We present The Odonate Phenotypic Database (OPD): an online data resource of dragonfly and damselfly phenotypes (Insecta: Odonata). The database consists of multiple morphological, life-history and behavioral traits, and biogeographical information collected from literature sources. We see taxon-specific phenotypic databases from Odonata and other organismal groups as becoming an increasing valuable resource in comparative studies. Our database has phenotypic records for 1011 of all 6400 known odonate species. The database is accessible at http://www.odonatephenotypicdatabase.org/, and a static version with an information file about the variables in the database is archived at Dryad.

Background & Summary
The Odonate Phenotypic Database is an online data resource for dragonfly and damselfly phenotypes (Insecta: Odonata). The database consists of a variety of morphological, life-history and behavioral traits, and biogeographical information collected from various sources in the literature. The database is not intended for species identification, but for comparative analysis of insect groups or to be combined with data from other taxonomic groups. The database is provided along with a large phylogenetic tree (1322 taxa, 21% of known odonates https://www.pugetsound.edu/academics/academic-resources/slater-museum/biodiversity-resources/dragonflies/world-odonata-list2/, accessed in November 2015): “The Odonate Super Tree”. This phylogenetic tree was constructed using DNA-sequences from GenBank in combination with a traditional (morphologically-based) odonate taxonomy as our backbone.

Comparative analyses are becoming an increasing common part of evolutionary studies, as researchers attempt to bridge the gap between microevolutionary processes and macroevolutionary patterns. Most comparative analyses require high-quality phenotypic data collected from the literature, and often a large amount of time is spent collecting such data. Trait information can be obtained from measurements of live and field-caught individuals, museum specimens or literature sources but often important covariates are missing, such as behavioural information or habitat data. It is therefore in the interest of many of us working in the field to collect and curate such phenotypic data in a coherent fashion so that such data can be used in future studies and combined with multiple other sources of phenotypic information, particularly in light of the explosion in phylogenetic comparative methods the last decades.

Paradoxically, as a community we have been much more successful at storing and cataloguing genotypes, and DNA-sequences often exist in GenBank for many species, but not even basic phenotypic data (such as body size) exist in an easily accessible form for many organismal groups. This lack of information is most likely due to there being no clear structure or well-established routines agreed upon in how to store phenotypic data and which aspects of phenotypes should be measured. Phenotypic databases, because of the high-dimensional nature of...
most phenotypes, are also very different from a genetic database such as GenBank. Recent calls for a new field of “phenomics” – i.e. obtaining high-throughput phenotypic data in a similar fashion as in genomics – will always have to prioritize what aspects of the phenotypes to measure. There are many challenges in developing such a general and cross-taxonomic research programme in phenomics, in particular the difficulties of establishing general and operational phenotype ontologies between distantly related organisms across the Tree of Life.

From an evolutionary viewpoint, phenotypes are arguably just as important and interesting as genotypes, if not more so, as selection operates on phenotypes, regardless of their genetic basis. Increasingly integrative research practices in evolutionary biology will need not only access to high-quality genomic, molecular and phylogenetic resources, but will also need high-quality phenotypic and biogeographic data, fossil information for time-calibration of phylogenetic trees and other general data provided by biodiversity informatics. The difficulty of these tasks and the size of each individual project should ideally not prevent the establishment of accessible structures needed to store the data.

One way forward is to create taxon-specific phenotypic databases, as we have done here. Having such databases available that focus on a certain taxonomic group, also allows the recorded phenotypes to be tailored to fit the needs of the specific group and have the advantages that trait definitions are less ambiguous. Examples of such open databases with various forms of phenotypic, biogeographic and phylogenetic data include AmphiBIO for amphibian ecological traits, panTHERIA and EltonTraits for birds and mammals, a global invasion atlas of birds, a database on eukaryotic sex determination systems, a recent database on thermal developmental plasticity of reptiles and a global database on plants. However, in the case of animals, such databases are largely focused on vertebrate groups, while the most speciose animal group – the insects – have few such open databases available. One exception is the Freshwater Biological Traits Database which contains trait data for North American macroinvertebrate taxa and includes habitat, life history, mobility, morphology, and ecological trait data, although not in all cases down to the species level.

