

OVERVIEW RESULTS: Harmonized invertebrate grouping feature databases

Introduction

Explaining and predicting how aquatic communities are shaped by environmental factors is a main goal of freshwater ecology. Organismal traits, defined as measurable properties of an organism ([1]), are increasingly incorporated into freshwater ecology to support this goal, e.g. by relating macroinvertebrate trait composition to environmental factors or as trait metrics in biomonitoring ([2–5]). Traits evolve through adaptations (e.g., physiological, behavioral) of organisms to their environment and indicate direct or indirect linkages between biological response of an organism to its environment ([6, 7]). Besides providing a mechanistic explanation of species-environment relationships, trait-based approaches may be suitable for large scale analysis since variability in trait responses is less than for taxonomic responses ([8]).

In the last decades, freshwater ecologists developed comprehensive invertebrate trait databases for various biogeographic regions ([9–14]). The availability of invertebrate trait data from different biogeographic regions enables comparisons of trait variation and their relation to environmental factors across large scales. However, such analysis have been carried out mostly within one biogeographic region, using information from one or two trait databases. For example, [8] compared trait composition for mediterranean and temperate regions in Europe using traits from [9], [2] characterized trait composition across sites in the Western US using traits from [15], and [16] investigated the effect of salinity on invertebrate traits in different sites in South Australia using trait data from [15] and [17]. Rarely have analysis on invertebrate traits been carried out that synthesize information on traits and grouping features from more than two different biogeographic regions. A grouping feature is a general property (e.g. feeding mode) that comprises a "group of related traits (e.g., predator, shredder, etc.) that vary among species or among individuals within a species" ([18]). To our knowledge, only [19] harmonized grouping features from more than two geographically distant invertebrate trait databases in a study on the influence of decreasing glacier cover in nine biogeographic regions on functional diversity and community assembly of invertebrates.

We suspect that the heterogeneity of descriptions of the invertebrate trait databases is likely to be one of the reasons why information of these databases is rarely combined. In order to harmonize trait data from different regions, first commonly accepted and unambiguous trait definitions are required ([20]). In the best case, grouping features would be classified into the same traits across databases or they could easily be harmonized using standardized terminology. However, a lack of standardized terminology of trait definitions and poor metadata quality in many trait databases is a common issue throughout the field of trait-based ecology ([20, 21]). Secondly, consistent coding of traits facilitates the compatibility of trait data from different databases. Traits can be binary (i.e. trait exists or not), continuous, or fuzzy coded variables. Fuzzy codes represent the affinity of an

organism to express a certain trait. They are used to account for plasticity in traits and are usually converted to percentages. Continuous values are typically used for grouping features that can be measured, like body size. However, invertebrate trait databases are heterogeneous with regard to the coding they use for their traits ([22], see also Table 5). [19] harmonized grouping features based on trait databases from Europe, North America, and New Zealand since in these trait databases identical grouping features are classified differently into traits depending on the database. Also, because traits from North America were differently coded (binary) than those from Europe and New Zealand (fuzzy coded), the authors consulted experts to assign traits from North America or inferred them from the European database. Thus, it becomes apparent that using invertebrate trait data from several regions requires extensive data processing prior to the actual data analysis.

Differing taxonomic resolutions between the observed taxa in a study and the used trait database is another challenge when working with trait data. When observations are on a lower taxonomic level than data available in the trait databases (e.g. observations on species-level, trait data on genus-level) trait data of the higher taxonomic level are often assigned (e.g. [3]). Conversely, if trait information is only available on lower taxonomic levels than the observed taxa, traits are aggregated to a higher taxonomic level (e.g. [3, 15, 23]). Thereby, trait aggregation is often done by using the median or the mode. Up to now, studies on how and to which extent different trait aggregation methods compare to assessments by trait experts are missing.

In this study we (1) present four harmonized invertebrate grouping feature databases for seven grouping features based on information from trait databases of the regions Europe, North America, New Zealand and Australia. Furthermore, we (2) compare to family-level aggregated trait values obtained by different trait aggregation methods to trait values assigned at family-level by experts. We (3) analyze how the usage of harmonized and/or aggregated trait data alters the results compared to non harmonized and non aggregated trait data by re-analyzing data on the effect of anthropogenic salinization on biological traits by [3]. Finally, we (4) present an overview of discrepancies in trait definition between the used invertebrate trait databases and discuss challenges of trait data synthesis.

Methods

Selection of traits and harmonization of trait databases

We established four harmonized grouping feature databases using information available in invertebrate trait databases from Europe, North America, Australia, and New Zealand. Trait information for Europe was obtained by the Freshwaterecology database ([10]) and complemented by Tachet ([9]) when information was missing (e.g. for the grouping feature size). Trait information for North America was obtained from Laura Twardochleb (*in prep.*) and complemented by Vieira et al. 2006 ([11]). Philippe Usseglio-Polatera provided information on body form for European and North American taxa. For Australia and New Zealand, we used trait databases from Kefford et al. ([13]) and Philips and Smith respectively ([12]).

We selected traits of seven grouping features based on the availability of trait information across databases and so that they describe different parts of the biology of an organism: life history (Voltinism), morphology (Respiration, Body form, Size), ecology (Locomotion, Feeding mode) and reproduction (Oviposition). We did not include ecological traits that describe habitat preferences (e.g. temperature preference) as these traits are not reported in the New Zealand trait database. Since the grouping features were differently classified across the databases we harmonized them

into 26 traits (Table 1). Harmonization was undertaken by amalgamating similar traits into one trait (e.g. crawlers and sprawlers into crawlers). Thereby, the highest trait affinity was assigned to the harmonized trait per taxa.

We used fuzzy coded traits for establishing our harmonized databases unless data quality prohibited and then we used binary traits, i.e. categorical and continuous traits were converted into binary traits. Fuzzy codes are reported with different ranges in the trait databases (e.g. freshwater ecology 0 to 10, Tachet 0 to 3 or 0 to 5). We standardized them to a range between 0 and 1 and converted trait affinities to percentages.

Prior harmonization we amalgamated duplicate taxa on species, genus or family-level if present. We omitted taxa with a lower taxonomic resolution than family-level.

Table 1: Traits of the harmonized grouping features. The last column indicates traits that were amalgamated (no amalgamation needed if empty).

Grouping feature	Trait	Amalgamated traits
Voltinism	Semivoltine	<1 generation per year
	Univoltine	1 generation per year
	Bi/multivoltine	>1 generation per year
Body Form	Cylindrical	Round (humped)
	Flattenend	Dorsoventrally flattened
	Spherical	Tubular
	Streamlined	Fusiform
Size	Small	<9 mm (Tachet: <10 mm)
	Medium	9 - 16 mm (Tachet: 10 - 20 mm)
	Large	>16 mm (Tachet: >20 mm)
Respiration	Gills	Temporary air store, tracheal gills
	Plastron/Spiracle	Spiracular gills, atmospheric breathers, plant breathers, functional spiracles, air (plants), aerial
	Tegument	Cutaneous
Locomotion	Burrower	Interstitial, boring
	Crawler	Sprawler, walking, climber, clinger
	Sessil	Attached
	Swimmer	Skating, diving, planctonic
Feeding mode	Filterer	Active/passive filterer, absorber, filter-feeder, collector-filterer
	Gatherer	Deposit-feeder, collector-gatherer, detrivore
	Herbivore	Grazer, scraper, piercer herbivore, algal piercer
	Parasite	
	Predator	Piercer
	Shredder	Miner, xylophagus, shredder detrivore
Oviposition	Aquatic eggs	Eggs attached to substrate/plants/stones, free/fixed eggs/clutches
	Ovoviviparity	
	Terrestrial eggs	Terrestrial clutches

Trait aggregation

Traits of the harmonized grouping feature databases were aggregated to family-level using three approaches. I) we directly aggregated taxa to family-level giving equal weight to every species. We denote this aggregation as *direct_agg*. For the *direct_agg* we tested aggregating with the median and the mean. We added *median* or *mean* to *direct_agg* to indicate when we used which method. II)

taxa were aggregated stepwise, i.e first to the genus-level and subsequently to the family-level. By using this approach, we give equal weights to each genus. Hereafter, we abbreviate this aggregation type as *stepwise_agg*. We tested the *stepwise_agg* using the mean and the median, using the same naming as for the *direct_agg*. III) taxa were aggregated using a weighted mean approach, denoted as *weighted_agg*. This method weights the genera according to the number of their species present in the databases. As mentioned above, trait affinities ranged between 0 and 1. Hence, the maximum differences possible is 1 or -1 (corresponds to 100 %). For convenience, we report absolute trait differences.

