Digitalization of the tardigrade collection of Giuseppe Ramazzotti

Silvia Zaupa,[1*] Diego Fontaneto,[2,3] Roberta Piscia,[2] Lyudmila Kamburska[2,3]

[1]National Research Council of Italy (CNR), Water Research Institute (IRSA), Taranto; [2]National Research Council of Italy (CNR), Water Research Institute (IRSA), Verbania Pallanza; [3]National Biodiversity Future Center (NBFC), Palermo, Italy

ABSTRACT

The Collection of Tardigrada by Giuseppe Ramazzotti, preserved at the Water Research Institute of the National Research Council (CNR-IRSA) in Verbania Pallanza, Italy, is composed of numerous slides and microscopic preparations of tardigrades, including animals and their eggs. It is a historical collection that includes samples taken over a period of years from 1938 to 1964 in Europe, Africa, North and South America by one of the Italian taxonomists working on the phylum Tardigrada. The dataset of the Ramazzotti Collection is composed of 3319 records, prepared within the Project ITINERIS (ITalian INtegrated Environmental Research Infrastructures System) dedicated to build the Italian Hub of Research Infrastructures in the environmental scientific domain. The purpose of creating this dataset is to protect the information through the digitalization of the samples, and to facilitate sharing the information on the animals collected by Ramazzotti and stored at CNR-IRSA, Verbania Pallanza. For this reason, each record in the database is associated with a photograph of the microscopic slide and with a photograph of the animal. The data is organized into a standardized Darwin Core Archive format, thus available globally, through the Global Biodiversity Information Facility infrastructure (GBIF). The Collection is also registered in the Global Registry of Scientific Collections (GRSciColl).

INTRODUCTION

The ITINERIS (ITalian INtegrated Environmental Research Infrastructures System) project has as its main objective to build the network of Italian Environmental Research Infrastructures, connecting the user community and establish access, through the ITINERIS HUB, to the wide range of knowledge, data, analytical tools and services produced by the participating research infrastructures (https://itineris.cnr.it).

In the framework of the project, the Ramazzotti Collection of

Tardigrada, preserved at the Water Research Institute, National Research Council of Italy (CNR-IRSA) in Verbania Pallanza, Italy, has been the subject of digitization. It is a historical collection that includes samples taken over a period ranging from 1938 to 1964 from different areas of the world. Most of the samples were taken in Europe, in particular from Italy, but also from Germany, France, Finland, Norway, Poland, Austria, Switzerland, Sweden, the Netherlands and the United Kingdom. Other samples are from Africa (Uganda and Congo), North America (United States and Mexico) and South America (Chile) (Fig. 1). According to the information on the slide of the samples from aquatic environments, they were collected by pumps, while for terrestrial ones the sampling methods are unknown.

Different mounting media and colorants were used for the preparation and preservation of the slides: mainly Faure and Polyvinyl, but also Formalin, Lactophenol, Carmine Acetic, Chlorazol Black, Eosin methylene, Potassium hydroxide, Turtox CMC-S. In total, the collection contains 3319 specimens either as individuals or eggs.

The Ramazzotti Collection is composed of ten boxes, numbered in progressive order with Roman numerals, containing a variable number of slides. There are a total of 859 slides divided as follows: box I contains 100 slides, II 98, III 98, IV 100, V 99, VI 94, VII 87, (VIII is missing), IX 83, X 83 and box XI contains 17 slides. The collection has a scientific value per se, being a collection of slides of a poorly studied group of animals, the phylum Tardigrada, in addition to a historical value. The collection was prepared by Giuseppe Ramazzotti (with the exception of 4 slides prepared and identified by Walter Maucci), an Italian amateur taxonomist who lived from 1898 to 1986. Giuseppe Ramazzotti was the only child of Ausano Ramazzotti, the founder of the liqueur factory Amaro Ramazzotti, a famous Italian alcoholic drink all over the world (Granata, 2015). In 1923 Giuseppe Ramazzotti married Angelina Buzzati, sister of Adriano Buzzati Traverso (a geneticist who worked at the same research institute in Verbania Pallanza currently hosting the collection: Capocci and Corbellini, 2002) and of Dino Buzzati (an Italian writer who is famous for his fiction stories on magical realism and social alienation

(Castelli, 2023). Giuseppe Ramazzotti remained famous for two of his passions: natural sciences and tobacco pipes. He was for a time the largest collector of pipes in Europe, publishing several books on pipes and on many other subjects, such as trains, spiders, the moon, and power stations. In 1962, he proposed the phylum Tardigrada, and in 1968, at the age of sixty, he became a professor at the University of Milan (Bertolani *et al.*, 2011).

