# Use Case: Biotic Interactions - Sottunga Island *Melitaea cinxia* Population Study

## Status: Reviewed (2023-11-26)

## Summary

This case study is based on the work described in Duplouy et al. 2021, which explores the long-term population dynamics of the butterfly *Melitaea cinxia* following its introduction to the Åland island Sottunga along with, unknown at the time, the parasitoid *Hyposoter horticola* and the hyperparasitoid *Mesochorus cf. stigmaticus*. In addition to this hierarchy of dependent relationships, the study also investigated the effect on the success of the hyperparasitoids based on the presence or absence of the endosymbiont bacteria *Wolbachia pipientis* in the parasitoid host. A total of 323 caterpillars from five distinct locations between 1992 and 2013 were collected, reared, and assessed for the presence of the parasitoids, hyperparasitoids and Wolbachia. The complex set of relationships between the host butterfly, parasitic wasp, hyperparasitic wasp and bacterial infection was shown to affect the phenotype of the host butterfly as well as population dynamics that ultimately maintain genetic diversity on the island despite multiple observed population bottlenecks.

### Highlights

- Events have multiple entities of interest between which multiple relationships can be asserted (relationships are not in and of themselves the entities of interest)
- Chains of proximate dependency of relationships can be captured
- Specimens and DNA extracts are left out of the model (they are captured in other use cases) to focus on the relationships.

### Concepts - see [Glossary](Glossary)

## Conceptual Model

There is a wide variety of biotic interactions that can be thought of as happening at distinct scales. Interactions tend to be described as relationships between dwc:Organisms and dwc:Taxa in any combination of those two types, plus their complementary inverse relationships. For example, if O is a dwc:Organism (e.g., a biological individual) and T is a dwc:Taxon (at whatever taxonomic rank), the following combinations can be found: O1 relatedTo O2, O2

inverseRelatedTo O1; O1 related to T1, T1 inverseRelatedTo O1; T1 relatedTo T2, T2 inverseRelatedTo T1. All of these can be modeled as dwc:Organism relationships, where relationships strictly involving identifiable dwc:Organisms can be thought of as involving "this" dwc:Organism while relationships involving dwc:Taxa can be thought of as involving "some" dwc:Organism. All relationships would use dwc:Organism identifiers and the relationshipType could distinguish between scales, while Taxon-scale subjects and objects would have no non-empty dwc:Organism properties.

A second feature of the model is to provide the capacity to posit that there is a sequential dependency among relationships using a link from one relationship to another one on which it directly depends. This is beyond the simple capacity to express the order in which relationships occurred, provided naturally from the Relationships being based on dwc:Events. It provides the ability to track dependencies that are more complex than just co-occurring or sequential. The model supports complex multiple co-occurrent relationships to be modeled independently as pairwise relationships.

Finally, the model treats a Relationship as an assertion applied to a subject and object, each of which are an EntityOfInterest in a dwc:Event. A distinct object for the relationship is deemed preferable to treating the Relationship as an EntityOfInterest itself or as a dwc:Event directly.
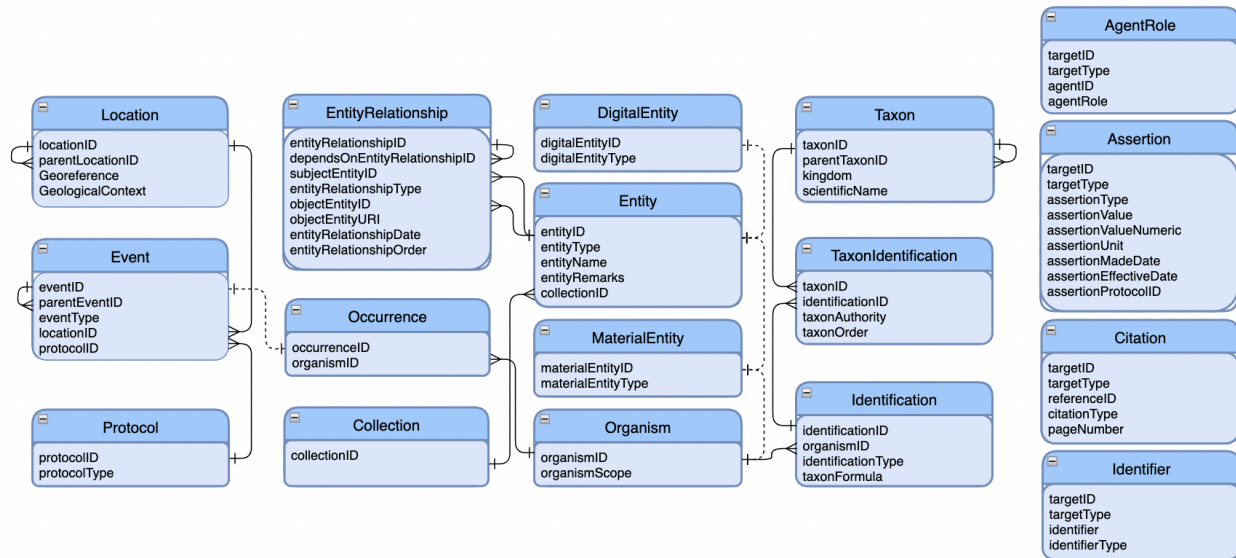


Figure 1: A conceptual model (2024-04-04) covering the assertions associated with an observed set of host / parasitoid / hyperparasitoid / endosymbiont relationships in *Melitaea cinxia / Hyposoter horticola / Mesochorus stigmaticus / Wolbachia pipientis* without reference to specific specimens or DNA extractions.