Data analysis with harmonized grouping features and aggregated traits

To investigate how harmonizing grouping features and aggregating invertebrate traits might change the results in the analysis of trait-environment relationships we replicated the data analysis of Szöcs et al. 2014 ([3]) using the harmonized grouping features *Body size*, *Feeding mode*, *Locomotion*, *Reproduction/Oviposition*, *Respiration* and *Voltinism* (21 grouping features have been used in total) and additionally aggregated traits using the aforementioned aggregation methods. The harmonized grouping features used are those that responded strongly to salinity in the study of Szöcs et. al. 2014, except for life cycle duration. For testing the effect of aggregated traits we assigned to each taxon in [3] the aggregated trait value from the established harmonized European grouping feature database for its corresponding family. We limit our analysis to the RDA of traits constrained by electric conductivity of the original publication (see appendix for a more in-depth comparison to the original results). An overview of the harmonization for the European trait databases can be found in the supporting information in section .

Results harmonized grouping feature databases

We used information on invertebrate traits for the regions Europe, North America, Australia, and New Zealand to establish harmonized grouping feature databases. Trait information for Europe was obtained by the Freshwaterecology database ([10]) and complemented by Tachet ([9]) when information was missing (e.g. for the grouping feature size). Trait information for North America was obtained from Laura Twardochleb and complemented by Vieira et al. 2006 ([11]). Philippe Usseglio-Polatera provided information on body form for European and North American taxa. For Australia and New Zealand, we used trait databases from Kefford et al. ([13]) and Philips and Smith respectively ([12]).

We selected seven grouping features that we harmonized into 26 traits. The grouping features were chosen based on the availability of trait information across databases and so that they describe different parts of the biology of an organism: life history (Voltinism), morphology (Respiration, Body form, Size), ecology (Locomotion, Feeding mode) and reproduction (Oviposition). We used fuzzy coded traits for establishing our harmonized databases unless data quality prohibited and then we used binary traits, i.e. categorical and continuous traits were converted into binary traits. Data processing prior harmonization also involved amalgamation of duplicate taxa on species, genus or family-level and conversion of trait affinities to percentages per trait. We omitted taxa with a lower taxonomic resolution than family-level. The following results relate to the harmonized grouping feature databases.

1 Taxonomic coverage of the trait databases

Regarding the taxonomical coverage, the New Zealand database has, as expected, the smallest taxon pool (478 taxa, Table 2). By contrast, the largest taxon pool is spanned by the European trait database with 4110 taxa followed by the North American trait database that contained trait information on 3753 taxa. The Australian database contains 1402 taxa. The European, New Zealand, and North American databases have most taxa on the highest taxonomical resolution while the Australian database has a similar number of taxa on species and genus-level.

Table 2: Number of taxa per harmonized database and per taxonomic level. Numbers in parenthesis show relative frequencies in percentage.

Database	Nr. of taxa	Species	Genus	Family	Nr. aquatic taxa
Europe	4110	3848 (93.63)	237 (5.77)	25 (0.61)	3579 (87.08)
North America	3753	2414 (64.32)	1163 (30.99)	176 (4.69)	3305 (88.06)
Australia	1402	564 (40.23)	578 (41.23)	260 (18.54)	1016 (72.47)
New Zealand	478	404 (84.52)	47 (9.83)	27 (5.65)	443 (92.68)

2 Completeness of trait information

The amount of entries with available information for the individual grouping features varied strongly for the European, North American, and Australian databases (Table 3). By contrast, the New Zealand database contained complete trait information for most of the investigated grouping features (between 94 % and 100 %).

Table 3: Displayed is the percentage of entries that have information for the individual grouping features per database.

Database	Body form	Oviposition	Voltinism	Locomotion	Size	Respiration	Feeding mode
EU	7	16	24	33	11	56	65
NOA	28	13	47	52	73	44	63
AUS	5	48	51	42	78	70	99
NZ	100	94	100	99	100	100	99

3 Description of trait aggregation methods

Traits of the harmonized grouping feature databases were aggregated to family-level using three approaches. I) we directly aggregated taxa to family-level giving equal weight to every species. We denote this aggregation as *direct_agg*. For the *direct_agg* we tested aggregating with the median and the mean. We added *median* or *mean* to *direct_agg* to indicate when we used which method. II) taxa were aggregated stepwise, i.e first to the genus-level and subsequently to the family-level. By using this approach, we give equal weights to each genus. Hereafter, we abbreviate this aggregation type as *stepwise_agg*. We tested the *stepwise_agg* using the mean and the median, using the same naming as for the *direct_agg*. III) taxa were aggregated using a weighted mean approach, denoted as *weighted_agg*. This method weights the genera according to the number of their species present in the databases.

Trait affinities ranged between 0 and 1. Hence, the maximum differences possible is 1 or -1 (corresponds to 100 %). For convenience, we report absolute trait differences.

4 Differences in trait affinities obtained by trait aggregation methods compared to traits assigned at family-level

Aggregated trait affinities using five trait aggregation methods (*direct_agg (median)*, *direct_agg (mean)*, *stepwise_agg (median)*, *stepwise_agg (mean)*, and *weighted_agg*) were compared to trait affinities assigned at family-level by experts, which were available for the Australian and North American database for a limited subset of grouping features and taxa. For the Australian database, we compared aggregated trait affinities with assigned trait affinities resolved at family-level for the grouping features feeding mode and size by using data from Chessman 2018 ([24]). We could carry out the comparison to all taxa available in Chessman 2018, which contained trait information for 220 families. Considering each factor combination of family and investigated trait this amounts to 1760 cases. For the North American database, we compared aggregated trait affinities with assigned trait affinities on family-level for the grouping features feeding mode, respiration, size, voltinism, and locomotion from Pyne et al. Trait information in the Pyne et al. database was available on the categorical scale and was converted to binary traits prior to the comparison with aggregated trait affinities. Trait information on family-level in the Pyne et al. database was available for 94 families of which all were present in the aggregated North American database (total number of cases 1598).

The percentage of differing cases of trait affinities obtained by the trait aggregation methods compared to trait affinities assigned at family-level varied between 16.18 % and 22.9 % for the Australian database. For the North American database, comparison of the trait aggregation methods to trait affinities assigned at family-level yielded between 15.3 % and 47 % differing cases (Table 4).

For both databases maximum differences of 1 occurred for all investigated grouping features (Figure 1 and Figure 2). In general, trait aggregation methods using the median yielded fewer cases with differences compared to approaches using the mean. However, using the median produced greater differences for both databases.

Table 4: Amount of differing cases, the minimum and maximum, and means and standard deviations of absolute differences between trait affinities assigned at family-level and aggregated trait affinities.

Database	Comparison to traits at fam.-lvl.	Differing cases [%]	Min. differences	Max. differences	Mean abs. differences	SD abs. differences
Australia (Chessman)	direct_agg (median)	16.53	0.01	1.00	0.45	0.27
	direct_agg (mean)	23.24	< 0.01	0.99	0.34	0.23
	stepwise_agg (median)	17.90	0.01	1.00	0.42	0.26
	stepwise_agg (mean)	23.24	< 0.01	0.99	0.33	0.22
	weighted_agg	23.24	< 0.01	1.00	0.34	0.24
North America (Pyne)	direct_agg (median)	15.33	0.17	1.00	0.70	0.26
	direct_agg (mean)	47.00	< 0.01	1.00	0.30	0.26
	stepwise_agg (median)	18.00	0.08	1.00	0.63	0.28
	stepwise_agg (mean)	47.00	< 0.01	1.00	0.30	0.27
	weighted_agg	47.00	< 0.01	1.00	0.31	0.28

