NHM Interactions Bank

Steen Dupont[1], Dataset Contributors*, Ed Baker[1,2]

[1] Natural History Museum, London
[2] Department of Electronic Engineering, University of York

s.dupont@nhm.ac.uk

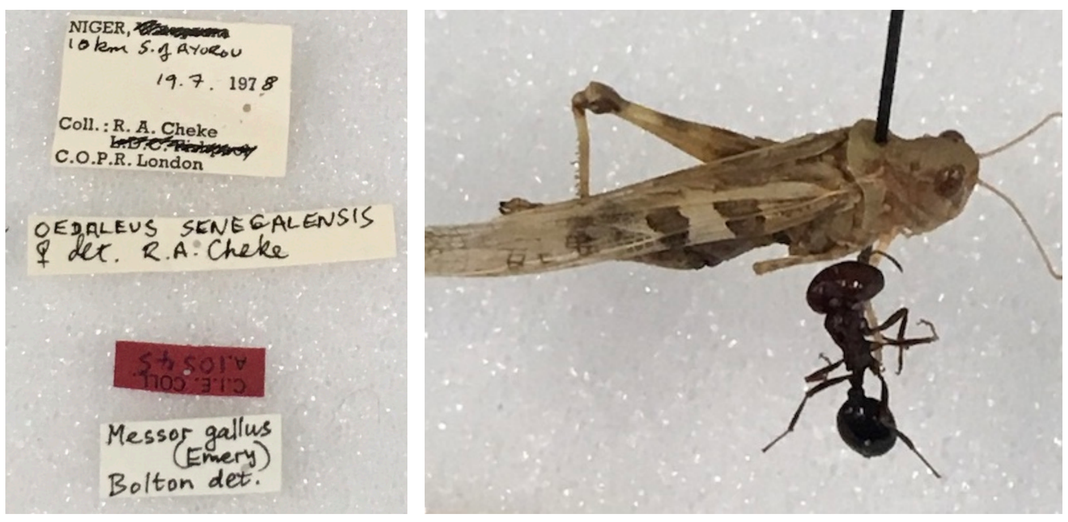
* Gaden S. Robinson, Phillip R. Ackery, Ian J. Kitching, George W. Beccaloni, Luis M. Hernandez, Vincent S. Smith, Amoret Whitaker, John S. Noyes, Chris Lyal

INTRODUCTION

In order to understand life on Earth it is essential to understand not only the distribution and traits of species, but how they interact with each other. Biodiversity informatics as a discipline has created stable, global infrastructures for both species distribution (GBIF) and trait (TraitBank) data. The infrastructure for ecological interaction datasets (Global Biotic Interactions: GloBI [1]) is not yet as mature or as well used.

Ecological interaction datasets are, like natural history specimens, fragmented and widely distributed. They can be found scattered through scientific literature and specimen labels in museums.

Numerous publications have synthesised interaction datasets for taxonomic groups (e.g. cockroaches [2], Lepidoptera [3]). Widespread adoption of computerised databases, and later the internet as a dissemination platform, bought these taxon-specific datasets into the electronic (e.g. cestodes [4]).



An example of a species interaction that has been collected and preserved. Both species have been identified by subject experts. Liberating such records form the collection and making them publicly available is one major goal of the NHM Interaction Bank.

DESIGN CONSIDERATIONS

The data model of ecological interactions for the NHM Interaction Bank must be able to handle diverse interaction datasets and show the relationships between these disparate datasets to make connections and trends discoverable.

Captive rearing records The Phasmida Interactions Database contains data on the hostplants of phasmids from both observations in the wild and captive rearing using alternative hosts. The system implemented allows for these interactions to be clearly identified, and included or excluded from analyses as appropriate.

Life-cycle stages Various groups (e.g. Hymenoptera) are parasites of a specific life-cycle stage of their host, while others (e.g. helminths) may have multiple hosts during depending on their life-cycle stage. Our implementation allows life-cycle stage of both organisms to be recorded.

Organism parts Many interactions are restricted to certain parts of organisms (e.g. lice on restricted areas of birds) and many herbivorous species eat only the leaves of plants. The data structure used allows these distinctions to be recorded.

DATA STANDARDS

The Relations Ontology [5] defines a large number of interactions between organisms. This allows for a hierarchy of standardised associations to be established (e.g. ‘parasite of’ is a parent of ‘endoparasite of’).

GLOBAL SCALE

The NHM Interactions Bank is regularly harvested by the aggregator GloBI that collects published interaction datasets from multiple web sources. These data are also accessible via the Enyclopedia of Life.

