sDiv working group meeting summary
"sDevTrait - streamlining development efforts in tools for ecological trait analyses"

PIs: Alexander Keller; Markus Ankenbrand

The goal of the sDevTrait working group is to unlock the potential of publicly available species traits. We want to particularly address the technical difficulties that hinder findability, reuse and integration of traits that have already been gathered and shared from different sources. Our group consists of scientists, software developers and data managers working across the trait data lifecycle, from trait collection over curation to database development and answering biological questions in syntheses. In a first, fully online workshop in late 2020 we discussed this trait lifecycle and the most promising approaches for improving the state of interoperability and the possibilities for reuse of trait databases. We also identified potential questions for a synthesis that we want to address together.

Many members of this working group are also part of the Open Traits Network which was founded as a community effort to make trait data more openly available in 2018 (Gallagher, R.V., Falster, D.S., Maitner, B.S. *et al.* Open Science principles for accelerating trait-based science across the Tree of Life. *Nat Ecol Evol* **4,** 294–303 (2020). https://doi.org/10.1038/s41559-020-1109-6).

This meeting was held in a hybrid format with six of us participating (at least partially) live at the iDiv and all other participants joining in remotely. During this workshop we focused on two main tasks:

1. Transforming the informal trait data lifecycle document drafted in the first online meeting into journal article format. The document has been shared online at https://opentraits.org/lifecycle in early 2021 and received comments and contributions from the Open Traits community. After further discussion, we found the "10 simple rules" series in Plos Computational Biology to be suitable in style https://collections.plos.org/collection/ten-simple-rules/. Therefore, we re-formulated our ideas as (sometimes not-so-simple) rules. Thus the document evolved over multiple iterations of focused work in small break-out groups alternating with joined group discussions. The result of our effort is a manuscript draft that needs some fine-tuning and consolidation before it can be submitted. Submission of the manuscript to Plos Computational Biology is planned in early 2022.

2. We started working on a joint synthesis project on one of the most fundamental traits - body size - across the tree of life. In order to do this, we first need to identify all the various ways how size is measured, e.g. body mass, body length, body width… Size, and the distribution of size across members of an ecological community, is an indicator of the kinds of ecological assembly mechanisms taking place within a community. It is also possible that size is, itself, a driver of the evolution of species, and may even limit the total number of species on Earth. Our goal is to use the resources and networks of the participants to collect as much size data as possible across the tree of life. Additionally, we will post a public call for contributions and distribute it through related initiatives. The synthesized dataset will be available for everyone to use. We will use it to identify taxonomic and geographic gaps in our knowledge about size, and make use of those data to answer research questions. By following principles of fair attribution and recognition we will make sure everyone has the opportunity to contribute and receive credit for their hard work collecting data. In particular, we are excited to try and use these data to understand the drivers of biodiversity across our planet. Armed with knowledge about the size of species, we think we will be better-placed to understand the limits to biodiversity across the planet.

In addition to our work on these two main topics, Jennifer Hammock presented her thoughts on the uses of trait data found in old sources, e.g. books in museums. Her talk was titled "What, if anything, are the uses for old trait data?" and was followed by interesting discussions. Furthermore, Florian Schneider informed us about the state of the "*traitdataform*" R package. The whole group discussed possibilities of long term sustainability of the package, despite the current lack of funding.

The unfamiliar hybrid style of the meeting posed some challenges to get everyone involved equally. Exchange between participants beyond the scientific topics of the meeting was easier for the live participants. Still the feedback from the remote participants confirmed our feeling, that they also felt well included and were able to contribute efficiently. Overall, the meeting was very fruitful. Multiple follow-up online meetings before the next annual meeting in Leipzig are planned and some have already taken place.

We thank iDiv staff for making this a nicely organized and safe experience, despite the challenges caused by the pandemic.