The long-term ecological study of the metapopulation dynamics of *Melitaea cinxia* in the Åland Islands can be captured as a parent-most Event. Within the project the specific study resulting in the cited manuscript can be captured as another Event. The five distinct localities in the study

form the set of Location records. The combination of Location and time of collection of host caterpillars form specific collecting Events within the study. Individual caterpillars collected in these Events and reared in the laboratory are target Organisms (each an EntityOfInterest). Each parasitoid and hyperparasitoid detected during rearing is also a target Organism (and also an EntityOfInterest).

Though the butterflies, parasitoids and hyperparasitoids were collected (together), that aspect of the study is not modeled here as it is covered in other use cases and allows this model to focus on the relationships between the Organisms. DNA extracts from the wasps identifying the presence of the endosymbiont bacteria are modeled simply as additional Organisms participating in the collecting Event because the extract, sequence and identification aspects of the study are also covered in other use cases. Thus, Taxon Identifications of Organisms here are a simplification that represents two distinct processes, one based on morphology and one based on DNA.

The relationships between Organisms are modeled as Assertions involving pairwise combinations, each joined with the specific relationship type that binds them (see Table 1). In this case study, for example, non-parasitized butterflies would have no relationship records; a parasitized butterfly would have a "parasitizedBy" relationship to a parasitoid and the parasitoid would have a complementary "parasitized" relationship to the butterfly; a hyperparasitized *Hyposoter* individual would have a "parasitizedBy" relationship to a hyperparasite *Mesochorus cf. stigmaticus* and the hyperparasite would have a "parasitized" relationship to the *Hyposoter* individual; *Hyposoter* individuals infected with the endosymbiont would have a "hadEndosymbiont" relationship to the bacteria (modeled as an Organism) and the bacteria would have a "wasEndoSymbiontOf" relationship to the *Hyposoter* individual.

| ID | Subject | Relationship Type* | Object |
|---|---|---|---|
| 1 | (A parasitized butterfly) Mc1 | parasitizedBy | (A parasitic wasp) Hh1 |
| 2 | Hh1 | parasitized | Mc1 |
| 3 | Hh1 | hadEndosymbiont | (Some bacteria) Wp |
| 4 | Wp | wasEndoSymbiontOf | Hh1 |
| 5 | Hh1 | parasitizedBy | (A parasitic wasp) Ms1 |
| 6 | Ms1 | parasitized | Hh1 |

Table 1. Examples of possible Relationships between hosts, parasitoids and a bacterium. The ID column represents an identifier for the Relationship rows. Mc1 identifies a specific individual *Melitaea cinxia,* Hh1 identifies a specific individual *Hyposoter horticola, Ms1* identifies a specific individual *Mesochorus cf. stigmaticus*, and *Wp* denotes the presence of the bacteria *Wolbachia pipientis.* *Note that string literal values for relationship type have been given here, but community-curated controlled vocabulary IRIs should be supported.

In addition to the pairwise relationships given above, it would be possible to posit (using the dependentOnRelationshipID) that a hyperparasitoid relationship existed between an individual butterfly and a *Mesochorus* individual and that this relationship depended on the relationship expressed in the parasitization of the butterfly by a *Hyposoter* individual. For example, the value of dependentOnRelationshipID for row 5 in Table 1 would be either 1 or 2.

An arbitrary number of **Assertions** can be made about each class. **Assertions** can be quantitative or qualitative and can have **Assertions** made about them as well. **Agents** can have roles with respect to any class as well, including **Assertions**. Instances of any class may be referenced in **Citations** and have alternate **Identifiers**. These four common aspects of all use cases can be found in Common Models.

