Natural England Commissioned Report NECR324

2020 Gap Analysis of the BOLD Database for Key English Invertebrates

First published October 2020

www.gov.uk/natural-england

Foreword

Natural England commission a range of reports from external contractors to provide evidence and advice to assist us in delivering our duties. The views in this report are those of the authors and do not necessarily represent those of Natural England.

Background

DNA – based methods offer a significant opportunity to change how we monitor and assess biodiversity. These techniques may provide cheaper alternatives to existing species monitoring or an ability to detect species that we cannot currently detect reliably.

However, for most species, there is still much development required before they can be used in routine monitoring. Natural England has been exploring the further use of these methods for environmental monitoring for several years, delivering a series of reports which focus on the development of DNA-based methods with potential in a particular area.

One area of focus has been terrestrial invertebrate assemblages, which are time-consuming and difficult to monitor traditionally, for several years.

One of the main barriers preventing the further uptake of these methods are the gaps in current DNA reference libraries for these species.

This report is the first step towards rectifying this by providing a gap analysis to identify which of the 13,690 taxa listed in Natural England's Pantheon database for invertebrates are represented in the widely used Barcode of Life Database (BOLD).

This report should be cited as: Macadam, C., Robins, J. & Thomson, T. (2020) 2020 Gap Analysis of the BOLD Database for Key English Invertebrates, *Natural England Commissioned Report Number 324*

Natural England Project Manager – Jon Webb, Principal Adviser, Natural England: jon.webb@naturalengland.org.uk

Contractor – Buglife - The Invertebrate Conservation Trust, Peterborough.

Keywords –DNA reference libraries, invertebrates, Barcode of Life, Tree of Life

Further information

This report can be downloaded from the Natural England Access to Evidence Catalogue: **http://publications.naturalengland.org.uk/**. For information on Natural England publications contact the Natural England Enquiry Service on 0300 060 3900 or e-mail **enquiries@naturalengland.org.uk**.

ISBN 978-1-78354-670-1

2020 Gap Analysis of the BOLD Database for Key English Invertebrates

21st May 2020

Craig Macadam, Jamie Robins & Tom Thomson

Saving the small things that run the planet

Contents

METHODS

RESULTS

CONCLUSIONS

Executive Summary

Recent work commissioned by Natural England (in prep) shows that DNA metabarcoding is currently not as reliable as traditional identification methods. Natural England wishes to reach a point where barcoding has at least a 95% success rate when compared to traditional identification methods. Commissioned by Natural England, Buglife undertook a gap analysis to identify which of the 13,690 taxa listed in Pantheon (https://www.brc.ac.uk/pantheon/) are represented in the widely used Barcode of Life Database (BOLD). For BOLD to be a useful tool, the provenance and accuracy of barcode data must be investigated. The most reliable records being those which are based on UK sourced specimens, which have been properly verified and with vouchers stored in recognised institutions. After analysis of the initial 5 million records in BOLD, 770 records associated with Pantheon taxa with a high confidence were identified, representing only 168 species. Medium confidence records were notably greater, with 100,701 records representing 3,025 species.

A workshop was held to bring together the entomological community and geneticists to discuss approaches to filling the gaps in barcoding identified above. Group sessions identified effective ways to fill the data gaps through novel opportunities for partnership working. There is clear potential for the better use of current monitoring projects to supply large volumes of genetic samples, but also for organisations and institutions to work with entomologists to undertake targeted sampling for barcoding purposes should they be connected and resources be made available.

The sizeable gap in the BOLD dataset makes it difficult to make meaningful conclusions regarding any bias in the dataset and to prioritise future targeted efforts to fill the data gaps. However, a centralised approach is urgently needed to build on the current momentum and funding for barcoding and it is recommended that future efforts align with the Darwin Tree of Life project to prevent duplication of effort. To support this, a regularly updated, dynamic list of barcoding gaps is required alongside agreed protocols for collection and sampling. Rapid filling of the data gaps could then be undertaken through organised sampling events and connections to active recording groups, as well as investigating the feasibility of using material collected through large national monitoring schemes to provide a potential source of bulk samples.

Background

Recent work commissioned by Natural England (NE) (Natural England, in prep.) shows that DNA metabarcoding is currently not as reliable as traditional identification methods for invertebrates. Natural England wishes to reach a point where barcoding has at least a 95% success rate when compared to traditional identification methods. Gaps in the Barcode of Life Data System (BOLD) database have been identified as one of the key blockers for enabling metabarcoding of invertebrates in England. It has been suggested that BOLD has barcode sequences for c.75% of the UK invertebrate fauna, but it is unclear what proportion of these are reliable. Over time, Natural England wants this to be increased to 100% coverage and reliability for all key taxa for metabarcoding to be workable in its monitoring programmes.

Buglife was commissioned to undertake:

- Module 1: Gap Analysis of invertebrates in Pantheon. Produce an Excel output for each taxa in Pantheon and the associated critical information, and the presence and reliability of any records in BOLD. This includes summary outputs for percentage cover of families and orders, and percentage coverage by Pantheon information such as habitat association, biotope and conservation status.
- Module 2: Workshop and an Outline Action Plan for Closing the Gaps in the BOLD database for Invertebrates. A workshop to bring together field active invertebrate taxonomists and geneticists to report on the methods and incentives which could be employed to fill gaps in the BOLD database.
- Module 3. Use the United Kingdom Species Inventory (UKSI) to describe the major groups and families not covered in Pantheon, indicate their coverage in BOLD and prioritise a future gap analysis project for this group.

This report details the findings of these three modules and makes recommendations for addressing the gaps identified.

Methods

Module 1

Data was obtained from two sources:

Pantheon (www.brc.ac.uk/pantheon) is an analytical tool developed by Natural England and the Centre for Ecology and Hydrology (CEH) to assist invertebrate conservation in England (Webb et al 2018). The tool allows you to determine the associated habitats, conservation status and ecology of inputted lists of invertebrate species.

The Barcode of Life Database (BOLD) (www.boldsystems.org/index.php) is a cloud-based repository for reference sequences from vouchered specimens of all species of life. The database currently holds over 5.4 million records of sequences from more than 205,000 species.

An initial data extract was made from BOLD. All public records for the target groups were downloaded and imported into a SQL database. This resulted in over 5 million records of 219,251 world taxa (Table 1)

Table 1 Number of records extracted from BOLD

Taxon group	Number of records	Number of species
Arthropoda	4,772,102	198,959
Mollusca	159,049	14,557
Annelida	45,921	3,181
Platyhelminthes	23,764	2,339
Bryozoa	2,151	215

A shared taxonomy was created by building a lookup table comparing values for the **Tax_ID** from BOLD and the **TVK field** from Pantheon. The records obtained from BOLD were then matched with taxa present in Pantheon to create a working dataset. This resulted in a dataset containing 314,100 records for 9,691 taxa.

This working dataset was further refined to include only records that have a valid sequence (defined as all records where the **nucleotides** field from BOLD is not blank). This resulted in a final working dataset of 307,414 records for 9,567 taxa.

Natural England was particularly interested in the provenance of sequences in BOLD. It was recognised that some identification may have originated from more reliable sources than others, and identifying this reliable subset was an essential element of this project. The highest confidence was to be given to records that are supported by curated specimens, identified by a credible taxonomist and linked with a publication. Following discussion with Natural England the following criteria were used to determine confidence in the provenance of the records:

- Whether a voucher specimen was available
- Whether the specimen was from the UK
- Whether the specimen is stored in an institution
- Whether the identification of the specimen was made by an experienced worker.

A set of filters were applied to the working dataset (see Figure 1) via scripts to populate new fields for these criteria as follows:

- **Voucher_class** - Binary field where 1 = all instance of the **Voucher_Status** field from BOLD that contains information of a morphological voucher specimen (not just a DNA or partial voucher) defined by manual creation of a lookup table classing all distinct values of **Voucher_Status** (see Annex 5).
- **Institution_storing_class** - Binary field where 1 = all instance of the **Institution_storing** field from BOLD that contains information of a recognised institution or faculty where a morphological voucher specimen is listed (not private collections) defined by manual creation of a lookup table classing all distinct values of **Institution_storing** (see Annex 5).
- **Country_class** - Binary field where 1 = all instance of the "**Country**" field from BOLD that contains UK, GB, England, Scotland, Wales or other permutations thereof.
- **Identifier_class –** Binary field where 1 = all instances of the "**Identification_provided_by**" field from BOLD that contains information of a recognised expert on the taxa to which the record relates.

These fields were then used to determine the confidence that could be applied to each record as follows:

> Medium confidence – defined as all records for species found in Pantheon that also have a valid sequence as above, and **voucher_class** and **Institution_storing_class** = 1.
>
> High confidence – defined as all records for species found in Pantheon that also have a valid sequence as above, and **voucher_class**, **Institution_storing_class**, and **Country_class** = 1.