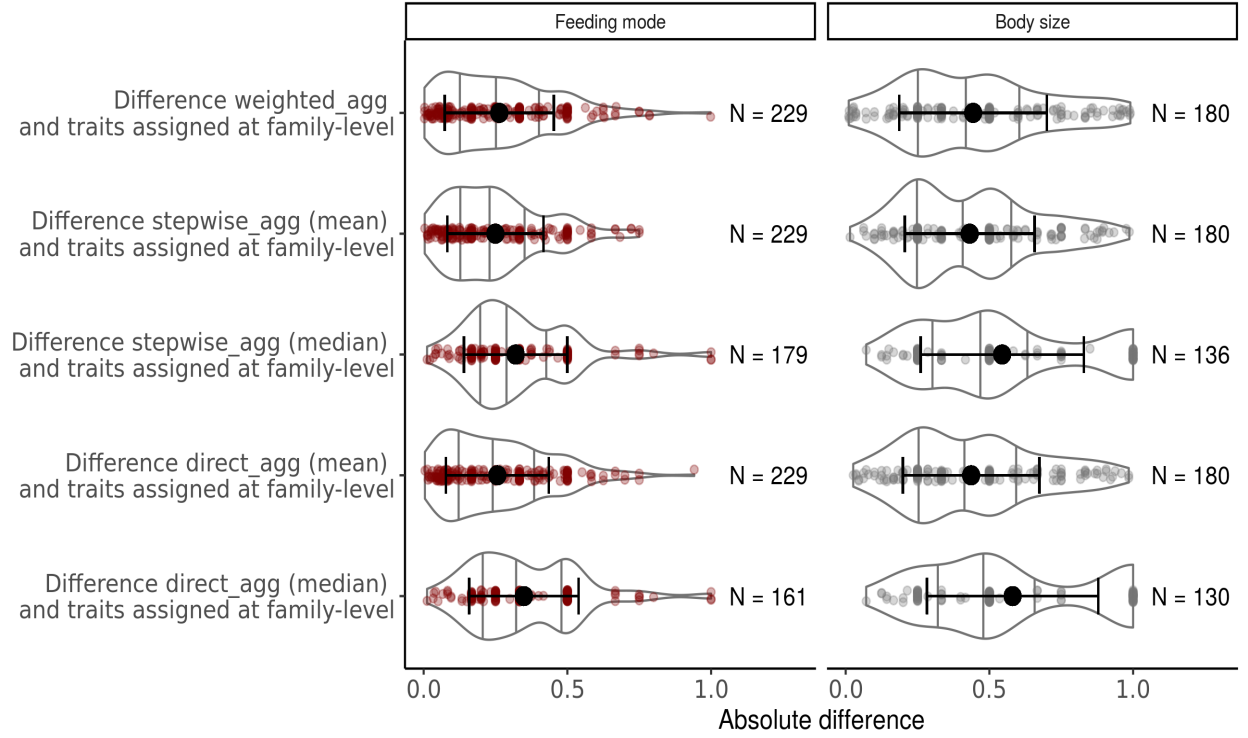


Figure 1: Absolute differences in trait affinities between aggregated traits and traits assigned at family-level by Chessman 2018 for the grouping features feeding mode and body size. N denotes the number of cases for each comparison. The black dot indicates the mean absolute difference, the error bars the standard deviation. The gray horizontal lines show the 25th, 50th and 75th quantile of the density estimate.

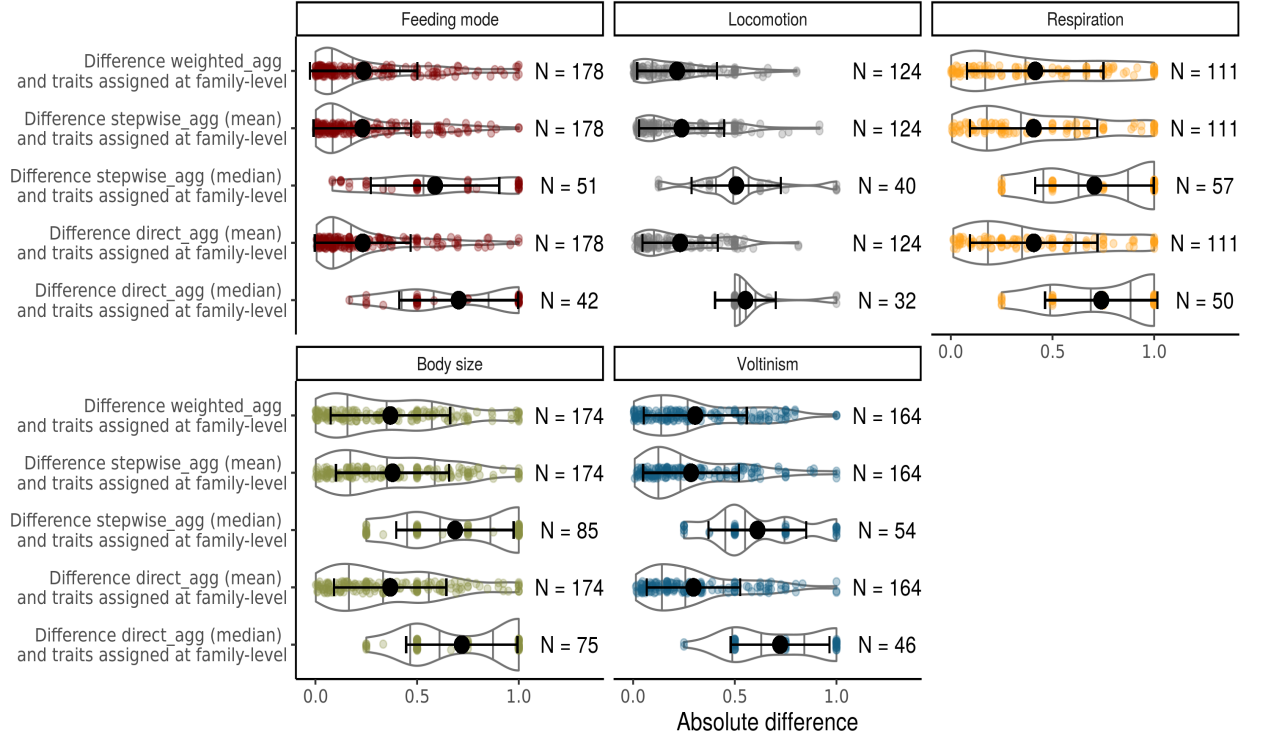


Figure 2: Absolute differences in trait affinities between aggregated traits and traits assigned at family-level by Pyne et al. for the grouping features feeding mode, locomotion, respiration, body size and voltinism. N denotes the number of cases for each comparison. The black dot indicates the mean absolute difference, the error bars the standard deviation. The gray horizontal lines show the 25th, 50th and 75th quantile of the density estimate.

5 Re-analysis of Szöcs et al. 2014 using harmonized and aggregated grouping features

To investigate how harmonizing grouping features and aggregating invertebrate traits might change the results in the analysis of trait-environment relationships we replicated the data analysis of Szöcs et al. 2014 ([3]) using the harmonized grouping features *Body size*, *Feeding mode*, *Locomotion*, *Reproduction/Oviposition*, *Respiration* and *Voltinism* (21 grouping features have been used in total) and aggregated traits using the aforementioned aggregation methods. The harmonized grouping features used are those that responded strongly to salinity in the study of Szöcs et. al. 2014, except for life cycle duration. For testing the effect of aggregated traits we assigned to each taxon in Szöcs et al. 2014 the aggregated trait value from the established harmonized European grouping feature database for its corresponding family.

Here, we limit our analysis to the RDA of traits constrained by electric conductivity of the original publication (see appendix for a more in-depth comparison to the original results). Overall, using the harmonized grouping features lead only to slightly different results in comparison to the original analysis. Sites with high salinity were characterized by multivoltine, oviparous, gill-respiring, and shredder species. Only species with the trait life cycle duration > 1 year fail to characterize sites with high salinization. Also, life cycle duration ≤ 1 year is not anymore characterizing sites not impacted by salinity. Like in the original analysis, transition and upstream sites from the point source are characterized by univoltine species and species that lay their eggs in an aquatic environment. The usage of aggregated traits yielded similar species scores for every aggregation method. Also, for every aggregation method compared, using at family-level aggregated traits did only slightly change the RDA species scores compared to not aggregated traits (Figure 3). Hence, the interpretation of the trait composition is the same as when only using harmonized grouping features. An overview of the harmonization for the European trait databases can be found in the supporting information in section .

