Data integration and provenance in the Encyclopedia of Life

Katja Schulz
Encyclopedia of Life
Smithsonian Institution
National Museum of Natural History
EOL aggregates data from content partners
Taxonomy

**Taxon** recognized by [AntWeb]:

**Arthropoda** +

**Insecta** +

**Hymenoptera** +

**Formicidae** +

**Adetomyrma** +

**Adetomyrma venatrix**

**Adetomyrma aureocuprea**

**Adetomyrma bressleri**

**Adetomyrma caputleae**

**Adetomyrma cassis**
General Description

All species of *Pelegrina* are easily distinguished from relatives by the male embolus, which have two rami placed retrolaterally to the opening. The embolus is further differentiated in *P. montana* by small denticles on its anterior surface. On average, *P. montana* is larger and of a darker colouration than related species. On males, the side band and forehead band common to all *Pelegrina* are a faded white and the cheek band is either weakly present or completely absent. The clypeus is brown, with dark setae overhanging the chelicerae. The white marks on the dorsal portion of the abdomen are absent except on the latter half, where they are reduced, forming thin, pale bars. Females possess the characteristic thickened epigynal flaps present in all *Pelegrina*. A prominent diagnostic characteristic is these flaps’ short length and convex shape. The first curve of the duct is narrow, and the second curve follows obliquely. Immediately behind the flaps is a prominent ridge shaped like a rounded rectangle. The carapace of the spider is dark, but covered in sparse, white hairs. These white scales are found in higher density on the clypeus. The abdomen is dark brown with small pairs of white spots (Maddison, 1996). Males: avg. body length 4.6 mm, carapace length 2.1 mm; females: avg. body length 5.9-6.9 mm, carapace length 2.5 mm (Maddison, 1996).

Source information

[License](https://creativecommons.org/licenses/by-nc/4.0) 2015 University of Alberta Museums
© University of Alberta Museums
View source
Supplier: University of Alberta Museums
Source: E.H. Strickland Entomological Museum
Author: Trevoy, S.
**Colobus satanas**

**Black Colobus**

**Taxonomy**

- **Body length**: 1,378 mm (female)

**Catalog Number**: 598560

- **Measurement Unit**: mm
- **Sex**: female

**Source**: [http://collections.mnh.si.edu/search/mammals/?nb=598560](http://collections.mnh.si.edu/search/mammals/?nb=598560)

**Verbatim Measurements**:
- Total Length: 1378mm
- Tail Length: 860mm
- Hind Tarsus Length: 174mm
- Ear Notch Length: 35mm
- Specimen Weight: 8600.0g

**Institution Code**: USNM

**Stomach Content**: digested leaves = 483.9g

**Smithsonian Institution, National Museum of Natural History, Department of Vertebrate Zoology, Division of Mammals.**

**Data for specimen 598560**

[http://collections.mnh.si.edu/search/mammals/](http://collections.mnh.si.edu/search/mammals/)

**Preparations**: Skin; Skull; Skeleton

**Occurrence Details**:
- Progeny Details: 1: embryo: 42: mm
- Mammals Field Number: 4055; Mammals Field Number: TU 4256

**Locality**
- Precise Locality: Evuenam; Province/State: Rio Muni; Country: Equatorial Guinea; Jones, Clyde

**Scientific Name**: Colobus satanas
At black light among mixed forest, farmed fields, tidal creek with salt marsh

Limestone cave in Sierra Cachiri NW of Maracacho, Distrito Mara, Estado Zulia, Venezuela.

Dacus psidii
in guava
Noumea New Caledonia VI-50
N. Krauss - 1X 50-11833

Bred
PARASITE
PHYLLATOMA NEMORATA

Dendroica petechia Chrysepoda NYK

Ovary: 4.5 x 4 mm, ova nimate, oviduct 20 mm, no brood patch present.
Skull: 100 x 200; ir is dark brown, maxilla dark gray, mandible medium brown, trace fat; pearl gray, tonsils; toes yellow/brown, pale yellow pads. Stomach saved; insect parts; no reangeRestricts molt; light/malar; molt, trace head molt.
Keel: 14.3 mm. interorbital distance 2.8 mm. Left wing chord: 64.2 mm; arc: 60 mm; tail: 52 mm; probable mate of NKK 366 shot in beach area 200 m from mangrove, foraging in isolated Avicennia.
Data integration: names reconciliation

- Name strings
- Synonym relationships
- Authority strings
- Ranks
- Parent taxa
- Child taxa
## Data integration: community standards

<table>
<thead>
<tr>
<th>Darwin Core</th>
<th>Dublin Core</th>
<th>Audubon Core</th>
<th>Species Profile Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>taxa</td>
<td>rights &amp; attribution</td>
<td>images</td>
<td>subjects</td>
</tr>
<tr>
<td>occurrences</td>
<td>language, format</td>
<td>videos</td>
<td></td>
</tr>
<tr>
<td>measurements or facts</td>
<td>general resource description</td>
<td>sounds</td>
<td></td>
</tr>
</tbody>
</table>
Data integration: semantic annotations