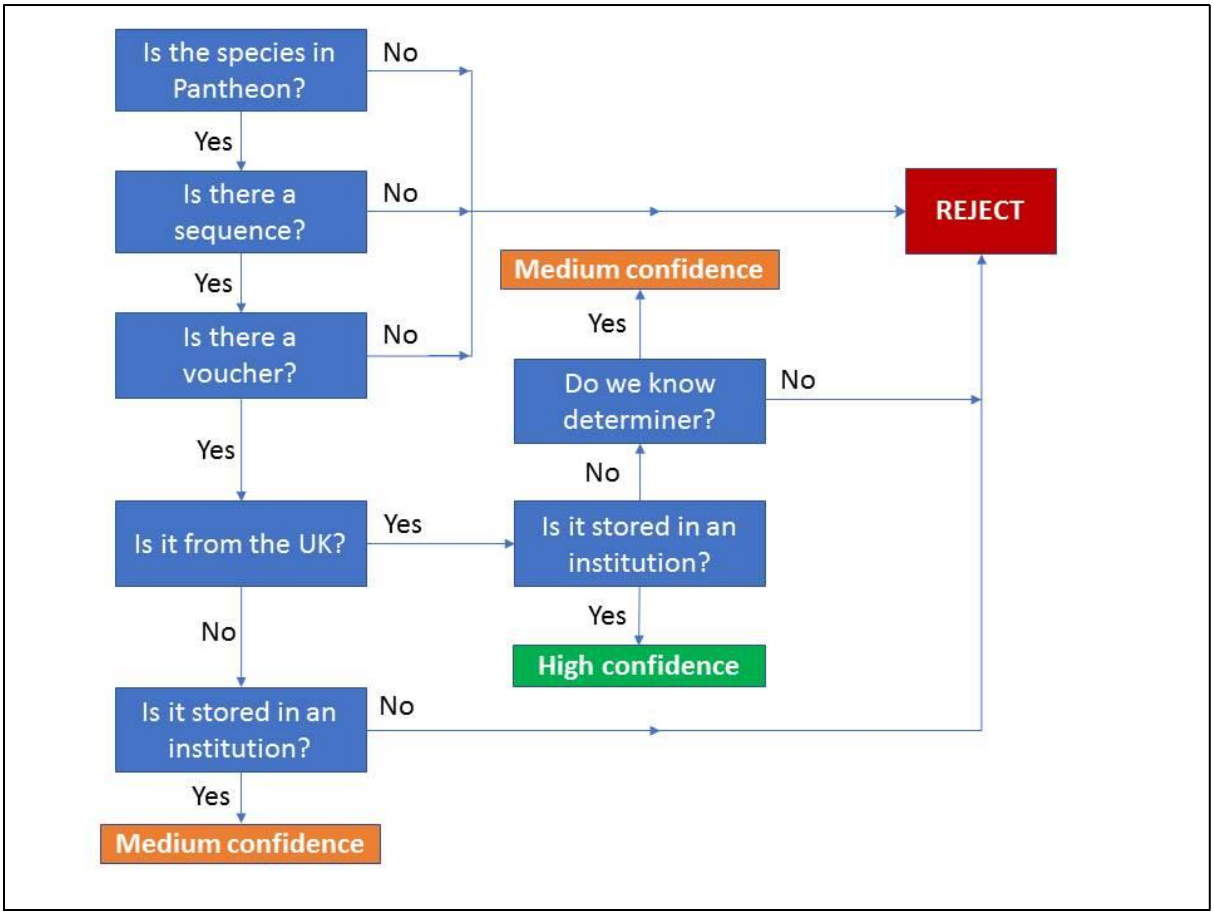

Figure 1 Process for analysing BOLD records

The resultant dataset of 101,471 records for 3,193 species was extracted to an Excel spreadsheet. In addition to the fields from BOLD, data on the Biotope association, Conservation Status and Site Quality Score, where available, were added from Pantheon (see Appendices).

Module 2

A workshop was organised in partnership with Natural England, engaging invertebrate taxonomists and geneticists, to identify the methods and incentives which could be employed to fill gaps in the BOLD database. A series of presentations were delivered considering the different methodologies of barcoding and genomic sequencing, including their applications and use in existing academic projects.

A roundtable session provided each individual or organisation with the opportunity to highlight their own interests and to start discussions identifying overlaps and novel ways of working together. Group sessions were then held to explore how entomologists and geneticists could collaborate, with the ultimate aim of finding ways to fill the gaps in the barcoding database. The development of an Outline Action Plan for filling gaps in the database was planned but following the results of Module 1 and discussions with Natural England, this was not progressed.

Module 3

Despite the wealth of information in Pantheon, there are many other terrestrial and freshwater species which are poorly represented in Pantheon, including sawflies (Symphyta), Parasitica (Ichneumons and their allies), Psocoptera (booklice) and many other groups, consisting of several thousand species. The intention had been to use the United Kingdom Species Inventory (UKSI) to describe the major groups and families not covered by Pantheon, summarise their representation in BOLD and to prioritise a future gap analysis. However, following on from the results of Module 1 and discussions with Natural England, this module was not progressed.

Results

Module 1: Gap Analysis

Figures 2 and 3 show the number of matching records and species respectively at each stage following the data sifting process. The number of BOLD records associated with Pantheon taxa in which we can have high confidence was small with 770 records representing 168 species. The number of records with medium confidence was considerably greater with 100,701 records representing 3,025 species.

The high confidence records were dominated by Hymenoptera (Bees, wasps, ants and sawflies) with 150 species (89% of records). The medium confidence records were dominated by three groups: Diptera (True flies) with 945 species (30%); Lepidoptera (Butterflies & moths) with 742 species (24%); and Coleoptera (Beetles) with 610 species (20%). Table 2 provides the full breakdown for each Class/Order.

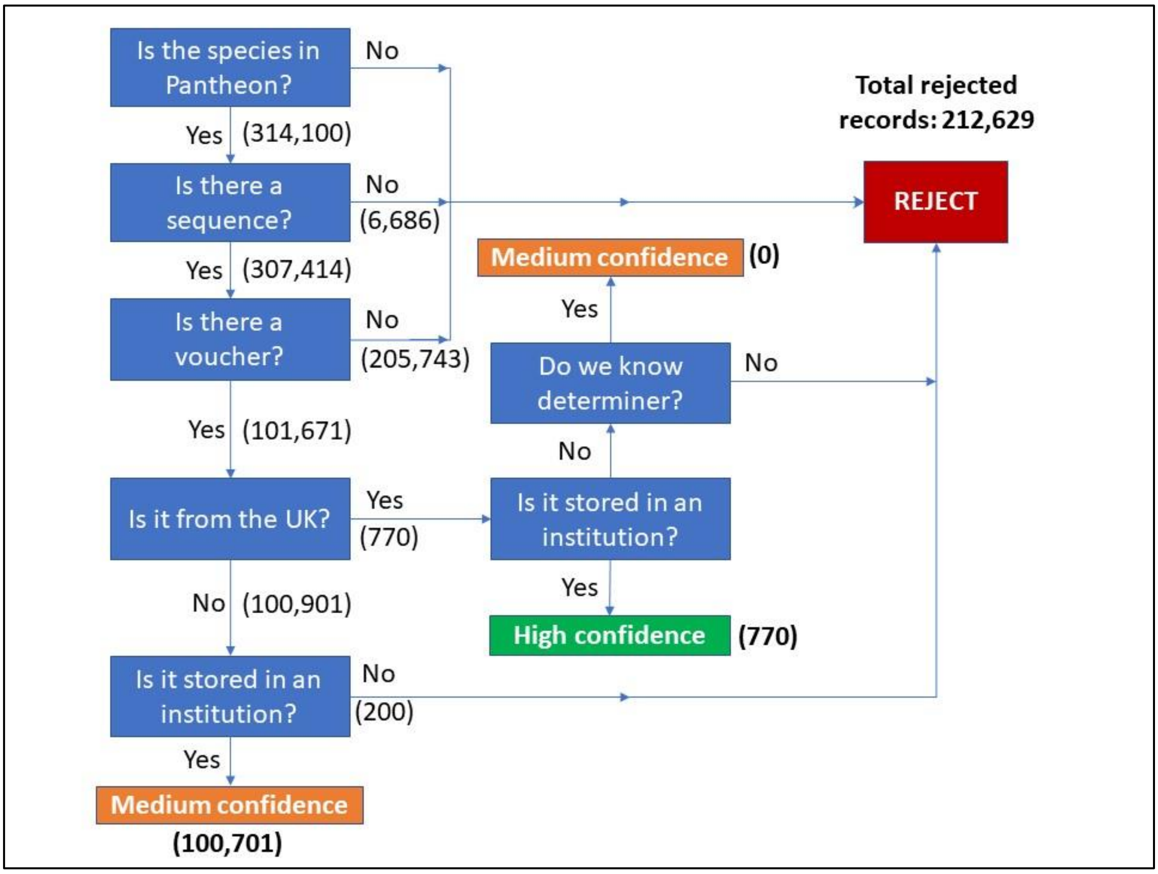

Figure 2 Number of records in each confidence category.