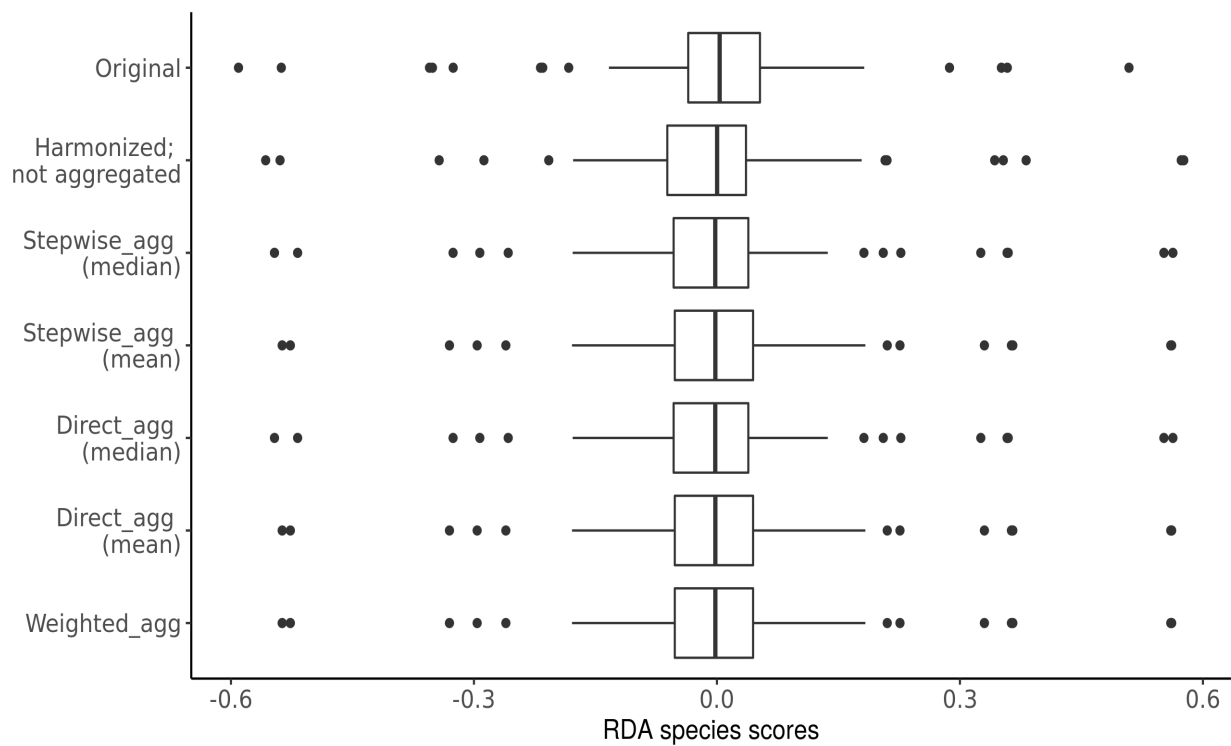


Figure 3: Species scores obtained by RDA from the original analysis ([3]), using harmonized grouping features, and using harmonized grouping features with assigned trait affinities from traits aggregated to family-level.

6 Discrepancies of invertebrate trait definitions

Definitions of grouping features and traits varied in their level of detail. The Tachet, Freshwaterecology, and North American (Twardochleb) trait databases provided more detailed descriptions of their trait information compared to the North American (Vieira) and New Zealand databases. An exception is the Australian trait database which is a collection of seven trait datasets ([13]). Thus, grouping features occur multiple times with varying differentiation into traits. Depending on the dataset trait information is described with more or less detail.

The definition of grouping features varied across databases mainly concerning their differentiation into traits but also in their scope. We provide a summary of discrepancies in trait definitions in the appendix (Table S9). Both, differences in differentiation and scope can lead to discrepancies in trait definitions. For example, for the grouping feature feeding mode discrepancies arise because traits are assigned in different ways. Tachet defines predators as carvers, engulfers and swallows. By contrast, in the North American (Twardochleb) database predators are defined as engulfers and carnivorous piercers. In turn, in the Tachet database, piercers are defined as a separate trait encompassing herbivorous and carnivorous piercers. Furthermore, the scope in the Freshwaterecology database for feeding mode is primarily on the food source of a species (except for filterers), while the other databases focus on the strategies of food acquisition. Therefore, the Freshwaterecology database defines e.g. predator as "eating from prey", while the other databases use the mouthpart morphology in their definition. The Tachet database captures the food source in an additional grouping feature. Varying levels of differentiation are also present in all other investigated grouping features between the trait databases (see for details Table 5 and Table S9). Locomotion definitions differ also in scope between databases. Freshwaterecology and New Zealand databases describe locomotion as the way of movement of an organism, Tachet as substrate relation, the North American (Vieira) as how organisms deal with flow, Australia as attachment, and the North American (Twardochleb) database includes among the way of movement also the location of movement. Similarly, regarding reproduction trait databases differ in their scope. Reproduction is captured in one grouping feature and defined as location of oviposit clutches and mode of reproduction in the Freshwaterecology and Tachet databases. North America (Vieira) provides information on the oviposition location but not on reproductive behavior. The Australian database report traits for reproductive behavior but also on oviposition site. The New Zealand database distinguishes three grouping features related to reproduction: reproductive technique, oviposition site (e.g. water surface, terrestrial), and egg/egg mass (e.g. free, cemented).

Various codings of traits are used throughout the databases (e.g. binary, fuzzy, continuous). The freshwaterecology and Australian use different codings in their databases. Tachet and the New Zealand database use exclusively fuzzy coding. Both North American trait databases contain categorical grouping features that can be converted into traits using a binary coding (Table 5). Binary coding represents a simple approach in which a taxon either expresses a trait or not. Fuzzy coding characterizes the affinity of an organism to exert a certain trait. It is used to account for plasticity in traits, e.g. taking into account that traits can change over the development time of an organism. Usually, fuzzy coded affinities are converted into proportional values. Continuous coding is used for traits like body size.

Table 5: Number of traits per grouping feature and type of coding of the traits for the respective grouping feature per database.

Database	Feeding Mode	Voltinism	Locomotion	Respiration	Reproduction	Size	Body Form
Freshwater-ecology	10	6	6	7	9	-	-
	10 point assignment system	single category assignment system	10 point assignment system	presence/absence assignment system	-	-	-
Tachet	7	3	8	5	8	7	-
North America (Twardochleb)	fuzzy [0 – 3]		fuzzy [0 – 5]	fuzzy [0 – 3]		-	-
	6	3	10	3	10	3	-
North America (Vieira)			binary		-	-	-
	8	3	9	8	10	3	5
Australia			binary		-	-	-
	16 ^a	7	9	10	13 ^b	9	4
New Zealand	binary; proportional [0 – 1]; fuzzy [0 – 3]		binary; fuzzy [0 – 3]	binary; proportional scale [0 – 1]; fuzzy [0 – 3]	categorical	binary; continuous; fuzzy [0 – 3]	fuzzy codes [0 – 3]
	6	3	4	4	4	5	4
fuzzy [0 – 3]							

^a Some of the traits were similar (e.g. trait *Shredder*, *Shredder*, *Detritivore*, and *Collector*, *Shredder*).

^b Many traits were rather comments than traits in the original database and were not considered.