Here, we make the data, metadata and all the images of the Collection available online for consultation and use, at https://www.gbif.org/dataset/e9aeed17-b33f-4af2-8471-9be87633d7e1 and at https://scientific-collections.gbif.org/collection/91ebb3a8-89cb-4403-8324-3ea2d6d445d1.

METHODS

The specimens were photographed using a microscope Zeiss Axioplan, with TiEsseLab TrueChrome HD IIS camera and Leica DM6 B, with Leica K3C camera. The photos have been named to give a unique code for each image, using: CollectionCode, the Roman numeral of the box, the slide number and the number of the specimen. ES: RamVB-III-31_4 = box III-slide 31_animal 4 (Fig. 2a). The unique code for each photo is the same as that assigned to the record in the database for the same specimen. The photos, taken in .tiff format, were then converted to JPEG format,
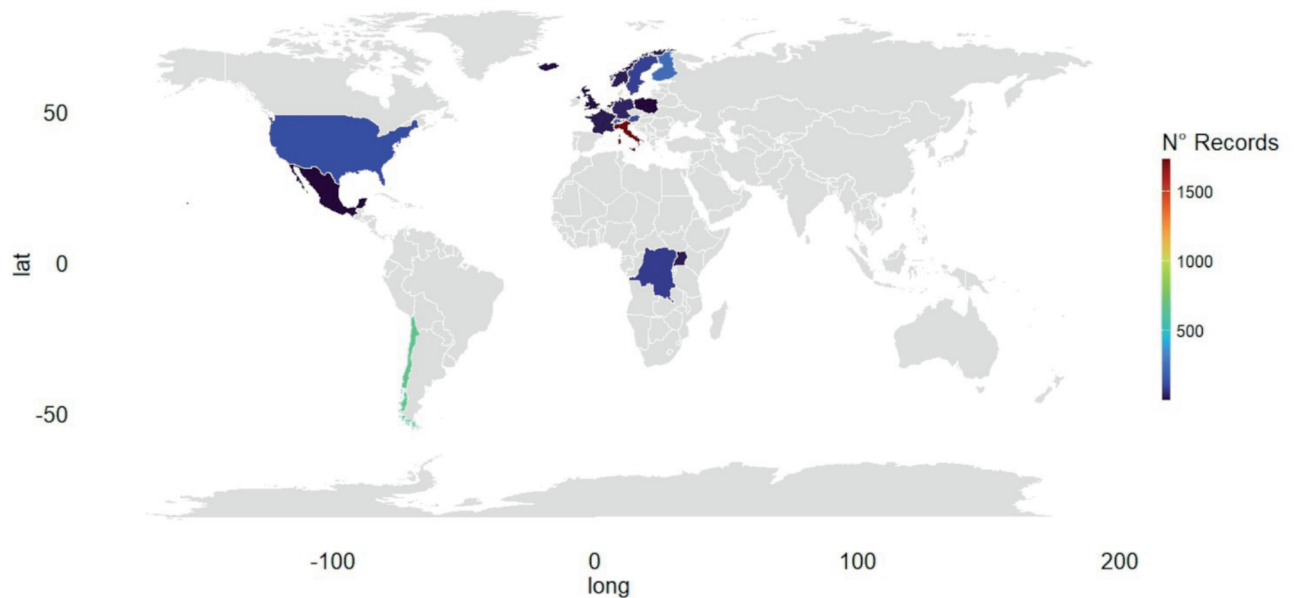