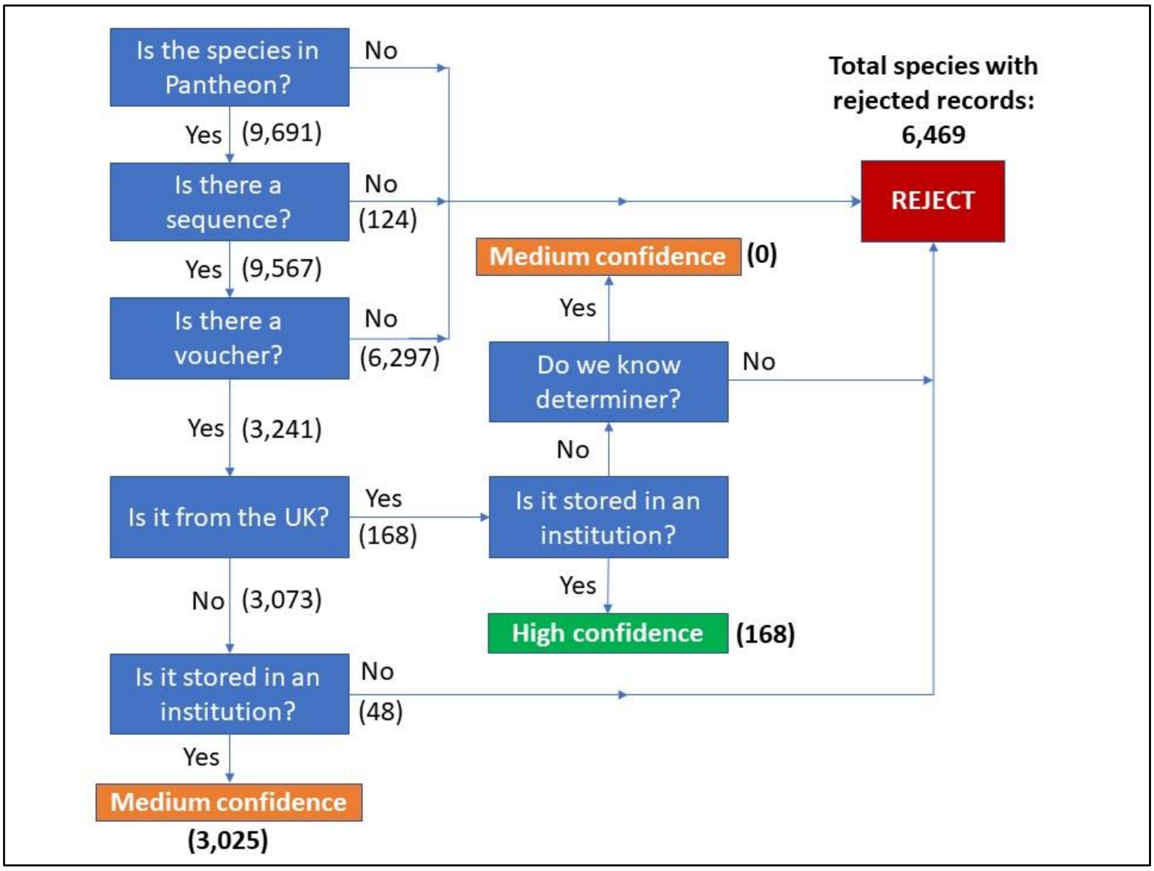

Figure 3 Number of species in each confidence category

It became evident during this analysis of the working dataset that barcodes originating from the Natural History Museum (NHM) FreshBase project (469 records) and the NE/NHM DNA Bioblitz at Ainsdale National Nature Reserve in 2019 (873 records) were not yet recorded on BOLD. These two datasets provide an additional 454 species with high confidence records, however the full details, including sequences are not available yet so these have not been included in the working dataset. A summary of the coverage of records expected from current barcoding projects is included in Table 2.

Table 2: Number of species with medium and high confidence records in BOLD and expected high confidence records from current barcoding projects. To avoid duplication any species present from more than one data source was only included once. All species from BOLD are included in the relevant columns of the table.

Class	Order	Common name	BOLD Medium	BOLD High	FreshBase	Bioblitz
Arachnida	Araneae	Spiders	188		1	1
Arachnida	Opiliones	Harvestmen	6			
Arachnida	Pseudoscorpiones	Pseudoscorpions	5			

Class	Order	Common name	BOLD Medium	BOLD High	FreshBase	Bioblitz
Bivalvia		Mussels & clams	12		8	
Branchiopoda	Cladocera	Water fleas	1		14	
Chilopoda		Centipedes	9	1		
Clitellata		Annelid worms	2		9	
Collembola		Springtails	1			
Diplopoda		Millipedes	13	2		
Gastropoda		Slugs & snails	25		22	
Gymnolaemata		Moss animals	1			
Hexanauplia		Copepods	2			
Hydrozoa	Anthoathecata				1	
Insecta	Blattodea	Cockroaches	8			
Insecta	Coleoptera	Beetles	610		100	8
Insecta	Dermaptera	Earwigs	2			
Insecta	Diptera	True-flies	945	11	34	20
Insecta	Ephemeroptera	Mayflies	4		44	
Insecta	Hemiptera	True bugs	217	1	20	
Insecta	Hymenoptera	Bees, wasps, ants & sawflies	233	150	1	
Insecta	Lepidoptera	Butterflies &moths	742	3	1	23
Insecta	Mantodea	Mantids	1			
Insecta	Mecoptera	Scorpionflies	2			
Insecta	Megaloptera	Alderflies			2	
Insecta	Neuroptera	Lacewings	13		1	
Insecta	Odonata	Dragonflies & damselflies	8		16	
Insecta	Orthoptera	Grasshoppers & crickets	7			
Insecta	Plecoptera	Stoneflies	7		23	
Insecta	Thysanoptera	Thrips	1			
Insecta	Trichoptera	Caddisflies	20		71	
Malacostraca	Amphipoda	Shrimps	12		9	
Malacostraca	Decapoda	Crabs and crayfish	15			
Malacostraca	Isopoda	Woodlice	13		2	
Malacostraca	Mysida	Mysid shrimps			1	
Malacostraca	Stomatopoda	Mantis shrimps	1			
Podocopa	Podocopida				1	
Polychaeta	Polychaeta	Bristle worms	5		1	
Polyplacophora	Polyplacophora	Chitons	1			

Class	Order	Common name	BOLD Medium	BOLD High	FreshBase	Bioblitz
Pycnogonida	Pycnogonida	Sea spiders	1			
Turbellaria	Seriata	Freshwater flatworms			4	
Total			**3133**	**168**	**403**	**51**

Coverage by habitat

Analysis of the Pantheon attributes for Biotope show that 1,089 species with high or medium confidence are associated with open habitats, 836 are associated with trees, and 370 species are associated with freshwater habitats (Table 3). 723 species had no biotope association noted in Pantheon. Taking in to account the additional species present in the FreshBase and NE Bioblitz datasets the number of freshwater associated species increases to 680 and there are small increases in other categories.

Table 3: Species coverage by broad biotope (as defined in Pantheon)

Habitat	BOLD only	All records
Open habitats	1089	1104
Tree-associated	836	840
No association listed	723	823
Wetland	370	680
Open habitats; tree-associated	96	96
Coastal	39	53
Tree-associated; wetland	38	46
Open habitats; wetland	16	16
Coastal; wetland	9	12
Coastal; open habitats	6	6
Open habitats; tree-associated; wetland	3	3

Module 2: Workshop Outputs

Following the initial data analysis, a workshop was held in Peterborough on 3rd December 2019, titled ‘A 2020 Gap Analysis of the BOLD Database for Key Invertebrates’ to bring together invertebrate taxonomists and geneticists and identify the most efficient way of filling gaps in the coverage of the BOLD database. Following the workshop, a brief workshop report was circulated (Annex 1) including a summary of key findings:

- A dynamic and up to date list of gaps for all target fauna needs to be maintained and easily accessible.
- It may be useful to pilot an approach to filling the gaps, to help identify the useful mechanisms and challenges.
- Common, agreed protocols on the collection and storage of samples are essential and need drafting and distributing to collectors. Good meta-data standards also need to be drafted and adhered to.
- Fresh specimens are best but dry museum specimens still have a role to play.
- Museums should lead on barcoding museum specimens, but coordination is needed to produce a dynamic database/list of species still requiring collection.
- Could regional hubs operate out of museums or record centres?
- Priority to ensure accurate species identification in all groups, but it might also be useful to utilise non-specialist collectors to provide materials in bulk initially.
- The use of by-catch and materials from ongoing monitoring represent a high value opportunity to get many common species.
- Recording groups represent a significant opportunity for targeted gap-filling particularly of less common species, but some volunteer-led groups may struggle to deliver the large, regular samples required of some projects.
- Multiple specimens would be preferred and from different localities to capture regional variation.
- Centralising Site of Special Scientific Interest (SSSI) permissions may help with collection of scarcer species.
- Do we need a decision tree for each taxa, where approaches or effort might vary?
- Regular communications (e.g. workshops, sharing contacts, website) is key to progress and maintaining enthusiasm of collectors. We could form a group from the workshop that keeps in contact.
- Feedback such as micro-publications and acknowledgements are key to incentivising engagement with collectors.
- Courier costs for passing on collected materials to museums needs to be covered as well as materials to collect and store specimens.