Discussion

References

- [1] Brian J. McGill et al. “Rebuilding community ecology from functional traits”. In: *Trends in Ecology & Evolution* 21.4 (Apr. 1, 2006), pp. 178–185. ISSN: 0169-5347. DOI: 10.1016/j.tree.2006.02.002. URL: <http://www.sciencedirect.com/science/article/pii/S0169534706000334> (visited on 08/07/2020).
- [2] N. LeRoy Poff et al. “Developing linkages between species traits and multiscaled environmental variation to explore vulnerability of stream benthic communities to climate change”. In: *Journal of the North American Benthological Society* 29.4 (Dec. 1, 2010), pp. 1441–1458. ISSN: 0887-3593. DOI: 10.1899/10-030.1. URL: <https://www.journals.uchicago.edu/doi/abs/10.1899/10-030.1> (visited on 06/21/2019).
- [3] Eduard Szöcs et al. “Effects of anthropogenic salinization on biological traits and community composition of stream macroinvertebrates”. In: *Science of The Total Environment* 468–469 (Jan. 2014), pp. 943–949. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2013.08.058. URL: <https://linkinghub.elsevier.com/retrieve/pii/S0048969713009728> (visited on 03/26/2020).
- [4] Avit Kumar Bhowmik and Ralf B. Schäfer. “Large Scale Relationship between Aquatic Insect Traits and Climate”. In: *PLoS ONE* 10.6 (June 16, 2015). ISSN: 1932-6203. DOI: 10.1371/journal.pone.0130025. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4469582/> (visited on 06/21/2019).
- [5] Salomé Menezes, Donald J. Baird, and Amadeu M. V. M. Soares. “Beyond taxonomy: a review of macroinvertebrate trait-based community descriptors as tools for freshwater biomonitoring”. In: *Journal of Applied Ecology* 47.4 (2010), pp. 711–719. ISSN: 1365-2664. DOI: 10.1111/j.1365-2664.2010.01819.x. URL: <https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2664.2010.01819.x> (visited on 11/28/2019).
- [6] T. R. E. Southwood. “Habitat, the Templet for Ecological Strategies?” In: *Journal of Animal Ecology* 46.2 (1977), pp. 337–365. ISSN: 0021-8790. DOI: 10.2307/3817. URL: <https://www.jstor.org/stable/3817> (visited on 06/18/2019).
- [7] W. C. E. P. Verberk, C. G. E. van Noordwijk, and A. G. Hildrew. “Delivering on a promise: integrating species traits to transform descriptive community ecology into a predictive science”. In: *Freshwater Science* 32.2 (June 1, 2013), pp. 531–547. ISSN: 2161-9549. DOI: 10.1899/12-092.1. URL: <https://www.journals.uchicago.edu/doi/10.1899/12-092.1> (visited on 04/16/2018).
- [8] Núria Bonada, Sylvain DoléDec, and Bernhard Statzner. “Taxonomic and biological trait differences of stream macroinvertebrate communities between mediterranean and temperate regions: implications for future climatic scenarios”. In: *Global Change Biology* 13.8 (Aug. 2007), pp. 1658–1671. ISSN: 1354-1013, 1365-2486. DOI: 10.1111/j.1365-2486.2007.01375.x. URL: <http://doi.wiley.com/10.1111/j.1365-2486.2007.01375.x> (visited on 06/25/2018).

- [9] Philippe Usseglio-Polatera et al. “Biomonitoring through biological traits of benthic macroinvertebrates: how to use species trait databases?” In: *Hydrobiologia* 422.0 (Apr. 1, 2000), pp. 153–162. ISSN: 1573-5117. DOI: 10.1023/A:1017042921298. URL: <https://doi.org/10.1023/A:1017042921298> (visited on 06/21/2019).
- [10] Astrid Schmidt-Kloiber and Daniel Hering. “www.freshwaterecology.info – An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences”. In: *Ecological Indicators* 53 (June 1, 2015), pp. 271–282. ISSN: 1470-160X. DOI: 10.1016/j.ecolind.2015.02.007. URL: <http://www.sciencedirect.com/science/article/pii/S1470160X15000758> (visited on 06/21/2019).
- [11] Nicole K M Vieira et al. “A Database of Lotic Invertebrate Traits for North America”. In: (2006), p. 19.
- [12] Ngaire Philips and Brian Smith. “New Zealand Freshwater Macroinvertebrate Trait Database”. In: *NIWA* (2018).
- [13] Ben J. Kefford et al. “An integrated database of stream macroinvertebrate traits for Australia: concept and application”. In: *Ecological Indicators* 114 (July 2020), p. 106280. ISSN: 1470160X. DOI: 10.1016/j.ecolind.2020.106280. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1470160X2030217X> (visited on 05/05/2020).
- [14] Sylvie Tomanova, Edgar Goitia, and Jan Helešić. “Trophic Levels and Functional Feeding Groups of Macroinvertebrates in Neotropical Streams”. In: *Hydrobiologia* 556.1 (Feb. 1, 2006), pp. 251–264. ISSN: 1573-5117. DOI: 10.1007/s10750-005-1255-5. URL: <https://doi.org/10.1007/s10750-005-1255-5> (visited on 08/14/2020).
- [15] N. LeRoy Poff et al. “Functional trait niches of North American lotic insects: traits-based ecological applications in light of phylogenetic relationships”. In: *Journal of the North American Benthological Society* 25.4 (Dec. 1, 2006), pp. 730–755. ISSN: 0887-3593. DOI: 10.1899/0887-3593(2006)025[0730:FTNNA]2.0.CO;2. URL: <https://www.journals.uchicago.edu/doi/10.1899/0887-3593%282006%29025%5B0730%3AFTNNA%5D2.0.CO%3B2> (visited on 01/07/2020).
- [16] Paul K. Botwe et al. “Effects of salinity and flow interactions on macroinvertebrate traits in temporary streams”. In: *Ecological Indicators* 89 (June 2018), pp. 74–83. ISSN: 1470160X. DOI: 10.1016/j.ecolind.2018.01.036. URL: <https://linkinghub.elsevier.com/retrieve/pii/S1470160X18300384> (visited on 08/13/2020).
- [17] Ralf B. Schäfer et al. “A trait database of stream invertebrates for the ecological risk assessment of single and combined effects of salinity and pesticides in South-East Australia”. In: *Science of The Total Environment* 409.11 (May 2011), pp. 2055–2063. ISSN: 00489697. DOI: 10.1016/j.scitotenv.2011.01.053. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0048969711001124> (visited on 04/12/2018).
- [18] Dénes Schmera et al. “A proposed unified terminology of species traits in stream ecology”. In: *Freshwater Science* 34.3 (Sept. 2015), pp. 823–830. ISSN: 2161-9549, 2161-9565. DOI: 10.1086/681623. URL: <https://www.journals.uchicago.edu/doi/10.1086/681623> (visited on 11/06/2019).
- [19] Lee E. Brown et al. “Functional diversity and community assembly of river invertebrates show globally consistent responses to decreasing glacier cover”. In: *Nature Ecology & Evolution* 2.2 (Feb. 2018), pp. 325–333. ISSN: 2397-334X. DOI: 10.1038/s41559-017-0426-x. URL: <http://www.nature.com/articles/s41559-017-0426-x> (visited on 02/04/2019).

- [20] Florian D. Schneider et al. “Towards an ecological trait-data standard”. In: *Methods in Ecology and Evolution* (Sept. 16, 2019). Ed. by David Orme, pp. 2041–210X.13288. ISSN: 2041-210X, 2041-210X. DOI: 10.1111/2041-210X.13288. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13288> (visited on 10/10/2019).
- [21] Donald J. Baird et al. “Toward a knowledge infrastructure for traits-based ecological risk assessment”. In: *Integrated Environmental Assessment and Management* 7.2 (2011). _eprint: <https://setac.onlinelibrary.wiley.com/doi/pdf/10.1002/ieam.129>, pp. 209–215. ISSN: 1551-3793. DOI: 10.1002/ieam.129. URL: <https://setac.onlinelibrary.wiley.com/doi/abs/10.1002/ieam.129> (visited on 08/11/2020).
- [22] Joseph M. Culp et al. “Incorporating traits in aquatic biomonitoring to enhance causal diagnosis and prediction”. In: *Integrated Environmental Assessment and Management* 7.2 (2011). _eprint: <https://setac.onlinelibrary.wiley.com/doi/pdf/10.1002/ieam.128>, pp. 187–197. ISSN: 1551-3793. DOI: 10.1002/ieam.128. URL: <https://setac.onlinelibrary.wiley.com/doi/abs/10.1002/ieam.128> (visited on 08/11/2020).
- [23] Pilière A. F. H. et al. “On the importance of trait interrelationships for understanding environmental responses of stream macroinvertebrates”. In: *Freshwater Biology* 61.2 (Jan. 5, 2016), pp. 181–194. ISSN: 0046-5070. DOI: 10.1111/fwb.12690. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/fwb.12690> (visited on 04/17/2018).
- [24] B. C. Chessman. “Dissolved-oxygen, current and temperature preferences of stream invertebrates estimated from field distributions: application to assemblage responses to drought”. In: *Hydrobiologia* 809.1 (Mar. 1, 2018), pp. 141–153. ISSN: 1573-5117. DOI: 10.1007/s10750-017-3455-1. URL: <https://doi.org/10.1007/s10750-017-3455-1> (visited on 09/08/2019).

