Response to review:

Many thanks for mapping our dataset and assessing against the Darwin Core standards used by GBIF. The version of the dataset that you reviewed has undergone changes to ensure that it meets the requirements of GBIF and I hope that these changes will allow ingest into GBIF. I have been in contact with the data support team at GBIF and expect that the current format will allow incorporating both the host and pathogen data, but please let me know if this is not the case.

The data has been converted into a long format with each row representing a single occurrenceID and therefore we no longer have the multiple tabs.

The following are specific changes suggested in the review that have been completed. A more substantial change is the above mentioned conversion to long format which should be compliant with GBIFs Darwin Core.

datasetName has now been harmonised based on the title of the contributing study

parentEventID now refers to a unique identifier for datasetName

samplingProtocol has not been implemented as there isn’t an easy way to standardise the description across records

bibliographicCitation is now more complete combining, authors, year, title, journal and subsequent identifiers if available

copyright information has been split into license and accessRights. I am not too clear generally about licensing, my understanding is that the information from the articles can be used as long as the publication itself isn’t copied for those that are all rights reserved. I cannot find any clear information about this online.

identificationRemarks have been added for pathogen records only.

eventID has been added which as unique identifier of the event based on sampling visit and location within a study.

Year and month have been combined into the relevant iso standard date in eventDate

Region of trapping has been entered as locality and verbatimLocality (for the more habitat description sites)

All coordinates have been converted to decimalLatitude and decimalLongitude with the correct geodeticDatum added and coordinate uncertainty estimates

Descriptives of trapping location have been removed

Count of individual host species has been renamed to individualCount

associatedOccurrences have been used for pathogens with organismQuantity rather than individualCount for the number of tested individuals that were positive. The number of individuals tested has been added as an occurrenceRemark.

Sampling effort has not been included.