Conclusions

Module 1

Despite BOLD containing 314,100 records with nucleotide sequences for species in Pantheon, only a very small subset were of either medium or high confidence. In total, BOLD contained high confidence records for only 168 species in Pantheon and medium confidence records for 3,025 species. This small number of high confidence records represent the records of UK origin, verified and with voucher specimens retained in recognised institutions. The criteria used to identify these high confidence records should become the benchmark for future barcoding efforts for English invertebrates. Of the 9,567 Pantheon species with BOLD sequences, 66% failed to meet the criteria for either medium or high confidence on the basis of there being no retained voucher specimen, a key step in ensuring that barcoding is a reliable method moving forward.

The sizeable gap in the BOLD dataset makes it difficult to make meaningful conclusions regarding any bias in the dataset. Although open habitat, tree-associated, no association and wetland associated species were the dominant broad biotopes represented by Pantheon species in BOLD, there is insufficient data for any trends to be identified. Among orders, there is some variation in the percentage coverage in the BOLD database, but this is largely the result of small orders being relatively well addressed. For example, although the majority of high confidence records are of Hymenopteran species, only 1.89% of species are represented due to the large family size (over 8,000) species. In contrast the small families of Scolopendromorpha (predatory centipedes) and Chordeumatida (millipedes) have higher levels of coverage with 11.1% and 13.3% respectively. It is therefore feasible for some smaller orders to be relatively rapidly barcoded should the relevant taxon groups wish to engage. A good example of this approach is the NHM FreshBase project which has engaged with the Riverfly Recording Schemes and has achieved almost complete coverage in the collection of species of Ephemeroptera and Plecoptera found in the UK. Of the major orders, there are clearly significantly more medium confidence records with 11.7% of Hemiptera, 12.9% of Diptera, 14.3% of Coleoptera, 26% of Lepidoptera and 27% of Araneae for example, but still notable gaps across all taxa.

There was some notable difficulty in linking BOLD with Pantheon, due to synonyms and taxonomic descriptions such as sub-species, varieties and forms and sensu lato, aggregations and sensu stricto multi-listings. A single custom taxonomy was developed using the UKSI as the benchmark for taxa and the NBN species dictionary as a higher taxonomy classifier to assign species to the correct families and orders. The inclusion of full taxonomic information in the UKSI would make this process less labour intensive in the future.

Module 2

There is clearly an appetite within the entomological community to work with geneticists and fulfil Natural England’s ambition to obtain barcodes for invertebrate species. Such a partnership could source UK samples which can be appropriately stored and verified to ensure the quality of future barcoding. However, the current data indicates there are large gaps in all invertebrate groups. The view of participants at the workshop is that a purely ad hoc approach to filling the sizeable barcoding gap is unlikely to succeed. There are significant resources linked to ongoing projects such as the Darwin Tree of Life Project and BIOSCAN projects and a strong willingness for cooperation between geneticists and entomologists. However, it is important to align this effort with the interests and capacity of the entomological community. Many of the current drives to collect barcode data require large numbers of samples to be provided to make the analysis economically viable, however entomological groups are unlikely to be able to sustain this scale of effort. Some existing sources of invertebrate material from ongoing large-scale environmental monitoring projects may offer the opportunity to source samples at the required scales, but there may still be a shortage of experts to identify material.

As demonstrated by the additional records from FreshBase and Bioblitz data in Table 2, organised barcoding activities can make significant in-roads to filling the current barcoding gaps. Whilst the primary aim of these initiatives was similar, i.e. to collect specimens for metabarcoding, the different approaches taken resulted in quite different coverage of species. Whilst the Bioblitz approach collected large numbers of specimens in a short period of time, many of the species collected already had sequences available in BOLD, resulting in fewer gaps being covered. In contrast, the FreshBase approach to target individual species for collection has resulted in good specimen coverage for selected groups, albeit over a longer period of time.

It is likely that more common species could be relatively easily sourced and identified if efforts are made to obtain by-catch from large ongoing monitoring schemes. This would require considerable entomological support to process samples, but with the support of some of the larger taxon specialist groups could lead to some orders being quickly addressed. Targeted survey work using specialist groups provided with appropriate materials for field surveys are likely to be essential to collect and correctly identifying species associated with rare or restricted habitat types. Such visits could produce lists similar to the FreshBase and Bioblitz outputs examined in this report, but would require coordination and incentivising, notably through providing the required field materials and providing feedback such as micro-publishing and acknowledging collectors within barcode entries.

It is also highly likely that significant barcoding data is not finding its way to BOLD. Moving forward it is essential that a single repository of data such as BOLD be used as part of a concerted effort to build a UK barcode library. For a single tool to be the main library for UK invertebrate species barcoding, it is essential that all current workers in the field recognise the merit of a coordinated approach and use it as their focal point for storing barcode data.

To aid this process and in particular to muster efforts in the entomological community, there is a need for a dedicated coordinator that can make connections with entomologists, provide feedback, and coordinate collection efforts to make them as efficient as possible. This would, for example, include maintaining an up to date list of the species that are required to prevent duplication of effort.

The workshop did not decide upon a specific focus for filling gaps, but rather it was agreed that we should work towards a partnership approach to sourcing and processing samples wherever suitable opportunities arise. The substantial gaps in BOLD for the UK invertebrate fauna does however mean that any increased collection and sequencing activity will be very useful and will ultimately lead to an improvement in the usefulness of metabarcoding techniques in invertebrate surveys.

Recommendations

- With limited coverage of UK invertebrates in BOLD at this stage, it is recommended that future barcoding is aligned with the Darwin Tree of Life project (https://www.darwintreeoflife.org/). This will prevent duplication of effort and will ensure that approaches to sampling are standardised and key potential partners have a single point of contact for metabarcoding.
- A regularly updated list of species collected should be maintained and made available to entomologists and recording groups to ensure that collecting effort is focussed on gaps in the barcode library.
- Standard protocols for collection, handling of samples and data recording are required to ensure that material and the associated meta-data obtained is of the highest quality.
- Targeted sampling should be prioritised for specific habitats or specific taxon groups where there are rare species that are unlikely to be encountered otherwise.
- Organised sampling events in different geographical areas should be used to collect a wide range of both common and geographically restricted species.
- The process of gaining permission for collecting on SSSIs should be reviewed to see if a general permit could be issued to collectors for all SSSIs in an area or region rather than individual permissions having to be sought.
- The feasibility of accessing samples from large monitoring schemes (e.g. the Rothamsted light trap network or the Pollinator Monitoring Scheme pan trap network) should be explored, to provide a potential source of common species in bulk.
- Regular feedback and acknowledgment of the effort of collectors is essential to maintain engagement with the entomological community.
- Improvements to the linkages between the UKSI and BOLD, in particular the nomenclature and taxonomy used would be useful to assist with future repeats of this exercise.

Acknowledgements

Buglife thanks Jon Webb, Debbie Leatherland and Katie Clark at Natural England for their advice and support during the completion of this report. Thanks also to the workshop speakers and participants who contributed their expertise and experience. Thanks to Martin Harvey from the Centre for Ecology and Hydrology for assistance with the Pantheon Tool.

References

Webb, J., Heaver, D., Lott, D., Dean, H.J., Van Breda, J., Curson, J., Harvey, M.C., Gurney, M., Ry, D.B., Van Breda, A., Drake, M., Alexander, K.N.A. and Foster, G. 2018. Pantheon - database version 3.7.6.

List of Annexes

Annex 1: Gap Analysis Workshop Report

Annex 2: The BOLD-Pantheon Dataset

Annex 3: Data processing technical flowcharts

Appendices

Appendix 1: Percentage coverage of different orders from the UK Species Inventory represented in Pantheon and in the BOLD database.

Order name	No of Species in UK (UKSI & NBN)	No of Species in Pantheon	% of Family represented in Pantheon	% of Family represented in BOLD by species with >1 High Confidence record	% of Family represented in BOLD by species with >1 Medium Confidence record
[Annelida]	1519	42	2.76%		0.86%
[Bryozoa]	333	7	2.10%		0.30%
[Mollusca]	2316	272	11.74%		1.30%
Amphipoda	649	50	7.70%		1.85%
Anostraca	3	2	66.67%		
Aphelenchida	20	1	5.00%		
Araneae	712	677	95.08%		26.97%
Arguloidea	3	1	33.33%		
Bathynellacea	2	2	100.00%		
Bonelliida	4	1	25.00%		
Bothriocephalidea	15	1	6.67%		
Calanoida	352	2	0.57%		
Caryophyllidea	8	3	37.50%		
Chordeumatida	15	15	100.00%	13.33%	6.67%
Coleoptera	4320	2951	68.31%		14.31%
Collembola	359	4	1.11%		
Cyclopoida	157	1	0.64%		
Decapoda	234	35	14.96%		6.41%
Dermaptera	9	8	88.89%		22.22%
Dictyoptera	28	13	46.43%		28.57%
Diplostomida	66	1	1.52%		
Diplostraca	107	2	1.87%		
Diptera	7382	3625	49.11%	0.15%	12.91%
Dorylaimida	91	1	1.10%		
Ephemeroptera	52	51	98.08%		7.69%
Geophilomorpha	36	34	94.44%		13.89%
Glomerida	4	4	100.00%		
Gyrodactylidea	1	1	100.00%		
Hemiptera	1857	1123	60.47%	0.05%	11.69%
Hymenoptera	8116	374	4.61%	1.89%	2.97%
Isopoda	259	70	27.03%		5.41%
Ixodida	24	1	4.17%		