Supporting Information

Comparison of the trait aggregation methods with each other

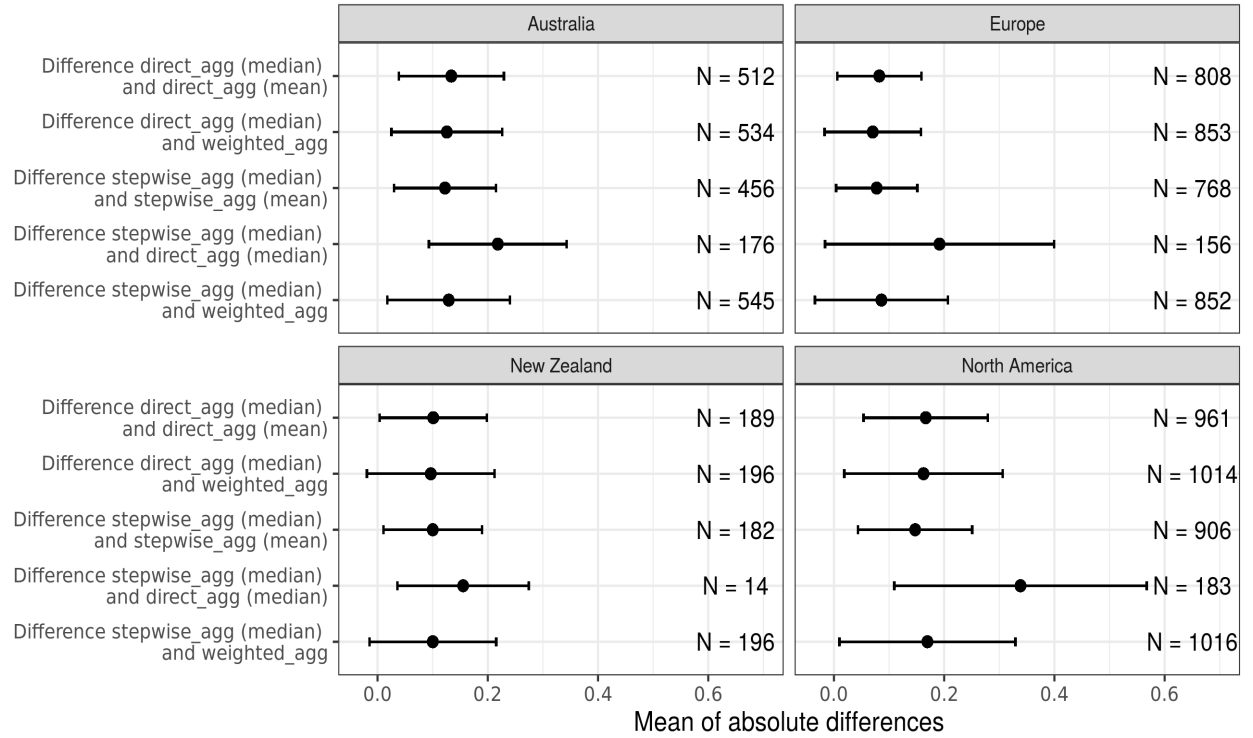


Figure 4: Means of absolute differences in trait affinities with standard deviations per region for all grouping features. Compared aggregation methods are displayed on the y-axis. N indicates the number of cases where differences occurred.

Re-analysis of Szöcs et al. 2014 using harmonized and aggregated grouping features.

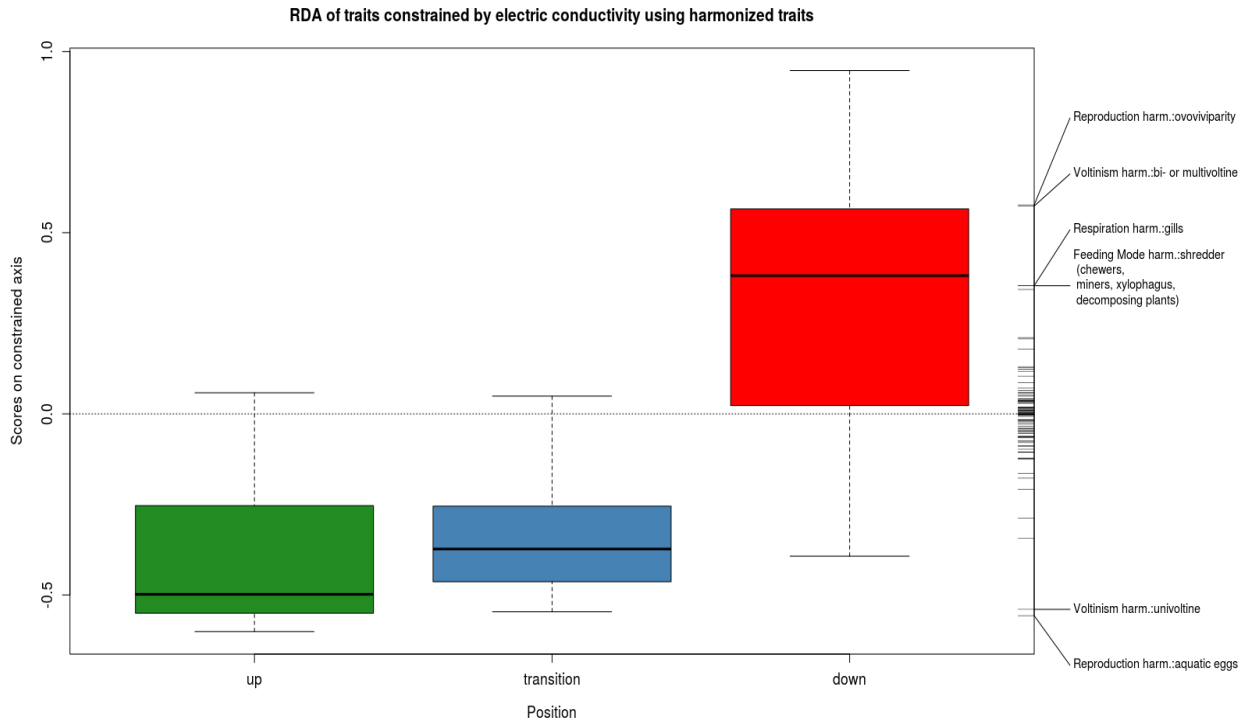


Figure 5: RDA of traits constrained by electric conductivity using harmonized grouping features. Boxplot of site scores along the conductivity axis (31.44% explained variance, $p = 0.001$, 1000 permutations). Rug on the left indicates trait scores on the conductivity axis. Only traits with a mahalanobis distance greater than 5.02 were labeled in accordance to the procedure in Szöcs et al. 2014.