The OBO Foundry

Open Biological and Biomedical Ontology

envo environments
oba biological attributes
pato phenotypic quality
to plant traits
uberon multi-species anatomy
vt vertebrate traits
ro relations
uo units of measurements
stato statistics
chebi chemical entities

Marineregions.org

Semanticscience
Integrated Ontology (SIO)

GRBio
GLOBAL REGISTRY OF BIODIVERSITY REPOSITORIES
### Data integration: semantic annotations

<table>
<thead>
<tr>
<th><strong>body length</strong></th>
<th><strong>cm</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OBA:VT0001256</strong>&lt;br&gt;The length of a multicellular organism.</td>
<td><strong>6</strong>&lt;br&gt;<strong>UO:0000015</strong>&lt;br&gt;A length unit which is equal to one hundredth of a meter or $10^{-2}$ m.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>male</strong></th>
<th><strong>mature</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PATO:0000384</strong>&lt;br&gt;A biological sex quality inhering in an individual or a population whose sex organs contain only male gametes.</td>
<td><strong>PATO:0001701</strong>&lt;br&gt;A quality inhering in a bearer by virtue of the bearer's exhibiting complete growth, differentiation, or development.</td>
</tr>
</tbody>
</table>
Dissemination to diverse audiences

humans

machines

{"identifier":128534,"scientificName":"Orcinus orca (Linnaeus, 1758)","richness_score":900.0,"taxonConcepts": [{"identifier":15256726,"scientificName":"Orcinus orca (Linnaeus, 1758)","nameAccordingTo":"Species 2000 & ITIS Catalogue of Life: April 2013","canonicalForm":"Orcinus orca","sourceIdentifier":6850349,"taxonRank":"Species"}, {"identifier":89338079,"scientificName":"Orcinus orca (Linnaeus, 1758)","nameAccordingTo":"Paleobiology Database","canonicalForm":"Orcinus orca","sourceIdentifier":urn:paleodb:btn64541","taxonRank":"Species"}, {"identifier":56582740,"scientificName":"Orcinus orca (Linnaeus, 1758)","nameAccordingTo":"WORMS Species Information (Marine Species)","canonicalForm":"Orcinus orca","sourceIdentifier":137102,"taxonRank":"Species"}, {"identifier":61335391,"scientificName":"Orcinus orca (Linnaeus, 1758)","nameAccordingTo":"WORMS Species Information (Marine Species)","canonicalForm":"Orcinus orca","sourceIdentifier":137102,"taxonRank":"Species"}, {"identifier":24950241,"scientificName":"Orcinus orca","nameAccordingTo":"IUCN Red List (Species Assessed for Global Conservation)","canonicalForm":"Orcinus orca","sourceIdentifier":IUCN-15421"}, {"identifier":65596031,"scientificName":"Orcinus orca (Linnaeus, 1758)","nameAccordingTo":"Integrated Taxonomic Information System (ITIS)","canonicalForm":"Orcinus orca","sourceIdentifier":180469,"taxonRank":"Species"}, {...}]}

<table>
<thead>
<tr>
<th>Darwin Core</th>
<th>Dublin Core</th>
</tr>
</thead>
<tbody>
<tr>
<td>institutionCode</td>
<td>bibliographicCitation</td>
</tr>
<tr>
<td>collectionCode</td>
<td>publisher</td>
</tr>
<tr>
<td>catalogNumber</td>
<td>contributor</td>
</tr>
<tr>
<td>recordedBy</td>
<td>creator</td>
</tr>
<tr>
<td>identifiedBy</td>
<td>source</td>
</tr>
<tr>
<td>measurementDeterminedBy</td>
<td>agent – role:</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>author</td>
</tr>
<tr>
<td></td>
<td>funder</td>
</tr>
<tr>
<td></td>
<td>compiler</td>
</tr>
<tr>
<td></td>
<td>illustrator</td>
</tr>
<tr>
<td></td>
<td>copyright holder</td>
</tr>
<tr>
<td></td>
<td>metadata contact</td>
</tr>
<tr>
<td></td>
<td>curator</td>
</tr>
<tr>
<td></td>
<td>photographer</td>
</tr>
<tr>
<td></td>
<td>data contributor</td>
</tr>
<tr>
<td></td>
<td>provider</td>
</tr>
<tr>
<td></td>
<td>data manager</td>
</tr>
<tr>
<td></td>
<td>recordist</td>
</tr>
<tr>
<td></td>
<td>editor</td>
</tr>
<tr>
<td></td>
<td>reviewer</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Audubon Core</td>
<td></td>
</tr>
<tr>
<td>furtherInformationURL</td>
<td></td>
</tr>
<tr>
<td>derivedFrom</td>
<td></td>
</tr>
<tr>
<td>rights:owner</td>
<td></td>
</tr>
</tbody>
</table>
Attribution can be complex

creator: Michele Esposito
recorded by: Epifânio, S.; Chaul, J.
identified by: Chaul, J.
rights holder: California Academy of Sciences
dataset: AntWeb
publisher: California Academy of Sciences
supplier: Global Biodiversity Information Facility (GBIF)
collection code: ANTWEB
institution code: CAS
owner institution code: UFV-LABECOL
catalogue number: ufv-labecol-007525
occurrence id: CAS:ANTWEB:ufv-labecol-007525
identifier: CAS:ANTWEB:ufv-labecol-007525
**Colobus satanas**  
*Black Colobus*