Order name	No of Species in UK (UKSI & NBN)	No of Species in Pantheon	% of Family represented in Pantheon	% of Family represented in BOLD by species with >1 High Confidence record	% of Family represented in BOLD by species with >1 Medium Confidence record
Julida	35	32	91.43%		22.86%
Lepidoptera	2852	2559	89.73%	0.11%	26.02%
Lithobiomorpha	21	20	95.24%		4.76%
Mantodea	3	2	66.67%		33.33%
Mecoptera	4	4	100.00%		50.00%
Megaloptera	3	3	100.00%		
Monopisthocotylea	3	3	100.00%		
Monostilifera	42	3	7.14%		
Myodocopida	25	1	4.00%		
Mysida	66	6	9.09%		
Neuroptera	74	62	83.78%		17.57%
Notostraca	1	1	100.00%		
Odonata	117	58	49.57%		6.84%
Opiliones	31	28	90.32%		19.35%
Orthoptera	56	50	89.29%		12.50%
Pantopoda	72	1	1.39%		1.39%
Phasmatodea	7	6	85.71%		
Plagiorchiida	258	1	0.39%		
Plecoptera	35	36	102.86%		20.00%
Podocopida	293	1	0.34%		
Poecilostomatoida	203	8	3.94%		
Polydesmida	18	16	88.89%		16.67%
Polyxenida	1	1	100.00%		100.00%
Polyzoniida	2	2	100.00%		
Pseudoscorpiones	27	29	107.41%		18.52%
Raphidioptera	4	4	100.00%		
Scalpelliformes	4	1	25.00%		
Scolopendromorpha	9	5	55.56%	11.11%	33.33%
Scutigeromorpha	1	1	100.00%		
Seriata	43	12	27.91%		
Sessilia	28	4	14.29%		7.14%
Siphonophorida	1	1	100.00%		
Siphonostomatoida	176	3	1.70%		
Solitaria	41	1	2.44%		
Spirobolida	2	2	100.00%		
Spirurida	1	1	100.00%		
Stomatopoda	2	1	50.00%		50.00%
Thysanoptera	188	1	0.53%		0.53%

Order name	No of Species in UK (UKSI & NBN)	No of Species in Pantheon	% of Family represented in Pantheon	% of Family represented in BOLD by species with >1 High Confidence record	% of Family represented in BOLD by species with >1 Medium Confidence record
Trichoptera	201	197	98.01%		9.95%
Triplonchida	14	1	7.14%		
Tylenchida	138	2	1.45%		

Appendix 2: Percentage of species in families covered for high profile invertebrate orders in Pantheon and in the BOLD database.

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Coleoptera	Aderidae	5	3	60%	0	0%	1	20%
Coleoptera	Aegialiidae	3	0	0%	0	0%	0	0%
Coleoptera	Alexiidae	2	2	100%	0	0%	0	0%
Coleoptera	Anobiidae	60	58	97%	0	0%	4	7%
Coleoptera	Anthicidae	16	13	81%	0	0%	2	13%
Coleoptera	Anthribidae	10	9	90%	0	0%	2	20%
Coleoptera	Aphodiidae	61	13	21%	0	0%	0	0%
Coleoptera	Apionidae	89	52	58%	0	0%	9	10%
Coleoptera	Attelabidae	2	2	100%	0	0%	0	0%
Coleoptera	Biphyllidae	2	2	100%	0	0%	0	0%
Coleoptera	Bolboceratidae	1	1	100%	0	0%	1	100%
Coleoptera	Bostrichidae	5	4	80%	0	0%	0	0%
Coleoptera	Bothrideridae	5	5	100%	0	0%	0	0%
Coleoptera	Buprestidae	24	12	50%	0	0%	8	33%
Coleoptera	Byrrhidae	13	13	100%	0	0%	1	8%
Coleoptera	Byturidae	2	2	100%	0	0%	1	50%
Coleoptera	Cantharidae	45	42	93%	0	0%	15	33%
Coleoptera	Carabidae	398	140	35%	0	0%	99	25%
Coleoptera	Cerambycidae	84	66	79%	0	0%	20	24%
Coleoptera	Cerophytidae	1	0	0%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Coleoptera	Cerylonidae	6	4	67%	0	0%	0	0%
Coleoptera	Cetoniidae	13	6	46%	0	0%	0	0%
Coleoptera	Chrysomelidae	309	296	96%	0	0%	61	20%
Coleoptera	Ciidae	23	22	96%	0	0%	0	0%
Coleoptera	Clambidae	11	3	27%	0	0%	0	0%
Coleoptera	Cleridae	14	13	93%	0	0%	1	7%
Coleoptera	Coccinellidae	66	44	67%	0	0%	20	30%
Coleoptera	Colydiidae	2	2	100%	0	0%	0	0%
Coleoptera	Corylophidae	14	8	57%	0	0%	0	0%
Coleoptera	Cryptophagidae	109	32	29%	0	0%	7	6%
Coleoptera	Cucujidae	3	2	67%	0	0%	0	0%
Coleoptera	Curculionidae	531	366	69%	0	0%	74	14%
Coleoptera	Cybocephalidae	1	0	0%	0	0%	0	0%
Coleoptera	Dascillidae	1	1	100%	0	0%	0	0%
Coleoptera	Dasytidae	9	9	100%	0	0%	0	0%
Coleoptera	Dermestidae	43	22	51%	0	0%	7	16%
Coleoptera	Derodontidae	1	1	100%	0	0%	0	0%
Coleoptera	Drilidae	1	1	100%	0	0%	1	100%
Coleoptera	Dryophthoridae	4	4	100%	0	0%	0	0%
Coleoptera	Dryopidae	9	1	11%	0	0%	0	0%
Coleoptera	Dynastidae	1	0	0%	0	0%	0	0%
Coleoptera	Dytiscidae	126	83	66%	0	0%	10	8%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Coleoptera	Elateridae	77	66	86%	0	0%	12	16%
Coleoptera	Elmidae	12	12	100%	0	0%	0	0%
Coleoptera	Endomychidae	8	7	88%	0	0%	0	0%
Coleoptera	Erirhinidae	14	14	100%	0	0%	0	0%
Coleoptera	Erotylidae	7	7	100%	0	0%	1	14%
Coleoptera	Eucinetidae	2	1	50%	0	0%	0	0%
Coleoptera	Eucnemidae	6	6	100%	0	0%	1	17%
Coleoptera	Geotrupidae	8	7	88%	0	0%	0	0%
Coleoptera	Gyrinidae	15	12	80%	0	0%	4	27%
Coleoptera	Haliplidae	19	2	11%	0	0%	2	11%
Coleoptera	Heteroceridae	9	8	89%	0	0%	0	0%
Coleoptera	Histeridae	56	41	73%	0	0%	2	4%
Coleoptera	Hydraenidae	33	17	52%	0	0%	1	3%
Coleoptera	Hydrophilidae	109	61	56%	0	0%	11	10%
Coleoptera	Hygrobiidae	1	1	100%	0	0%	0	0%
Coleoptera	Kateretidae	10	10	100%	0	0%	2	20%
Coleoptera	Laemophloeidae	11	6	55%	0	0%	1	9%
Coleoptera	Lampyridae	3	3	100%	0	0%	0	0%
Coleoptera	Latridiidae	56	28	50%	0	0%	1	2%
Coleoptera	Leiodidae	104	43	41%	0	0%	4	4%
Coleoptera	Limnichidae	1	1	100%	0	0%	0	0%
Coleoptera	Lucanidae	4	4	100%	0	0%	1	25%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Coleoptera	Lycidae	4	4	100%	0	0%	2	50%
Coleoptera	Lyctidae	6	6	100%	0	0%	0	0%
Coleoptera	Lymexylidae	2	2	100%	0	0%	0	0%
Coleoptera	Malachiidae	17	17	100%	0	0%	0	0%
Coleoptera	Melandryidae	18	18	100%	0	0%	4	22%
Coleoptera	Meloidae	11	10	91%	0	0%	0	0%
Coleoptera	Melolonthidae	9	7	78%	0	0%	0	0%
Coleoptera	Monotomidae	22	2	9%	0	0%	2	9%
Coleoptera	Mordellidae	20	8	40%	0	0%	5	25%
Coleoptera	Mycetophagidae	15	15	100%	0	0%	4	27%
Coleoptera	Mycteridae	1	1	100%	0	0%	0	0%
Coleoptera	Nanophyidae	2	2	100%	0	0%	0	0%
Coleoptera	Nemonychidae	1	1	100%	0	0%	0	0%
Coleoptera	Nitidulidae	90	40	44%	0	0%	9	10%
Coleoptera	Noteridae	2	2	100%	0	0%	0	0%
Coleoptera	Oedemeridae	10	6	60%	0	0%	5	50%
Coleoptera	Phalacridae	17	9	53%	0	0%	0	0%
Coleoptera	Phloiophilidae	1	1	100%	0	0%	0	0%
Coleoptera	Platypodidae	2	2	100%	0	0%	0	0%
Coleoptera	Psephenidae	1	1	100%	0	0%	0	0%
Coleoptera	Ptiliidae	79	27	34%	0	0%	1	1%
Coleoptera	Ptilodactylidae	1	1	100%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Coleoptera	Pyrochroidae	3	3	100%	0	0%	1	33%
Coleoptera	Pythidae	1	1	100%	0	0%	0	0%
Coleoptera	Raymondionymidae	1	1	100%	0	0%	0	0%
Coleoptera	Rhynchitidae	20	16	80%	0	0%	0	0%
Coleoptera	Ripiphoridae	1	1	100%	0	0%	0	0%
Coleoptera	Rutelidae	5	3	60%	0	0%	0	0%
Coleoptera	Salpingidae	12	12	100%	0	0%	1	8%
Coleoptera	Scarabaeidae	11	2	18%	0	0%	5	45%
Coleoptera	Scirtidae	22	22	100%	0	0%	0	0%
Coleoptera	Scraptiidae	15	4	27%	0	0%	5	33%
Coleoptera	Scydmaenidae	32	15	47%	0	0%	4	13%
Coleoptera	Silphidae	23	21	91%	0	0%	3	13%
Coleoptera	Silvanidae	12	6	50%	0	0%	1	8%
Coleoptera	Sphaeritidae	1	1	100%	0	0%	0	0%
Coleoptera	Sphaeriusidae	1	1	100%	0	0%	0	0%
Coleoptera	Sphindidae	2	2	100%	0	0%	2	100%
Coleoptera	Staphylinidae	1148	857	75%	0	0%	138	12%
Coleoptera	Tenebrionidae	49	48	98%	0	0%	3	6%
Coleoptera	Tetratomidae	4	4	100%	0	0%	0	0%
Coleoptera	Thanerocleridae	1	1	100%	0	0%	0	0%
Coleoptera	Throscidae	6	6	100%	0	0%	2	33%
Coleoptera	Trogidae	3	3	100%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Coleoptera	Trogossitidae	5	5	100%	0	0%	0	0%
Coleoptera	Zopheridae	10	10	100%	0	0%	0	0%
Diptera	Acartophthalmidae	2	2	100%	0	0%	1	50%
Diptera	Acroceridae	4	3	75%	0	0%	0	0%
Diptera	Agromyzidae	407	6	1%	0	0%	0	0%
Diptera	Anisopodidae	5	4	80%	0	0%	3	60%
Diptera	Anthomyiidae	252	43	17%	0	0%	16	6%
Diptera	Anthomyzidae	20	19	95%	0	0%	3	15%
Diptera	Asilidae	31	29	94%	0	0%	7	23%
Diptera	Asteiidae	7	7	100%	0	0%	1	14%
Diptera	Atelestidae	2	2	100%	0	0%	0	0%
Diptera	Athericidae	3	3	100%	0	0%	0	0%
Diptera	Aulacigastridae	1	1	100%	0	0%	0	0%
Diptera	Bibionidae	20	20	100%	0	0%	5	25%
Diptera	Bolitophilidae	16	6	38%	0	0%	4	25%
Diptera	Bombyliidae	12	11	92%	0	0%	3	25%
Diptera	Borboropsidae	1	1	100%	0	0%	0	0%
Diptera	Braulidae	2	0	0%	0	0%	0	0%
Diptera	Calliphoridae	39	39	100%	0	0%	24	62%
Diptera	Camillidae	5	0	0%	0	0%	0	0%
Diptera	Campichoetidae	2	2	100%	0	0%	0	0%
Diptera	Canacidae	11	11	100%	0	0%	0	0%