~~An arbitrary number of **Assertions** can be made about each class. **Assertions** can be quantitative or qualitative and can have **Assertions** made about them as well. **Agents** can have roles with respect to any class as well, including **Assertions**. These two common aspects of all use cases can be found in Common Models.~~
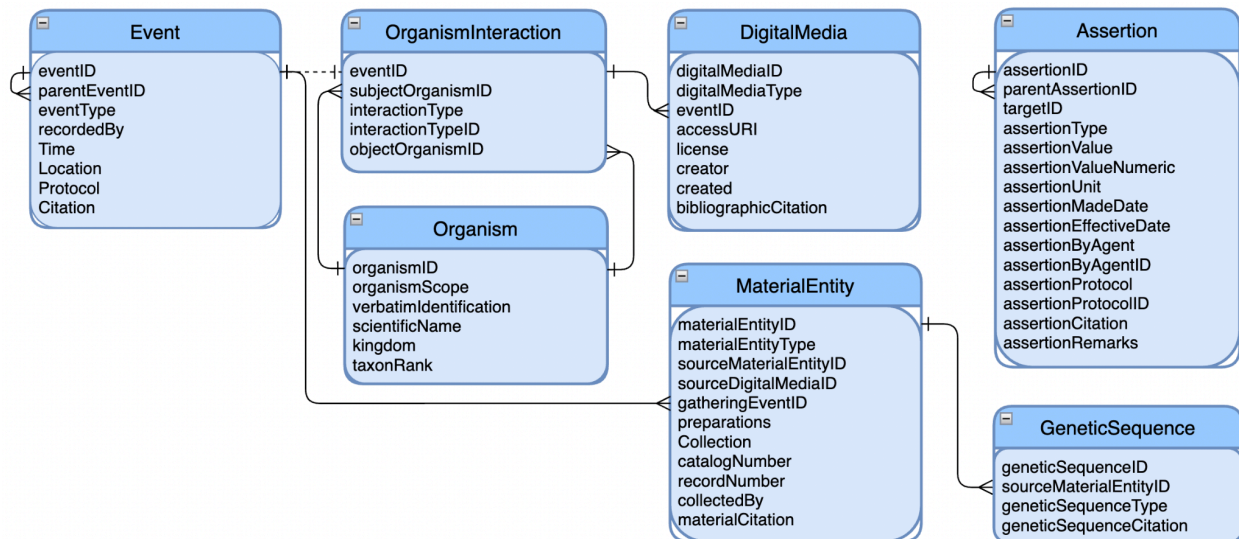
# Publishing Model



Figure 2. A simplification of the conceptual model for the purpose of publishing Events associated with biotic interactions as interactions between Organisms. Included fields are not exhaustive. Each class may be accompanied by Assertions, AgentRoles, Citations, and Identifiers, of which only the example of Assertions is shown (2023-11-26).

**Location** - one row for each distinct Location at which at least one **Event** occurred.
    **Terms** - all terms from *dwc:Location* plus *preferredSpatialRepresentation.*
**Event** - one row for each **Event** in which at least one **Organism** is present.
    **Terms** - *eventType* to declare the nature of the **Event**. All terms from *dwc:Event* plus *recordedBy.*

**Organism** - one row for each **Organism** present in the **Event**.
>**Terms** - *organismScope* can be used to distinguish that hosts and parasitoids are individuals, and that the bacteria is present but uncountable.

**MaterialEntity** - one row for each **Organism** present in the **Event**. Potentially one row for each derivative **MaterialEntity** as well.
>**Terms** - *materialEntityType* ("ORGANISM") when referring to a **dwc:Organism**. Can be used to distinguish parts of **Organism**, such as tissue samples used for DNA sequencing.

**OrganismInteraction** - one row for each interaction **Event** establishing the relationship of a subject **Organism** to an object **Organism**.
>**Terms**: event*ID* (primary key to for interaction **Events** observed between **Organisms**), *subjectOrganismID* (foreign key to the **Organism** on the subject side of the **OrganismInteraction**), *interactionType* to establish the relationship of the subject **Organism** to the object **Organism**, *objectOrganismID* (foreign key to the **Organism** on the object side of the **OrganismInteraction**).

# References

Common Models

https://github.com/gbif/rs.gbif.org/tree/master/core
https://rs.gbif.org/extension/
https://tools.gbif.org/dwca-validator/extensions.do
https://tools.gbif.org/dwca-validator/extension.do?id=dwc:Event
https://rs.gbif.org/core/dwc_event_2016_06_21.xml
https://tools.gbif.org/dwca-validator/extension.do?id=dwc:Identification
https://rs.gbif.org/extension/dwc/identification.xml
https://tools.gbif.org/dwca-validator/extension.do?id=dwc:ResourceRelationship
https://github.com/gbif/rs.gbif.org/blob/master/extension/dwc/resource_relation_2018_01_18.xml
https://github.com/tdwg/interaction
https://github.com/globalbioticinteractions/globalbioticinteractions

Duplouy A, Nair A, Nyman T, van Nouhuys S. 2021. Long-term spatiotemporal genetic structure of an accidental parasitoid introduction, and local changes in prevalence of its associated Wolbachia symbiont. Molecular Ecology 30(18), 4368-4380. doi:10.1111/mec.16065

Poelen JH, Simons JD, Mungall CJ. 2014. Global Biotic Interactions: An open infrastructure to share and analyze species-interaction datasets. *Ecological Informatics*, 24, 148–159. doi:10.1016/j.ecoinf.2014.08.005

José A Salim, Antonio M Saraiva, Paula F Zermoglio, Kayna Agostini, Marina Wolowski, Debora P Drucker, Filipi M Soares, Pedro J Bergamo, Isabela G Varassin, Leandro Freitas, Márcia M Maués, Andre R Rech, Allan K Veiga, Andre L Acosta, Andréa C Araujo, Anselmo Nogueira,

# Acknowledgements