Fig. 1. Origin of the samples in the Ramazzotti Collection. The colour indicates the number of records per country.
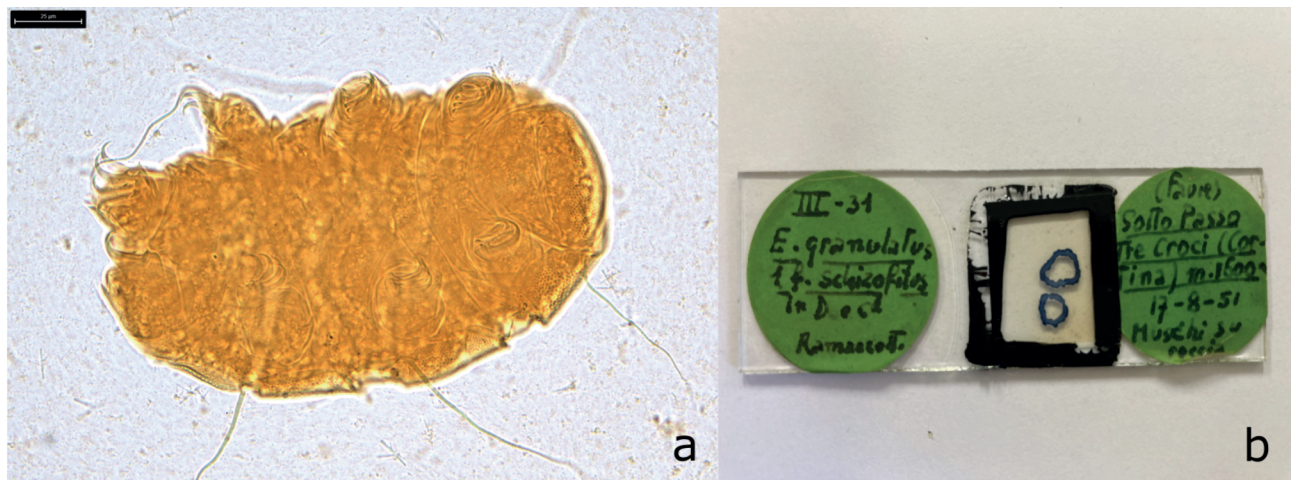

Fig. 2. Example of photos from the Ramazzotti Collection: (a) specimen number 4 of slide 31: *Echiniscus granulatus* (RamVB-III-31_4); (b) slide 31 Box III (RamVB-III-31_S).
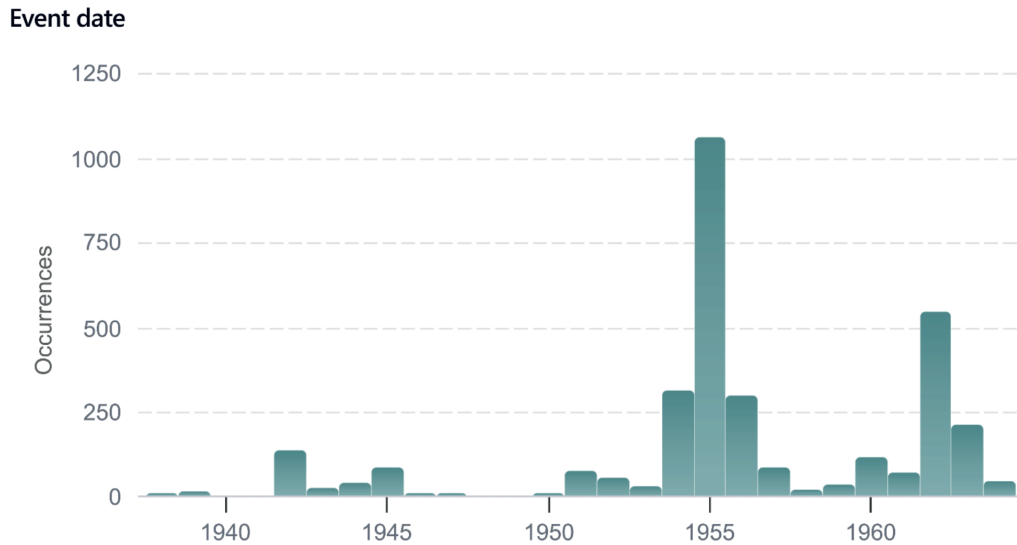
Event date

Fig. 3. Tardigrades occurrence records per year in the Ramazzotti Collection.

using the R Project software R (R Core Team, 2024), so as to make them available for viewing on the GBIF webpage.