Order Name	**Family name**	**No of Species in Family in the UK (UKSI/NBN data)**	**No of Species in Pantheon in Family (Pantheon data)**	**Percentage of Family represented in Pantheon**	**Number of Species in Family represented in BOLD by species with at least one High Confidence record**	**Percentage of Family represented in BOLD by species with at least one High Confidence record**	**Number of Species in Family represented in BOLD by species with at least one Medium Confidence record**	**Percentage of Family represented in BOLD by species with at least one Medium Confidence record**
Diptera	Carnidae	13	6	46%	0	0%	0	0%
Diptera	Cecidomyiidae	667	7	1%	0	0%	1	0%
Diptera	Cecidomyiinae	1	1	100%	0	0%	0	0%
Diptera	Ceratopogonidae	187	1	1%	0	0%	0	0%
Diptera	Chamaemyiidae	32	32	100%	0	0%	2	6%
Diptera	Chaoboridae	6	6	100%	0	0%	2	33%
Diptera	Chironomidae	714	13	2%	0	0%	5	1%
Diptera	Chiropteromyzidae	1	1	100%	0	0%	0	0%
Diptera	Chloropidae	188	176	94%	0	0%	33	18%
Diptera	Chyromyidae	10	2	20%	0	0%	0	0%
Diptera	Clusiidae	10	10	100%	0	0%	3	30%
Diptera	Coelopidae	3	3	100%	0	0%	1	33%
Diptera	Conopidae	26	23	88%	0	0%	6	23%
Diptera	Cryptochetidae	1	0	0%	0	0%	0	0%
Diptera	Culicidae	37	34	92%	8	22%	23	62%
Diptera	Cylindrotomidae	4	4	100%	0	0%	2	50%
Diptera	Diadocidiidae	3	2	67%	0	0%	0	0%
Diptera	Diastatidae	6	6	100%	0	0%	3	50%
Diptera	Ditomyiidae	3	3	100%	0	0%	0	0%
Diptera	Dixidae	15	15	100%	0	0%	1	7%
Diptera	Dolichopodidae	308	297	96%	0	0%	29	9%
Diptera	Drosophilidae	68	21	31%	0	0%	6	9%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Diptera	Dryomyzidae	3	3	100%	0	0%	0	0%
Diptera	Empididae	224	218	97%	0	0%	49	22%
Diptera	Ephydridae	155	140	90%	0	0%	35	23%
Diptera	Fanniidae	60	60	100%	0	0%	31	52%
Diptera	Helcomyzidae	1	1	100%	0	0%	0	0%
Diptera	Heleomyzidae	59	57	97%	0	0%	27	46%
Diptera	Heterocheilidae	1	1	100%	0	0%	0	0%
Diptera	Hippoboscidae	15	1	7%	0	0%	0	0%
Diptera	Hybotidae	172	170	99%	0	0%	57	33%
Diptera	Keroplatidae	51	25	49%	0	0%	3	6%
Diptera	Lauxaniidae	62	58	94%	0	0%	13	21%
Diptera	Limoniidae	244	217	89%	0	0%	27	11%
Diptera	Lonchaeidae	46	45	98%	0	0%	9	20%
Diptera	Lonchopteridae	7	7	100%	0	0%	2	29%
Diptera	Megamerinidae	1	1	100%	0	0%	0	0%
Diptera	Micropezidae	10	10	100%	0	0%	2	20%
Diptera	Microphoridae	3	3	100%	0	0%	1	33%
Diptera	Milichiidae	19	3	16%	0	0%	0	0%
Diptera	Muscidae	290	261	90%	0	0%	105	36%
Diptera	Mycetobiidae	3	3	100%	0	0%	0	0%
Diptera	Mycetophilidae	494	226	46%	0	0%	62	13%
Diptera	Mythicomyiidae	1	0	0%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Diptera	Nycteribiidae	3	3	100%	0	0%	0	0%
Diptera	Odiniidae	9	9	100%	0	0%	2	22%
Diptera	Oestridae	13	11	85%	0	0%	0	0%
Diptera	Opetiidae	1	1	100%	0	0%	0	0%
Diptera	Opomyzidae	17	17	100%	0	0%	5	29%
Diptera	Pallopteridae	14	13	93%	0	0%	3	21%
Diptera	Pediciidae	20	19	95%	0	0%	0	0%
Diptera	Periscelididae	6	6	100%	0	0%	1	17%
Diptera	Phaeomyiidae	2	2	100%	0	0%	0	0%
Diptera	Phoridae	358	23	6%	0	0%	2	1%
Diptera	Piophilidae	15	15	100%	0	0%	7	47%
Diptera	Pipunculidae	95	91	96%	0	0%	25	26%
Diptera	Platypezidae	35	33	94%	0	0%	4	11%
Diptera	Platystomatidae	2	2	100%	0	0%	1	50%
Diptera	Pseudopomyzidae	1	1	100%	0	0%	1	100%
Diptera	Psilidae	29	26	90%	0	0%	0	0%
Diptera	Psychodidae	99	3	3%	0	0%	0	0%
Diptera	Ptychopteridae	7	7	100%	0	0%	1	14%
Diptera	Rhagionidae	16	15	94%	0	0%	4	25%
Diptera	Rhinophoridae	8	7	88%	0	0%	5	63%
Diptera	Sarcophagidae	64	62	97%	0	0%	23	36%
Diptera	Scathophagidae	55	53	96%	0	0%	14	25%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Diptera	Scatopsidae	45	0	0%	0	0%	0	0%
Diptera	Scenopinidae	4	3	75%	0	0%	0	0%
Diptera	Sciaridae	274	1	0%	0	0%	1	0%
Diptera	Sciomyzidae	70	68	97%	0	0%	18	26%
Diptera	Sepsidae	29	28	97%	0	0%	15	52%
Diptera	Simuliidae	32	0	0%	0	0%	0	0%
Diptera	Sphaeroceridae	139	10	7%	0	0%	5	4%
Diptera	Stratiomyidae	54	49	91%	0	0%	11	20%
Diptera	Strongylophthalmyiidae	1	1	100%	0	0%	0	0%
Diptera	Syrphidae	300	292	97%	3	1%	114	38%
Diptera	Tabanidae	37	31	84%	0	0%	5	14%
Diptera	Tachinidae	284	107	38%	0	0%	21	7%
Diptera	Tanypezidae	1	1	100%	0	0%	1	100%
Diptera	Tephritidae	87	82	94%	0	0%	19	22%
Diptera	Tethinidae	1	1	100%	0	0%	2	200%
Diptera	Thaumaleidae	3	3	100%	0	0%	1	33%
Diptera	Therevidae	14	14	100%	0	0%	3	21%
Diptera	Tipulidae	88	87	99%	0	0%	16	18%
Diptera	Trichoceridae	11	6	55%	0	0%	3	27%
Diptera	Trixoscelididae	4	4	100%	0	0%	0	0%
Diptera	Ulidiidae	20	20	100%	0	0%	2	10%
Diptera	Xylomyidae	3	3	100%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Diptera	Xylophagidae	3	3	100%	0	0%	2	67%
Hemiptera	Acanthosomatidae	5	5	100%	0	0%	0	0%
Hemiptera	Adelgidae	8	3	38%	0	0%	2	25%
Hemiptera	Aleyrodidae	13	3	23%	0	0%	0	0%
Hemiptera	Alydidae	3	1	33%	0	0%	0	0%
Hemiptera	Anthocoridae	38	34	89%	0	0%	0	0%
Hemiptera	Aphelocheiridae	1	1	100%	0	0%	0	0%
Hemiptera	Aphididae	614	73	12%	0	0%	17	3%
Hemiptera	Aphrophoridae	11	9	82%	0	0%	0	0%
Hemiptera	Aradidae	7	7	100%	0	0%	0	0%
Hemiptera	Asterolecaniidae	5	0	0%	0	0%	0	0%
Hemiptera	Berytidae	9	9	100%	0	0%	0	0%
Hemiptera	Calophyidae	1	0	0%	0	0%	0	0%
Hemiptera	Ceratocombidae	1	1	100%	0	0%	0	0%
Hemiptera	Cercopidae	1	1	100%	0	0%	0	0%
Hemiptera	Cicadellidae	307	303	99%	0	0%	8	3%
Hemiptera	Cicadidae	1	1	100%	0	0%	0	0%
Hemiptera	Cimicidae	4	4	100%	0	0%	0	0%
Hemiptera	Cixiidae	16	14	88%	0	0%	0	0%
Hemiptera	Coccidae	32	8	25%	0	0%	0	0%
Hemiptera	Coreidae	13	11	85%	0	0%	0	0%
Hemiptera	Corixidae	40	37	93%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Hemiptera	Cydnidae	12	9	75%	0	0%	0	0%
Hemiptera	Delphacidae	82	77	94%	0	0%	0	0%
Hemiptera	Diaspididae	11	5	45%	0	0%	0	0%
Hemiptera	Dipsocoridae	2	2	100%	0	0%	0	0%
Hemiptera	Eriococcidae	3	0	0%	0	0%	0	0%
Hemiptera	Gerridae	10	10	100%	0	0%	0	0%
Hemiptera	Hebridae	2	2	100%	0	0%	0	0%
Hemiptera	Homotomidae	1	1	100%	0	0%	0	0%
Hemiptera	Hydrometridae	2	2	100%	0	0%	0	0%
Hemiptera	Issidae	2	2	100%	0	0%	0	0%
Hemiptera	Leptopodidae	1	0	0%	0	0%	0	0%
Hemiptera	Lygaeidae	101	92	91%	0	0%	0	0%
Hemiptera	Margarodidae	1	1	100%	0	0%	0	0%
Hemiptera	Membracidae	2	2	100%	0	0%	0	0%
Hemiptera	Mesoveliidae	1	1	100%	0	0%	0	0%
Hemiptera	Microphysidae	7	7	100%	0	0%	0	0%
Hemiptera	Miridae	249	238	96%	0	0%	0	0%
Hemiptera	Nabidae	14	13	93%	0	0%	0	0%
Hemiptera	Naucoridae	2	2	100%	0	0%	0	0%
Hemiptera	Nepidae	2	2	100%	0	0%	0	0%
Hemiptera	Notonectidae	6	5	83%	0	0%	0	0%
Hemiptera	Ortheziidae	2	0	0%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Hemiptera	Pentatomidae	33	26	79%	0	0%	1	3%
Hemiptera	Phoenicococcidae	1	1	100%	0	0%	0	0%
Hemiptera	Phylloxeridae	7	1	14%	0	0%	0	0%
Hemiptera	Piesmatidae	4	2	50%	0	0%	0	0%
Hemiptera	Pleidae	1	1	100%	0	0%	0	0%
Hemiptera	Pseudococcidae	47	2	4%	0	0%	0	0%
Hemiptera	Psyllidae	72	1	1%	0	0%	0	0%
Hemiptera	Pyrrhocoridae	1	1	100%	0	0%	0	0%
Hemiptera	Reduviidae	10	9	90%	0	0%	0	0%
