

Table S1. Darwin Core fields used, descriptions in the Darwin Core were taken from the TDWG website (<https://dwc.tdwg.org/terms/>; accessed on 29 July 2022).

Field	Definition
occurrenceID	An identifier for the Occurrence (as opposed to a particular digital record of the occurrence). In the absence of a persistent global unique identifier, construct one from a combination of identifiers in the record that will most closely make the occurrenceID globally unique.
kingdom	The full scientific name of the kingdom in which the taxon is classified.
phylum	The full scientific name of the phylum or division in which the taxon is classified.
class	The full scientific name of the class in which the taxon is classified.
order	The full scientific name of the order in which the taxon is classified.
family	The full scientific name of the family in which the taxon is classified.
genus	The full scientific name of the genus in which the taxon is classified.
specificEpithet	The name of the first or species epithet of the scientificName.
infraspecificEpithet	The name of the lowest or terminal infraspecific epithet of the scientificName, excluding any rank designation.
scientificNameAuthorship	The authorship information for the scientificName formatted according to the conventions of the applicable nomenclaturalCode.
taxonRank	The taxonomic rank of the most specific name in the scientificName.
scientificName	The full scientific name, with authorship and date information if known. When forming part of an Identification, this should be the name in lowest level taxonomic rank that can be determined. This term should not contain identification qualifications, which should instead be supplied in the IdentificationQualifier term.

country	The name of the country or major administrative unit in which the Location occurs.
countryCode	The standard code for the country in which the Location occurs.
locality	The specific description of the place.
decimalLatitude	The geographic latitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are north of the Equator, negative values are south of it. Legal values lie between -90 and 90, inclusive.
decimalLongitude	The geographic longitude (in decimal degrees, using the spatial reference system given in geodeticDatum) of the geographic center of a Location. Positive values are east of the Greenwich Meridian, negative values are west of it. Legal values lie between -180 and 180, inclusive.
geodeticDatum	The ellipsoid, geodetic datum, or spatial reference system (SRS) upon which the geographic coordinates given in decimalLatitude and decimalLongitude as based.
eventDate	The date-time or interval during which an Event occurred. For occurrences, this is the date-time when the event was recorded. Not suitable for a time in a geological context.
day	The integer day of the month on which the Event occurred.
month	The integer month in which the Event occurred.
year	The four-digit year in which the Event occurred, according to the Common Era Calendar.
samplingProtocol	The names of, references to, or descriptions of the methods or protocols used during an Event.
taxonID	
basisOfRecord	Recommended best practice is to use the standard label of one of the Darwin Core classes.
institutionCode	The name (or acronym) in use by the institution having custody of the object(s) or information referred to in the record.
collectionCode	The name, acronym, coden, or initialism identifying the collection or data set from which the record was derived.

catalogNumber	An identifier (preferably unique) for the record within the data set or collection.
recordNumber	An identifier given to the Occurrence at the time it was recorded. Often serves as a link between field notes and an Occurrence record, such as a specimen collector's number.
recorderBy	A list (concatenated and separated) of names of people, groups, or organizations responsible for recording the original Occurrence. The primary collector or observer, especially one who applies a personal identifier (recordNumber), should be listed first.
individualCount	The number of individuals present at the time of the Occurrence.
preparations	A list (concatenated and separated) of preparations and preservation methods for a specimen.
identifiedBy	A list (concatenated and separated) of names of people, groups, or organizations who assigned the Taxon to the subject.
license	A legal document giving official permission to do something with the resource.
rightsHolder	A person or organization owning or managing rights over the resource.
typeStatus	A list (concatenated and separated) of nomenclatural types (type status, typified scientific name, publication) applied to the subject.