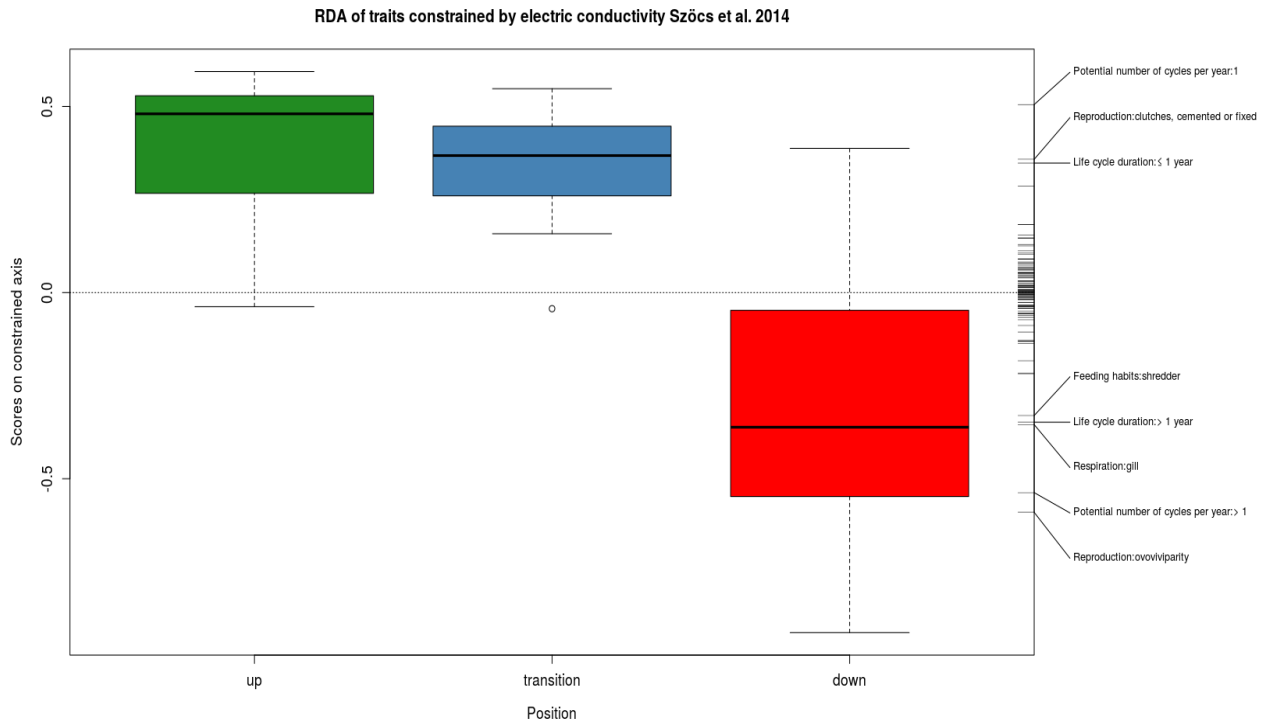


Figure 6: RDA of traits constrained by electric conductivity. Boxplot of site scores along the conductivity axis (30.09% explained variance, $p = 0.001$, 1000 permutations). Rug on the left indicates trait scores on the conductivity axis. Only traits with a mahalanobis distance greater than 5.02 were labeled.

Trait distribution along first RDA axis

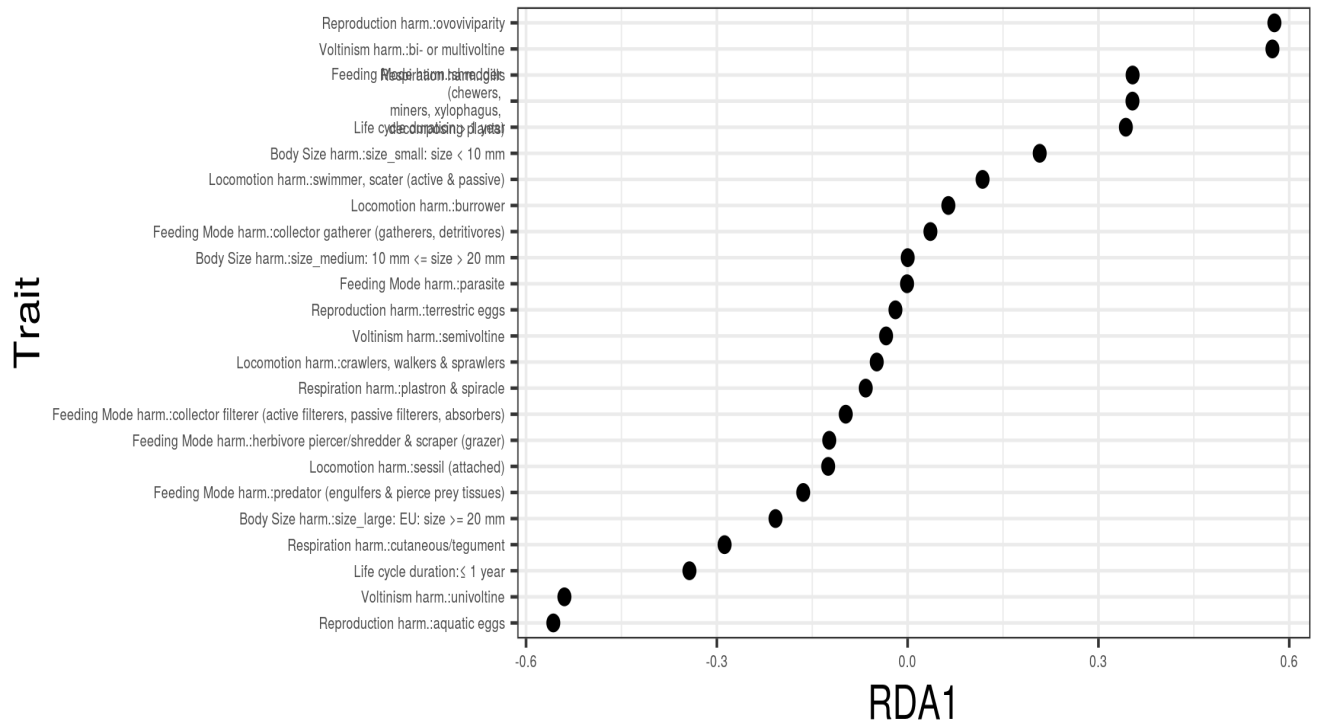


Figure 7: Trait scores on the first RDA axis for harmonized traits and traits of the grouping feature *life cycle duration*.

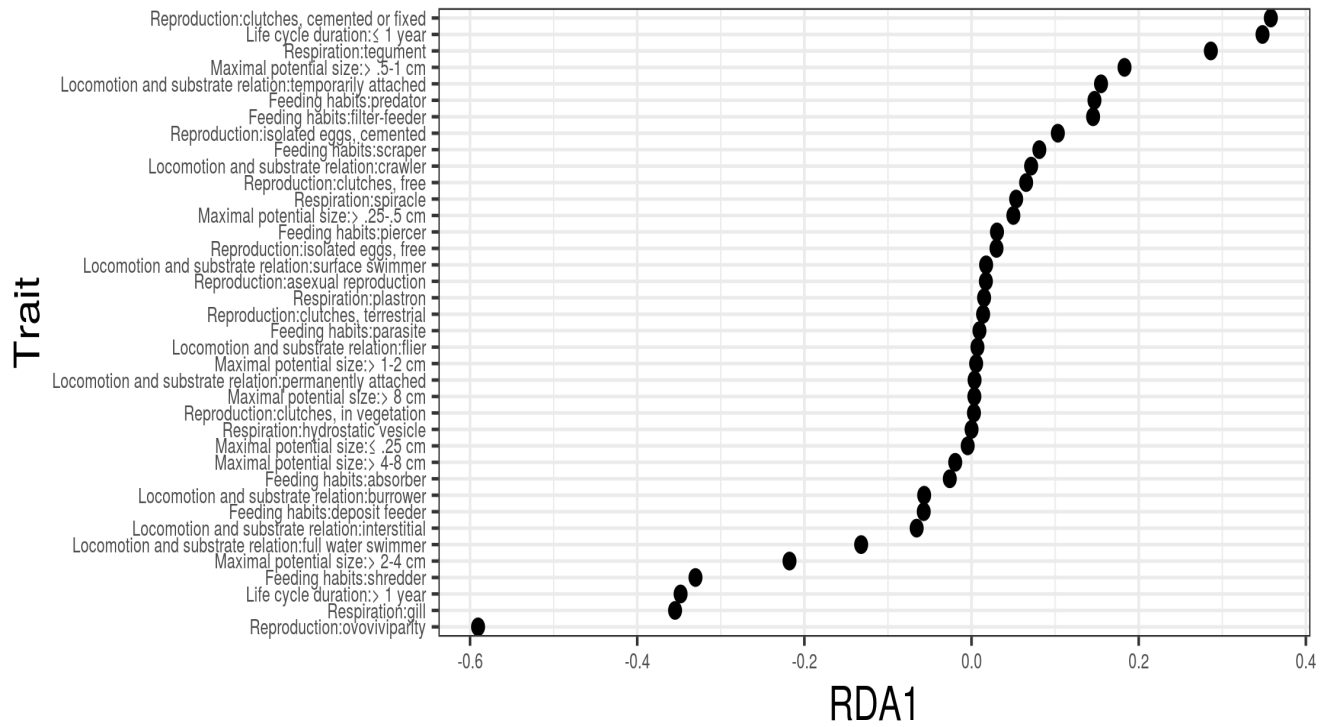


Figure 8: Trait scores on the first RDA axis for the traits responding to high salinity in Szöcs et al. 2014 .

Linear models of trait proportions

Linear models of trait proportions with harmonized traits:

Table 6: Results of linear models for the four selected harmonized traits and life cycle duration > 1 year. Trait proportions were logit transformed prior model building, estimates are on the logit scale. Although years were statistically not significant we kept this factor in the model to avoid temporal autocorrelation. Bold values indicate statistically significant effects ($p < 0.05$).