Hemiptera	Rhopalidae	14	11	79%	0	0%	0	0%
Hemiptera	Saldidae	23	23	100%	0	0%	0	0%
Hemiptera	Scutelleridae	5	5	100%	0	0%	0	0%
Hemiptera	Stenocephalidae	2	2	100%	0	0%	0	0%
Hemiptera	Tettigometridae	1	1	100%	0	0%	0	0%
Hemiptera	Thyreocoridae	1	1	100%	0	0%	0	0%
Hemiptera	Tingidae	25	25	100%	0	0%	2	8%
Hemiptera	Triozidae	19	2	11%	0	0%	0	0%
Hemiptera	Ulopidae	2	2	100%	0	0%	0	0%
Hemiptera	Veliidae	6	5	83%	0	0%	0	0%
Hymenoptera	Andrenidae	74	4	5%	41	55%	21	28%
Hymenoptera	Aphelinidae	43	0	0%	0	0%	0	0%
Hymenoptera	Apidae	106	36	34%	21	20%	12	11%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Hymenoptera	Argidae	22	1	5%	0	0%	1	5%
Hymenoptera	Aulacidae	1	0	0%	0	0%	0	0%
Hymenoptera	Azotidae	1	0	0%	0	0%	0	0%
Hymenoptera	Bethylidae	23	20	87%	0	0%	4	17%
Hymenoptera	Blasticotomidae	1	0	0%	0	0%	0	0%
Hymenoptera	Braconidae	1349	1	0%	0	0%	0	0%
Hymenoptera	Cephidae	13	0	0%	0	0%	0	0%
Hymenoptera	Ceraphronidae	77	0	0%	0	0%	0	0%
Hymenoptera	Chalcididae	10	0	0%	0	0%	0	0%
Hymenoptera	Chrysididae	40	33	83%	0	0%	9	23%
Hymenoptera	Cimbicidae	20	0	0%	0	0%	0	0%
Hymenoptera	Colletidae	22	1	5%	14	64%	12	55%
Hymenoptera	Crabronidae	131	81	62%	0	0%	4	3%
Hymenoptera	Cynipidae	199	2	1%	0	0%	0	0%
Hymenoptera	Diapriidae	300	0	0%	0	0%	0	0%
Hymenoptera	Diprionidae	9	0	0%	0	0%	0	0%
Hymenoptera	Dryinidae	37	35	95%	0	0%	9	24%
Hymenoptera	Embolemidae	1	1	100%	0	0%	0	0%
Hymenoptera	Encyrtidae	233	0	0%	0	0%	0	0%
Hymenoptera	Eucharitidae	1	0	0%	0	0%	0	0%
Hymenoptera	Eulophidae	516	0	0%	0	0%	0	0%
Hymenoptera	Eupelmidae	22	0	0%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Hymenoptera	Eurytomidae	101	0	0%	0	0%	0	0%
Hymenoptera	Evaniidae	2	0	0%	0	0%	0	0%
Hymenoptera	Figitidae	131	0	0%	0	0%	0	0%
Hymenoptera	Formicidae	83	72	87%	0	0%	25	30%
Hymenoptera	Gasteruptiidae	5	0	0%	0	0%	0	0%
Hymenoptera	Halictidae	66	20	30%	31	47%	20	30%
Hymenoptera	Heloridae	3	0	0%	0	0%	0	0%
Hymenoptera	Ibaliidae	2	0	0%	0	0%	0	0%
Hymenoptera	Ichneumonidae	2678	0	0%	0	0%	0	0%
Hymenoptera	Megachilidae	49	12	24%	13	27%	21	43%
Hymenoptera	Megalodontesidae	2	0	0%	0	0%	0	0%
Hymenoptera	Megaspilidae	16	0	0%	0	0%	0	0%
Hymenoptera	Melittidae	6	6	100%	4	67%	3	50%
Hymenoptera	Mutillidae	5	3	60%	0	0%	1	20%
Hymenoptera	Mymaridae	104	0	0%	0	0%	0	0%
Hymenoptera	Mymarommatidae	1	0	0%	0	0%	0	0%
Hymenoptera	Ormyridae	4	0	0%	0	0%	0	0%
Hymenoptera	Orussidae	1	0	0%	0	0%	0	0%
Hymenoptera	Pamphiliidae	22	0	0%	0	0%	0	0%
Hymenoptera	Perilampidae	10	0	0%	0	0%	0	0%
Hymenoptera	Platygastridae	258	0	0%	0	0%	0	0%
Hymenoptera	Pompilidae	46	14	30%	0	0%	14	30%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Hymenoptera	Proctotrupidae	43	0	0%	0	0%	0	0%
Hymenoptera	Pteromalidae	577	0	0%	0	0%	0	0%
Hymenoptera	Sapygidae	2	2	100%	0	0%	0	0%
Hymenoptera	Scelionidae	122	0	0%	0	0%	0	0%
Hymenoptera	Scoliidae	2	0	0%	0	0%	0	0%
Hymenoptera	Signiphoridae	2	0	0%	0	0%	0	0%
Hymenoptera	Siricidae	12	0	0%	0	0%	0	0%
Hymenoptera	Sphecidae	8	5	63%	0	0%	50	625%
Hymenoptera	Tenthredinidae	465	0	0%	0	0%	0	0%
Hymenoptera	Tetracampidae	8	0	0%	0	0%	0	0%
Hymenoptera	Tiphiidae	3	3	100%	0	0%	1	33%
Hymenoptera	Torymidae	114	0	0%	0	0%	0	0%
Hymenoptera	Trichogrammatidae	38	0	0%	0	0%	0	0%
Hymenoptera	Trigonalidae	1	0	0%	0	0%	0	0%
Hymenoptera	Vespidae	40	19	48%	0	0%	10	25%
Hymenoptera	Xiphydriidae	3	0	0%	0	0%	0	0%
Hymenoptera	Xyelidae	3	0	0%	0	0%	0	0%
Lepidoptera	Adelidae	15	15	100%	0	0%	5	33%
Lepidoptera	Alucitidae	1	1	100%	0	0%	1	100%
Lepidoptera	Arctiidae	9	9	100%	0	0%	7	78%
Lepidoptera	Argyresthiidae	28	24	86%	0	0%	13	46%
Lepidoptera	Autostichidae	5	4	80%	0	0%	0	0%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Lepidoptera	Batrachedridae	3	2	67%	0	0%	1	33%
Lepidoptera	Bedelliidae	1	1	100%	0	0%	1	100%
Lepidoptera	Blastobasidae	9	6	67%	0	0%	2	22%
Lepidoptera	Bucculatricidae	14	13	93%	0	0%	3	21%
Lepidoptera	Castniidae	1	0	0%	0	0%	0	0%
Lepidoptera	Chimabachidae	3	3	100%	0	0%	1	33%
Lepidoptera	Choreutidae	8	7	88%	0	0%	3	38%
Lepidoptera	Coleophoridae	110	109	99%	0	0%	20	18%
Lepidoptera	Cosmopterigidae	17	15	88%	0	0%	7	41%
Lepidoptera	Cossidae	3	3	100%	0	0%	1	33%
Lepidoptera	Crambidae	151	81	54%	0	0%	19	13%
Lepidoptera	Depressariidae	53	49	92%	0	0%	12	23%
Lepidoptera	Douglasiidae	2	2	100%	0	0%	0	0%
Lepidoptera	Drepanidae	28	16	57%	0	0%	0	0%
Lepidoptera	Elachistidae	50	48	96%	0	0%	13	26%
Lepidoptera	Endromidae	1	1	100%	0	0%	0	0%
Lepidoptera	Epermeniidae	9	8	89%	0	0%	3	33%
Lepidoptera	Erebidae	111	93	84%	0	0%	0	0%
Lepidoptera	Eriocraniidae	8	8	100%	0	0%	5	63%
Lepidoptera	Ethmiidae	6	6	100%	0	0%	0	0%
Lepidoptera	Gelechiidae	172	162	94%	0	0%	39	23%
Lepidoptera	Geometridae	465	318	68%	0	0%	111	24%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Lepidoptera	Glyphipterigidae	14	14	100%	0	0%	3	21%
Lepidoptera	Gracillariidae	104	94	90%	0	0%	36	35%
Lepidoptera	Heliodinidae	1	1	100%	0	0%	0	0%
Lepidoptera	Heliozelidae	5	5	100%	1	20%	1	20%
Lepidoptera	Hepialidae	9	5	56%	0	0%	2	22%
Lepidoptera	Hesperiidae	18	10	56%	0	0%	6	33%
Lepidoptera	Incurvariidae	5	5	100%	0	0%	5	100%
Lepidoptera	Lasiocampidae	19	12	63%	0	0%	0	0%
Lepidoptera	Limacodidae	2	2	100%	0	0%	0	0%
Lepidoptera	Lycaenidae	68	22	32%	0	0%	8	12%
Lepidoptera	Lyonetiidae	10	9	90%	0	0%	4	40%
Lepidoptera	Lypusidae	4	4	100%	0	0%	0	0%
Lepidoptera	Micropterigidae	5	5	100%	0	0%	0	0%
Lepidoptera	Momphidae	15	14	93%	0	0%	7	47%
Lepidoptera	Nepticulidae	107	100	93%	1	1%	36	34%
Lepidoptera	Noctuidae	500	391	78%	0	0%	117	23%
Lepidoptera	Nolidae	17	15	88%	0	0%	0	0%
Lepidoptera	Notodontidae	32	29	91%	0	0%	2	6%
Lepidoptera	Nymphalidae	129	33	26%	0	0%	12	9%
Lepidoptera	Oecophoridae	29	28	97%	0	0%	17	59%
Lepidoptera	Opostegidae	4	4	100%	0	0%	0	0%
Lepidoptera	Papilionidae	14	3	21%	0	0%	2	14%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Lepidoptera	Parametriotidae	6	6	100%	0	0%	1	17%
Lepidoptera	Peleopodidae	1	1	100%	0	0%	1	100%
Lepidoptera	Pieridae	43	15	35%	0	0%	8	19%
Lepidoptera	Plutellidae	7	7	100%	0	0%	0	0%
Lepidoptera	Praydidae	7	5	71%	0	0%	0	0%
Lepidoptera	Prodoxidae	7	7	100%	0	0%	0	0%
Lepidoptera	Psychidae	21	19	90%	0	0%	4	19%
Lepidoptera	Pterophoridae	47	44	94%	0	0%	6	13%
Lepidoptera	Pyralidae	99	65	66%	0	0%	16	16%
Lepidoptera	Riodinidae	1	1	100%	0	0%	0	0%
Lepidoptera	Roeslerstammiidae	2	2	100%	0	0%	0	0%
Lepidoptera	Saturniidae	3	2	67%	0	0%	0	0%
Lepidoptera	Schreckensteiniidae	1	1	100%	0	0%	1	100%
Lepidoptera	Scythrididae	12	11	92%	0	0%	0	0%
Lepidoptera	Scythropiidae	1	1	100%	0	0%	0	0%
Lepidoptera	Sesiidae	16	15	94%	0	0%	5	31%
Lepidoptera	Sphingidae	27	26	96%	0	0%	17	63%
Lepidoptera	Stathmopodidae	3	2	67%	0	0%	0	0%
Lepidoptera	Tineidae	65	50	77%	0	0%	14	22%
Lepidoptera	Tischeriidae	6	6	100%	0	0%	3	50%
Lepidoptera	Tortricidae	415	384	93%	1	0%	101	24%
Lepidoptera	Yponomeutidae	24	24	100%	0	0%	17	71%

