing council directives 69/464/eec, 74/647/eec, 93/85/eec, 98/57/ec, 2000/29/ec, 2006/91/ec and 2007/33/ec. *Off. J.* **317**, 4–104 (2016).
16. Murtaugh, M. P. *et al.* The science behind one health: at the interface of humans, animals, and the environment. *Tech. Rep.* (2017).
17. Ogden, N. H. *et al.* Emerging infectious diseases and biological invasions: a call for a one health collaboration in science and management. *Royal Soc. open science* **6**, 181577 (2019).
18. Roy, H. E. *et al.* Alien pathogens on the horizon: Opportunities for predicting their threat to wildlife. *Conserv. Lett.* **10**, 477–484 (2017).
19. Ikner, L. A., Gerba, C. P. & Bright, K. R. Concentration and recovery of viruses from water: a comprehensive review. *Food Environ. Virol.* **4**, 41–67 (2012).
20. Taylor, M. W. Introduction: A short history of virology. In *Viruses and Man: A History of Interactions*, 1–22 (Springer, 2014).
21. Thakur, M. P., Van der Putten, W. H., Cobben, M. M., van Kleunen, M. & Geisen, S. Microbial invasions in terrestrial ecosystems. *Nat. Rev. Microbiol.* **17**, 621–631 (2019).
22. Desprez-Loustau, M.-L. *et al.* The fungal dimension of biological invasions. *Trends ecology & evolution* **22**, 472–480 (2007).
23. Rivett, D. W. *et al.* Elevated success of multispecies bacterial invasions impacts community composition during ecological succession. *Ecol. Lett.* **21**, 516–524 (2018).
24. Dunn, A. M. & Hatcher, M. J. Parasites and biological invasions: parallels, interactions, and control. *TRENDS Parasitol.* **31**, 189–199 (2015).
25. Pyšek, P. *et al.* Macroecological framework for invasive aliens (mafia): disentangling large-scale context dependence in biological invasions. (2020).
26. Hulme, P. E. *et al.* Blurring alien introduction pathways risks losing the focus on invasive species policy. *Conserv. Lett.* **10**, 265–266 (2017).
27. Gilroy, J. J., Avery, J. D. & Lockwood, J. L. Seeking international agreement on what it means to be “native”. *Conserv. Lett.* **10**, 238–247 (2017).
28. Webber, B. L. & Scott, J. K. Rapid global change: implications for defining natives and aliens. *Glob. Ecol. Biogeogr.* **21**, 305–311 (2012).
29. CBD Secretariat. Decision VI/23: Alien species that threaten ecosystems, habitats and species. Document UNEP/CBD/COP/6/23 (2002).
30. World Health Organization. A brief guide to emerging infectious diseases and zoonoses. *Tech. Rep.* <https://apps.who.int/iris/handle/10665/204722> (2014).
31. Firrao, G. *et al.* Candidatus phytoblastus, a taxon for the wall-less, non-helical prokaryotes that colonize plant phloem and insects. *Int. J. Syst. Evol. Microbiol.* **54**, 1243–1255 (2004).
32. CBD. Pathways of introduction of invasive species, their prioritization and management (Secretariat of the Convention on Biological Diversity Montreal, 2014).
33. OIE. *Terrestrial Animal Health Code 2021* (OIE, 2021).
34. Magliozzi, C. *et al.* bacteria and viruses traits and species-related factors. *figshare* <https://doi.org/10.6084/m9.figshare.18550907.v2> (2022).
35. Katsanevakis, S. *et al.* Implementing the European policies for alien species: networking, science, and partnership in a complex environment. *Manag. Biol. Invasions* **4**, 3–6 (2013).
36. Tsiamis, K. *et al.* The EASIN Editorial Board: quality assurance, exchange and sharing of alien species information in europe. *Manag. Biol. Invasions* **7**, 321–328 (2016).
37. Wieczorek, J. *et al.* Darwin core: an evolving community-developed biodiversity data standard. *PloS one* **7**, e29715 (2012).
38. Darwin Core. Darwin Core quick reference guide. <https://dwc.tdwg.org/terms/> (2018).
39. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/> (2020).
40. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*, <https://ggplot2.tidyverse.org> (Springer-Verlag New York, 2016).
41. Shkarzl, T. *ggBubbles: Mini Bubble Plots for Comparison of Discrete Data with 'ggplot2'* R package version 0.1.4 (2019).
42. Moon, K. *R statistics and graphs for medical papers* (Hannarae Seoul, 2015).
43. Current, C. *Invasive species compendium*. Wallingford, UK: CAB Int. Available online: [www.cabi.org/isc](http://www.cabi.org/isc) (accessed on 19 August 2020) (2011).
44. Adams, M. J. & Antoniw, J. F. Dpvweb: An open access internet resource on plant viruses and virus diseases. *Outlooks on Pest Manag.* **16**, 268 (2005).
45. Adams, M. J. & Antoniw, J. F. Dpvweb: a comprehensive database of plant and fungal virus genes and genomes. *Nucleic acids research* **34**, D382–D385 (2006).
46. Benson, D. A. *et al.* Genbank. *Nucleic acids research* **41**, D36–D42 (2012).

## Acknowledgements

A.B., T.C. and M.A. were supported by the EU-funded contracts numbers C720026.X0, CT-EX2016D292351-101, CT-EX2020D401430-101, respectively. I.D. was supported by the Administrative Arrangement (AA) no. 070202/2020/831489/AA/ENV. D.2, between the Directorate General Environment (ENV) and the Joint Research Centre (JRC), on ‘Strengthening the policy support capacity of the European Alien Species Information Network (EASIN) and its notification system (NOTSYS) (EASIN6)’. Many thanks to data partners in Table 2.

## Author contributions

C.M. wrote the first draft of the manuscript, carried out the technical validation of the species occurrences at country level, and reviewed the manuscript. A.B., T.C. and M.A. carried out the taxa data collection as plants and animals’ pathologist or epidemiology experts and reviewed the manuscript. I.D. and F.D. helped with the data management in the EASIN web applications. A.C.C., E.G. and K.T. coordinated the work and reviewed the manuscript. All authors have approved the final article.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-022-01485-1>.

**Correspondence** and requests for materials should be addressed to C.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022