![Fig. 1: Phenotypic and taxonomic diversity of Odonata.](https://www.nature.com/scientificdata)
We see taxon-specific phenotypic databases as becoming an increasingly valuable resource in comparative studies. Our research background and expertise is in the insect order Odonata (dragonflies and damselflies) (Fig. 1). To this end, we have collected data on 35 phenotypic, ecological and biogeographical variables that we see as useful to the research community (Online-only Table 1). Odonata is a relatively small insect order that currently consists of about 6400 species belonging to 32 families\(^1\). Odonata are morphologically highly conserved with respect to their overall external morphology (all species in this insect order have six legs and four wings), but they show considerable diversity in terms of wing and body colouration and shape (Fig. 1). Our database has at least some phenotypic records for 1011 of all 6400 known odonate species\(^1\) and a total of 3978 records (i.e., multiple records for many species, from different literature sources). The database is accessible at http://www.odonatephenotypicdatabase.org/.

**Methods**

We have collected phenotypic data and data on habitats, climatic classifications and more coarse-grained biogeographic categories (e.g. ecozones\(^2\)), from the scientific literature and from odonate field guides (Online-only Table 1). The field guides from which we obtained our phenotypic data are listed in Online-only Table 2. Phenotypes were scored by following a specific set of instructions for each variable that are described in an accompanying file that is uploaded alongside this paper as Supporting Material (“VariableDefinitions.pdf”) which is available on the Dryad Digital Repository\(^3\). The descriptions of each variable can be found in the Data Records section below and in Online-only Table 1. The construction of the Odonate Phylogenetic Super Tree has been described in detail elsewhere\(^4\).

Developing general and meaningful phenotype ontologies that are generally applicable across many taxa is a challenging task that is beyond the scope of this paper. The phenotypic variables (size, behaviours, wing and body colouration, colour patterns; Online-only Table 1) that we have assembled from the literature are not always easily and straightforward to translate to other, more well-studied insect groups, including the classical model organism *Drosophila melanogaster*, although our classifications are largely consistent with the recommendations given in the Ontobee database (http://www.ontobee.org/) and the Phenotype and Trait Ontology (PATO) database (http://www.ontobee.org/ontology/PATO)(Online-only Table 1). Clearly a lot of work is needed before we can develop generally applicable phenotypic ontologies across the entire class Insecta, let alone across more distantly related organismal groups\(^5,6\). All of our size measurements are given in millimetres (mm), although we acknowledge the fact that even a simple variable like wing length can be measured in different ways in different studies.

With respect to the ecological and geographic variables, our database is more easily connected and comparable with similar initiatives from other organismal groups. The ecozone variable for the distribution data of the different species follows the recent updates and classifications of Wallace's classical zoogeographic regions by Holt et al.\(^2\). Our habitat classifications of water bodies (Online-only Table 1) are largely consistent with the
environmental ontologies given in the Environment Ontology database (ENVO: http://www.ontobee.org/ontology/ENVO). We define lakes, ponds, rivers and streams in the same way as ENVO, whereas our classification of Ephemeral wetlands encompass both ephemeral springs (http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO_00000204) and ephemeral rivers (http://www.ontobee.org/ontology/ENVO?iri=http://purl.obolibrary.org/obo/ENVO_01000979).

Data Records
The database has been deposited on the Dryad Digital Repository26, and additional material, including the code that is also uploaded on Github (see above) is also deposited with the most recent and updated database version at http://www.odonatephenotypicdatabase.org/. This data includes a PDF-file that describes each of the variables within the database and how they were collected (“VariableDefinitions.pdf”)26.

We note that our database contains information from only 1011 of all 6400 species, i.e. about 16% taxonomic coverage, and coverage varies among traits (Fig. 2). For instance, behavioural traits like male guarding of females and territoriality have considerably lower coverage than morphological traits like size and other phenotypic measurements (Online-only Table 1; Fig. 2). The reason for this is that morphological traits can easily be obtained also from dead specimens in museums, whereas behavioural traits would typically require time-consuming field studies of these insects28–29, particularly in the tropics where species diversity is high but where we still lack basic faunistic information even about which species occur in many countries30. It is particularly noteworthy that much data overall is lacking from several tropical damselfly families (Calopterygidae, Chlorocyphidae, Euphalidae, Megapodagrionidae, Polythoridae and Platyctenidae) and some dragonfly families, particularly Gomphidae (Fig. 2). Our dataset illustrated in Fig. 2 should hopefully serve as a basis for targeted studies on these families.