### Data about this record

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>catalog number</strong></td>
<td>598560</td>
</tr>
<tr>
<td><strong>measurement unit</strong></td>
<td>mm</td>
</tr>
<tr>
<td><strong>sex</strong></td>
<td>female</td>
</tr>
<tr>
<td><strong>source</strong></td>
<td><a href="http://collections.mnh.si.edu/search/mammals/?nb=598560">http://collections.mnh.si.edu/search/mammals/?nb=598560</a></td>
</tr>
<tr>
<td><strong>measurement remarks</strong></td>
<td>Verbatim measurements: Total Length: 1378mm; Tail Length: 860mm; Hind Tarsus Length: 174mm; Ear Notch Length: 35mm; Specimen Weight: 8600.0g</td>
</tr>
<tr>
<td><strong>institution code</strong></td>
<td>USNM</td>
</tr>
<tr>
<td><strong>field notes</strong></td>
<td>Stomach Content: digested leaves = 483.9g</td>
</tr>
<tr>
<td><strong>citation</strong></td>
<td>Smithsonian Institution, National Museum of Natural History, Department of Vertebrate Zoology, Division of Mammals. Data for specimen 598560. <a href="http://collections.mnh.si.edu/search/mammals/">http://collections.mnh.si.edu/search/mammals/</a></td>
</tr>
<tr>
<td><strong>preparations</strong></td>
<td>Skin; Skull; Skeleton</td>
</tr>
<tr>
<td><strong>occurrence remarks</strong></td>
<td>Progeny Details: 1: embryo: 42: mm; Mammals Field Number: 4055; Mammals Field Number: TU 4256</td>
</tr>
<tr>
<td><strong>locality</strong></td>
<td>Precise Locality: Evuenam; Province/State: Rio Muni; Country: Equatorial Guinea;</td>
</tr>
<tr>
<td><strong>recorded by</strong></td>
<td>Jones, Clyde</td>
</tr>
<tr>
<td><strong>scientific name</strong></td>
<td>Colobus satanas</td>
</tr>
</tbody>
</table>

**body length** 1,378 mm *(female)*
The time has come for Natural History Collections to claim co-authorship of research articles

Germinal Rouhan,¹ Laurence J. Dorr,² Laurent Gautier,³ Philippe Clerc,³ Serge Muller¹ & Myriam Gaudeul¹

¹ Muséum national d’Histoire naturelle, Institut de Systématique Évolution Biodiversité (ISYEB, UMR 7205 CNRS, MNHN, UPMC, EPHE), Sorbonne Universités, Herbier national, 57 rue Cuvier, CP39 – 75005, Paris, France
² National Museum of Natural History, Department of Botany, MRC-166, Smithsonian Institution, P.O. Box 37012, Washington, D.C. 20013-7012, U.S.A.
³ Conservatoire et Jardin botaniques de la Ville de Genève, C.P. 71, 1292 Chambéry, Switzerland

Author for correspondence: Germinal Rouhan, rouhan@mnhn.fr

DOI https://doi.org/10.12705/665.2
Developing a guideline to standardize the citation of bioresources in journal articles (CoBRA)

Elena Bravo, Alessia Calzolari, Paola De Castro, Laurence Mobile, Federica Napolitani, Anna Maria Rossi and Anne Cambon-Thomsen

BMC Medicine 2015 13:33
Received: 25 June 2014 Accepted: 2 January 2015 Published: 17 February 2015

Quantifying the use of bioresources for promoting their sharing in scientific research

Laurence Mobile, Raymond Dalgleish, Gudmundur A Thorisson, Mylène Deschênes, Robert Hewitt, Jane Carpenter, Elena Bravo, Mirella Filocamo, Pierre Antoine Gourraud, Jennifer R Harris ...

Published: 01 May 2013 Article history
Transitive Credit for contributors

Issues in Research Software

Transitive Credit as a Means to Address Social and Technological Concerns Stemming from Citation and Attribution of Digital Products

Author: Daniel S. Katz

Keywords: citation, credit, attribution, software

Transitive Credit for contributors

[Diagram showing relationships between authors, papers, software, and data, with numeric values indicating contributions.]
Transitive Credit for contributors

- unique identifiers for products
- unique identifiers for people
- unique identifiers for institutions
- standards to define contributor roles
- automated provenance systems
  - harvest contribution data
  - build contribution metrics
What can aggregators do?

- support attribution standards
- expose identifiers
- provide attribution guidance to data users
- provide metrics for individuals & institutions