In addition to the animals, individual slides were also photographed. Photos of slides have been renamed with the Roman numeral of the box, the number of the slide, and the letter S, to indicate that it is the image containing the slide (Fig. 2b).

The images were then uploaded to the D4Science digital infrastructure designed to offer diverse communities a comprehensive suite of services (https://www.d4science.org).

To view the photos online, GBIF requires the upload a multimedia file containing links to the different images. This file has been created and contains the IDs of the individual records, the type of multimedia file (still image in our case), the format, type of license and creator of the photos and the link to the image corresponding to the record.

RESULTS

For the specimens a total of 3319 photos were produced. The size of each image ranges from approximately 5 MB to 18 MB in Tiff format, and 4 MB after conversion into JPEG. The photos of the slide were taken in .jpeg format with a common camera, and the size of each is approximately 500-600 KB. In this case 859 images were taken. The codes used for these photos are not present in the dataset as records, because they do not contain taxonomic information, but they are still an important part of the Collection, to keep track of the tags on them.

Each specimen is therefore associated with two photos, one with only the animal and the other one with the whole slide on which it is preserved.

DATASET DESCRIPTION

The dataset was developed at CNR-IRSA in Verbania Pallanza, Italy, and organised in a systematic and coherent way according to Darwin Core Standards (Wieczorek *et al.*, 2012).

A total of 4178 records of occurrences for the period 1938-1964 (Fig. 3) are present in the dataset: 855 refer to slides, 3319 to the actual samples of Tardigrada. Of these, 1751 have a species name, subsumed from the information on the slide, whereas for the others a taxonomic revision is necessary, because the same slide contains more than one species, so a taxonomic expert is needed to attribute each sample to the correct species. However, these individuals have been assigned to a higher taxonomic level such as genus, order or family (Fig. 4). Nomenclature of the species was also updated to currently used synonyms, referring to the World Register of Marine Species, WoRMS (https://www.marinespecies.org).
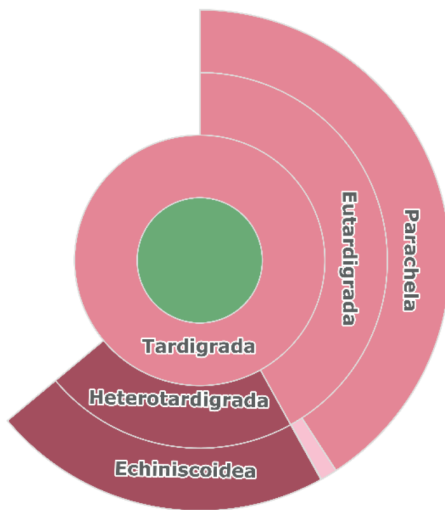

Fig. 4. Taxonomic distribution of occurrences of the Ramazzotti Collection.

All geographic locations with known coordinates, because indicated on the sample, were mapped in Google maps. Whereas for sampling points with unknown coordinates, the Locality centroid was used. A total of 204 locations are present in the dataset, but it was possible to trace the coordinates only for 88 of them.

We kept all available information for the Collection, thus the dataset includes 48 columns with each row containing a record of a tardigrade individual with taxonomy information.

The 48 columns are divided into 4 categories based on the type of information (Tab. 1):

Tab. 1. Description of the dataset available in GBIF with specific information relative to column names and description.