Order Name	Family name	No of Species in Family in the UK (UKSI/NBN data)	No of Species in Pantheon in Family (Pantheon data)	Percentage of Family represented in Pantheon	Number of Species in Family represented in BOLD by species with at least one High Confidence record	Percentage of Family represented in BOLD by species with at least one High Confidence record	Number of Species in Family represented in BOLD by species with at least one Medium Confidence record	Percentage of Family represented in BOLD by species with at least one Medium Confidence record
Lepidoptera	Ypsolophidae	16	16	100%	0	0%	0	0%
Lepidoptera	Zygaenidae	37	11	30%	0	0%	1	3%

Contact us: Buglife, The Lindens, 86 Lincoln Road, Peterborough, PE1 2SN

www.buglife.org.uk

Tel: 01733 201210

Email: info@buglife.org.uk

@buzz_dont_tweet

Photo credits L-R; Ladybird spider *(Eresus sandaliatus)* © S. Dalton, Jellyfish © D. Huffman, Tansy beetle (*Chrysolina graminis*) © S. Falk and Large garden bumblebee (*Bombus ruderatus*) © S. Falk

Saving the small things that run the planet

Buglife - The Invertebrate Conservation Trust is a registered charity at
The Lindens, 86 Lincoln Road, Peterborough, PE1 2SN
Company no. 4132695, Registered charity no. 1092293, Scottish charity no. SC040004