Finally, we also note that information about the occurrence of female colour polymorphisms, which is an important feature of many odonate species and which have important roles in frequency-dependent sexual conflict31–37, is missing for many species and genera. Documenting the presence or absence of female colour polymorphisms typically requires large sample sizes, particularly if one wishes to estimate the frequencies of certain morphs, such as male-mimicking females, and such information is typically available for mainly temperate species, such as those in the genera Ischnura and Enallagma35,38. There is clearly a need for more quantitative field studies and surveys to improve this situation and fill the missing gaps in our knowledge, particularly for tropical taxa.

Technical Validation
All of the phenotypic data were collected from published field guides or reliable internet sources. The field guides are listed in Online-only Table 2. All the field guides have been published by respected odonatologists and experts on species identification. Our database is therefore not static, and additional data will be added as it becomes available. Researchers interested in contributing to this project are encouraged to contact the author for correspondence on how to incorporate new data. We will accept data both from already published sources (e.g. scientific papers) even if it has already been deposited in other databases such as Dryad, as well as data that is not intended to be published elsewhere, as long as it can be tailored to the format of the Odonate Phenotypic Database. Each species has a reference list, which lists the references from where the data were gathered, so it is possible to check each entry against these primary sources.

Usage Notes
The database is intended to be used in future comparative analyses of odonate trait evolution. We therefore provide a previously published phylogenetic tree (Fig. 2) along with these phenotypic data. Previous questions that we have addressed in recent past using phylogenetic comparative methods and these and similar data include the relationship between latitude and wing pigmentation19 and the micro- and macroevolutionary dynamics of body size1. We have also investigated diversification dynamics (speciation and extinction rates) in relation to body size and wing pigmentation19. Body size is also a trait that is of interest for conservation biology, as extinction risk (as defined by IUCN redlist criteria) was recently demonstrated to be significantly related to body size in damselflies40. In addition, behavioural diversity (“ethodiversity”) is a neglected aspect of biodiversity and conservation biology29 and in the future, we hope to incorporate such information in this database.

Other interesting questions to pursue in even more depth in the future include the evolutionary and ecological dynamics of sexual conflict and its consequences for the maintenance of sex-limited colour polymorphism in females31–36, and how climatic niche conservatism31,42 can shape the evolution, ecology and biogeography of odonate range limits35–36.

Code availability
All code used to generate the website and Odonate Super Tree are available on Github (https://github.com/jhnwlr/shiny-server/tree/master/odonates).

Received: 6 June 2019; Accepted: 12 November 2019;
Published online: 12 December 2019

References
1. Waller, J. T. & Svensson, E. I. Body size evolution in an old insect order: No evidence for Cope’s Rule in spite of fitness benefits of large size. Evolution 71, 2178–2193 (2017).
2. Arnold, S. J., Pfrender, M. E. & Jones, A. G. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112–113, 9–32 (2001).
3. Arnold, S. J. Phenotypic evolution: the ongoing synthesis. Am. Nat. 183, 729–746 (2014).
Acknowledgements
We would like to acknowledge the students and laboratory assistants that helped us to collect the data from the literature, and in particular Anna Kell, Ev Poslin, Hanna Bensch, Robin Pranter, Kajsa Svensson, Lisa Winberg, Karolina Pehrson, Mireia Balesta and Tammy Ho. We would like to also thank the many authors of the field guides from which our database references. Our database is by no means nor intends to be a substitute for these valuable books. Funding for this study have been provided by research grants from The Swedish Research Council (VR: grant no. 2016-03356), Gyllenstiernska Krapperupstiftelsen (grant no. KR2018-0038) and Olle Engqvist Byggmästares Stiftelse to E.I.S and from a Faculty Mobility grant from the University of Costa Rica and a grant from the Schlumberger Foundation to B.W. Open access funding provided by Lund University.

Author contributions
John Waller: constructed the website as part of his PhD-work, collected data, helped create variable definitions, wrote the first draft of the manuscript. Beatriz Willink: collected data, helped plan and organize data collection. Maximilian Tschlor: collected wing pigment data, created variable definitions. Erika I. Svensson: created variable definitions, supervised data collection, came up with the original idea to create the database and secured funding for the project.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to E.I.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

The Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files associated with this article.

© The Author(s) 2019