	Column label	Column description
Basic	OccurrenceID	An identifier for the Occurence
	InstitutionCode	The acronym in use by the institution having custody of the collection and dataset
	InstitutionID	The link to insitute's collection at https://scientific-collections.gbif.org/institution/9443e624-7071-4d54-b05d-a69c268bfe15/collections
	CatalogCode	Acronym referring to the type of collection present at the institute's headquarters. Composed by Institute acronym + short of Institute site + initial of collection
	CollectionCode	The acronym of the collection from which the record was derived
	CatalogNumber	A unique identifier for the record within the dataset or collection. The code is composed from the Collection Code + the Serial number
	basisOfRecord	The specific nature of the data record
	verbatimLabel	The content of this term includes the original label describing the specimens
	verbatimIdentification	A string representing the taxonomic identification as it appeared in the original record
	scientificNameAuthorship	The authorship information for the column verbatimIdentification
	dateOfFirstIdentification	The date on which the subject was determined
	originalDeterminavit	A name who assigned the Taxon to the subject
	datasetID	The ID of dataset registered in GrSciColl, composed by InstitutionCode + CollectionCode
	ObjectType	General term to describe the kind of specimen
	Preparation	Medium or support in which the specimen is preserved
	verbatimLabelRemarks	Comments or notes about the column occurrence
Taxonomy	Kingdom	The full scientific name of the kingdom in which the scientificNameis classified
	Phylum	The full scientific name of the phylum or division in which the scientificName is classified
	Class	The full scientific name of the class in which the scientificName is classified
	Order	The full scientific name of the order in which the scientificName is classified
	Family	The full scientific name of the family in which the scientificName is classified
	Fenus	The full scientific name of the genus in which the scientificName is classified
	specificEpithet	The name of the first or species epithet of the scientificName
	scientificName	The full scientific name
	infraspecificEpithet	The name with the lowest or terminal infraspecific epithet of the scientificName
	taxonRank	The taxonomic rank information (e.g. genus, species)
	scientificNameAuthorship	The authorship information for the scientificName
	MethodOIidentification	Method of identification applied
Morphological	Sex	The sex of the biological individual(s) represented in the Occurrence records
	lifeStage	The age class or life stage of the organism(s) at the time the occurrence was recorded
Sampling	Legit	A name of who collected the sample
	eventDate	The date-time or interval during which an event occurred
	Continent	The name of the continent in which the location occurs
	Country	The name of the country or major administrative unit in which the location occurs
	CountryCode	country codes deliver from: http://rs.gbif.org/vocabulary/iso/3166-1_alpha2.xml
	StateProvince	The name of the next smaller administrative region than the country (state, province, canton, department, region, etc.) in which the location occurs
	Region	The name of the administrative region in which the location occurs
	Locality	The specific description of the place
	verbatimSiteName	The sampling station code
	decimalLatitude	The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a location
	decimalLongitude	The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a location
	geodeticDatum	Spatial reference system: WGS84
	verbatimElevation	Vertical distance above sea level
	verbatimDepth	Depth below the local surface
	GeolocalizedGeographicObject	The identification of the geographical position (sampling point or centroid)
	Habitat	Habitat where the taxon was found
	eventRemarks	Comments or notes about the event
	SamplingTool	Type of instrument used for sampling

i) Basic information regarding the institution to which the collection belongs, with names, acronyms and type of collection. They also refer to the label, including the name of the species and the person who originally performed the determination, as well as the date of first identification, and the type of preparation. The information is structured in 18 columns: occurrenceID InstitutionCode, InstitutionID, CatalogCode, CollectionCode, CatalogNumber, basisOfRecord, verbatimLabel, verbatimIdentification, scientificNameAuthorship, DateOfFirstIdentification, OriginalDeterminavit, DatasetID, ObjectType, Preparation, and verbatimLabelRemarks;

ii) Taxonomy information of the species after taxonomic revision following WoRMS as reference site, with categories included in 12 columns: kingdom, phylum, class, order, family, genus, specificEpithet, scientificName, infraspecificEpithet, taxonRank, scientificNameAuthorship and Method OfIdentification;

iii) Morphology category about the digitized specimen, with two columns: gender and lifeStage;

iv) Sampling information, such as location, date, who collected the sample and sampling method, with 18 columns: legit, eventDate, Continent, Country, CountryCode, StateProvince, Region, Locality, verbatimSiteName, decimalLatitude, decimalLongitude, geodeticDatum, verbatimElevation, verbatimDepth, GeolocalizedGeographicObject, Habitat, event Remarks and SamplingTool.

Dataset object name: Tardigrade collection of Ramazzotti

Data licence: Creative Commons Attribution Non Commercial Licence (CC-BY-NC 4.0)

Data set citation: Zaupa S, Fontaneto D, Piscia R, Kamburska L (2025). Tardigrade Collection of Ramazzotti. Version 1.26. Consiglio Nazionale delle Ricerche - Istituto di Ricerca sulle Acque. *Occurrence dataset:* https://doi.org/10.15468/wngz4x accessed via GBIF.org.