FUTURE PLANS

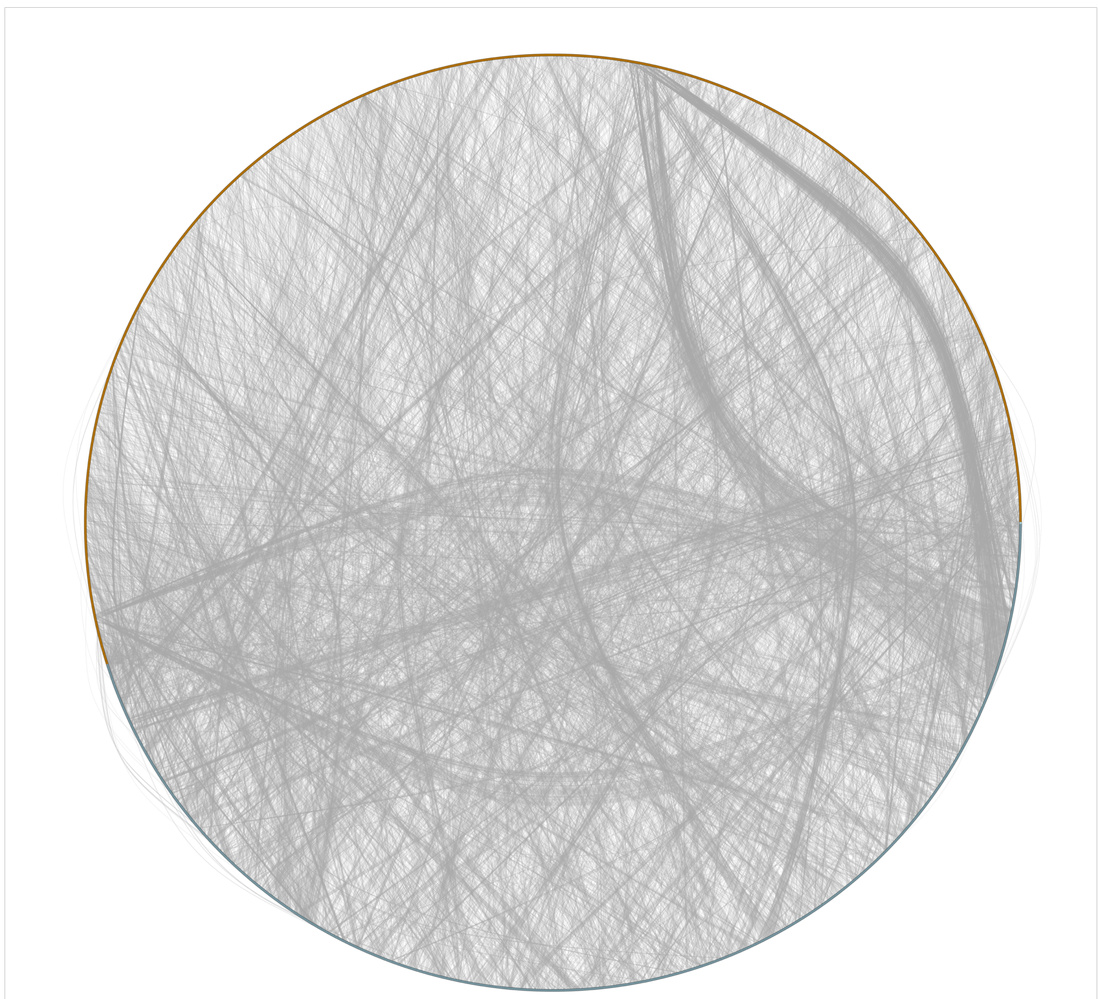
Future developments will incorporate datasets beyond insects, including the Host-Parasite database of parasitic worms and the Cestode Life Cycle Database. Digitisation of the Index to the Pittioni Bee Collection will add historic data of flower visits by bumble bees. Ongoing improvements to the source data projects (including cleaning the bibliographic data in HOSTS) will automatically be aggregated by NHM Interactions Bank and harvested by GloBI Unique record identifiers allow end users to easily incorporate changes in the source data into their ongoing projects.

CONTRIBUTIONS

This project was conceived by Ed Baker as part of the NHM’s Digital Collections Programme. It is currently overseen by Steen Dupont.

REFERENCES

1. Poelen JH, Simons JD & Mungall CJ (2014) Global biotic interactions: an open infrastructure to share and analyse species-interaction datasets. *Ecological Informatics* 24:148-159
2. Roth L & Willis E (1960) The Biotic Associations of Cockroaches. *Smithsonian Miscellaneous Collections* 141.
3. Beccaloni GW, Viloria AL, Hall SK & Robinson GS (2008) catalogue of the hostplants of the Neotropical butterflies. Sociedad Entomologica Aragonesa.
4. Lefebvre F, Georgiev B, Bray R & Littlewood DTJ (2009) Developing a dedicated cestode life cycle database: lessons from the hymenolepidids. *Helminthologia* 46(1):21-27
5. https://www.ebi.ac.uk/ols/ontologies/ro



Visualisation of Interactions related to Phthiraptera (blue) and their mammal and bird hosts (red). The broad patterns of coevolution and host-specificity can be seen (interactions: n=12,536)

http://data.nhm.ac.uk/dataset/nhm-ib