DOI of the dataset: https://doi.org/10.15468/wngz4x

Distribution (permanent link) GBIF: https://www.gbif.org/dataset/ e9aeed17-b33f-4af2-8471-9be87633d7e1

Data format version: 1.26

Character encoding: UTF-8

Data format: csv

Date of dataset creation: 3 June 2024

Date of publication: 27 January 2025

Date of last revision: 15 April 2025

Metadata language: English

Metadata managers: Silvia Zaupa (silvia.zaupa@cnr.it), Diego Fontaneto (diego.fontaneto@cnr.it)

Licence of use: both access and use are free to any non-commercial user (CC-BY-NC 4.0). The authors would appreciate users to provide a link to the original dataset in GBIF (https://www.gbif.org/dataset/e9aeed17-b33f-4af2-8471-9be87633d7e1) and to cite the present paper when using the data. Stakeholders interested in additional information can contact the authors via information provided in the metadata.

Update policy: GBIF policy rules. Future changes to the dataset due to quality control activities might change its content, however the updates of the dataset are not planned.

Geographic coverage

Continent: Europe, Africa, North America, South America.

Study area by continents and countries: Europe (Austria, Finland, France, Germany, Italy, Netherlands, Norway, Sweden, Switzerland and the United Kingdom) Africa (Congo and Uganda), North America (Mexico and United States) and South America (Chile).

When possible, the coordinates were found by reporting the locations written on the label in Google maps. Since the specific coordinates of the sampling points are not known, the Locality centroid was used. In particular, for well-defined locations such as small lakes, refuges and mountain peaks this criterion was applied. Instead, if the location is represented by a large and not well-defined area, such as a city, the centroid was not used, and the record does not have coordinates. Data are georeferenced according to WGS 84 datum.

Sampling design: as reported on the slide, the samples were collected by pumps in an aquatic environment, while for terrestrial ones the sampling methods are unknown.

Habitat type: mosses and lichens in different substrates, soil, water.

Quality control for geographic data: reliability of coordinates where available was checked in Google maps. Geographic coordinate format, coordinates within country/provincial boundaries, and the absence of anomalous in the dataset were also double checked.

Taxonomic coverage

The dataset is composed of taxa belonging to the tardigrade phylum.

Taxonomic ranks: species, genus, family, order and class rank were included in the dataset.

Taxonomic methods: all collected organisms were preserved by preparation of slides by Giuseppe Ramazzotti. The taxonomy and nomenclature of the dataset follows the original one, with no updates.

Quality control for taxonomic data: the revision of the species nomenclature was done using WoRMS as a reference site.

Taxon specialists: Giuseppe Ramazzotti, Walter Maucci.

Management details

Project title: digitalization of the tardigrade collection of Ramazzotti.

Temporal coverage: 1938-1964

Database managers: Silvia Zaupa, Diego Fontaneto, Lyudmila Kamburska

Basis of record: preserved specimens

Curator of the collection: Roberta Piscia

Collection home page: https://scientific-collections.gbif.org/collection/91ebb3a8-89cb-4403-8324-3ea2d6d445d1

ACKNOWLEDGEMENTS

REFERENCES

Bertolani R, Nelson DR, 2011. Changes in scientific approaches as seen through 35 years of tardigrade symposia. J Zool Syst Evol Re. 49:2-5.

Castelli A, 2023. A limbo between Beckett and Kafka: The Tartar Steppe. Critique-Stud Contemp Fict 64:728-736.

Capocci M, Corbellini G, 2002. Adriano Buzzati-Traverso and the foundation of the International Laboratory of Genetics and Biophysics in Naples (1962-1969). Stud Hist Philos Sci C 33:489-513.

Granata M, 2015. The capital of innovation, pp. 185-226. In: Smart Milan. Cham, Springer.

R Core Team 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. Available from: https://www.R-project.org

Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, et al, 2012. Darwin Core: An evolving community-developed biodiversity data standard. PLoS One 7:e29715.