	Feeding mode: shredder	Life cycle duration: > 1 year	Voltinism: bi- or multivoltine	Reproduction: ovoviviparity	Respiration: gills
Intercept (= upstream)	-1.041	-0.486	0.375*	-0.823	0.092
Downstream	0.926	0.605	1.376	1.684	0.854
Downstream x 2008	-0.117	0.106	-0.235	-0.088	-0.317
Downstream x 2009	0.030	-0.056	0.001	0.245	0.180
Year 2008	-0.167	-0.115	0.033	-0.182	-0.151
Year 2009	0.175	0.086	-0.088	0.246	0.141

* $p.value = 0.055$

Linear models of trait proportions Szöcs et al. 2014:

Table 7: Results of linear models for the five selected traits for Szöcs et al. 2014. Trait proportions were logit transformed prior model building, estimates are on the logit scale. Although years were statistically not significant we kept this factor in the model to avoid temporal autocorrelation. Bold values indicate statistically significant effects ($p < 0.05$).

	Feeding habits: shredder	Life cycle duration: > 1 year	Cycles per year: > 1	Reproduction: ovoviviparity	Respiration: gills
Intercept (= upstream)	-0.853	-0.478	0.603	-0.838	0.111
Downstream	0.819	0.594	1.297	1.679	0.839
Downstream x 2008	-0.155	0.102	-0.227	-0.070	-0.314
Downstream x 2009	0.073	-0.053	-0.020	0.248	0.176
Year 2008	-0.122	-0.112	0.026	-0.192	-0.154
Year 2009	0.167	0.084	-0.104	0.250	0.139

Trait proportions over time

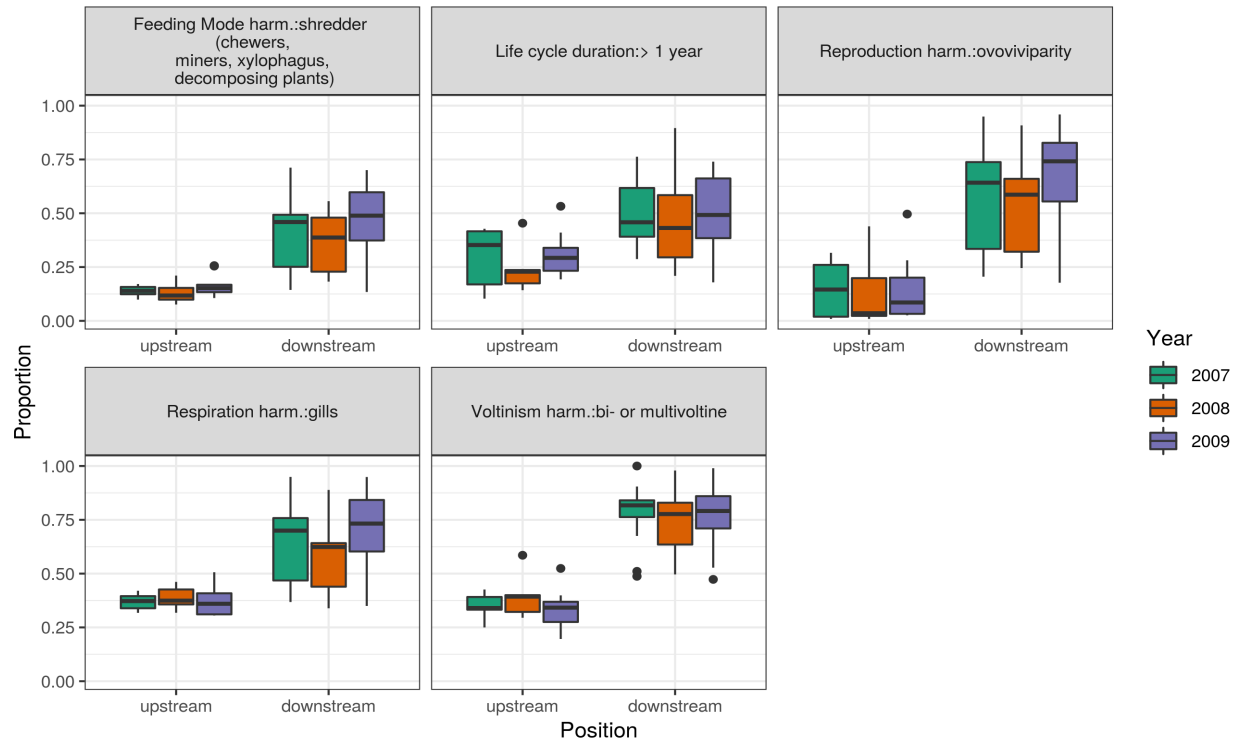


Figure 9: Proportions for the four harmonized traits that have been promoted by salinization and life cycle duration > 1 year for down- and upstream sites.

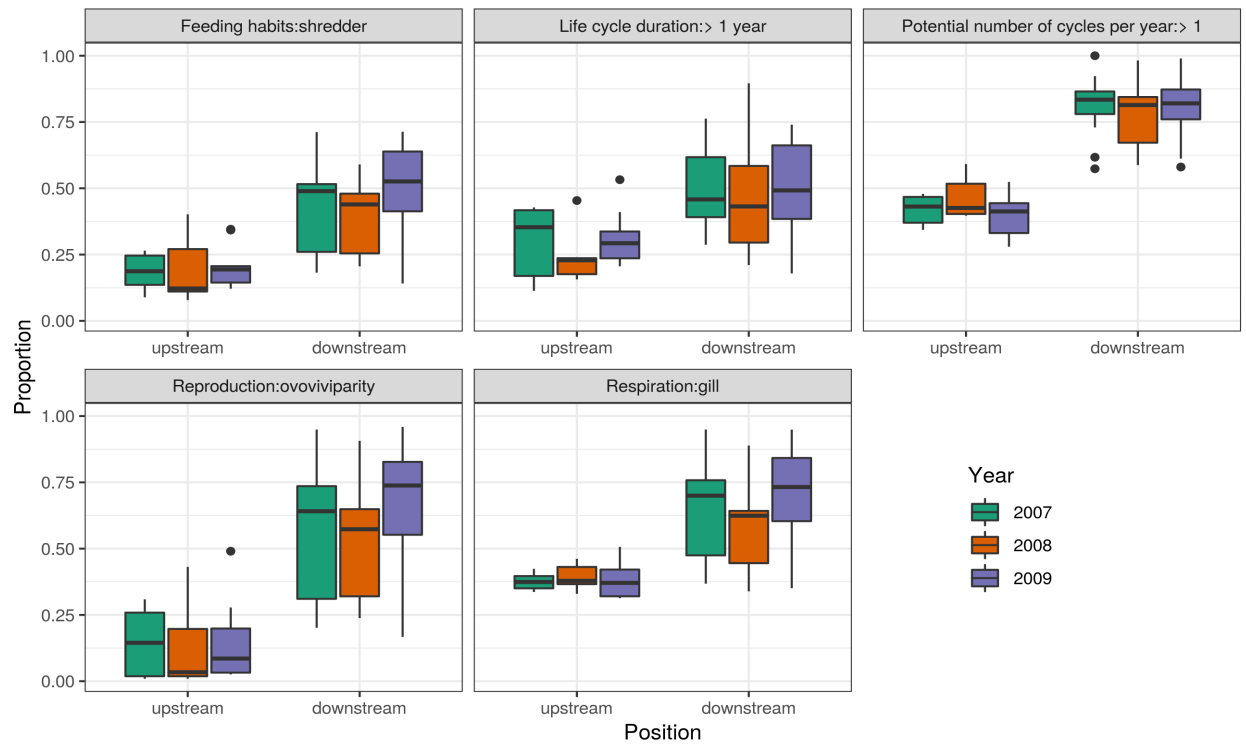


Figure 10: Proportions for five selected traits for down- and upstream sites (traits that have been promoted by salinization) from Szöcs et al. 2014.

Harmonization of the European trait databases

Table 8: Representation of traits per grouping feature and their harmonization for the European trait databases. The color coding indicates traits that have been harmonized. Cyan colored traits have not been used because they either represented ambiguous traits or traits that were not compatible with the traits of the other databases. Harmonization was done by assigning the maximum affinity of the allocated traits for the respective taxa to the harmonized trait.

Grouping feature	Freshwater ecology	Tachet	Harmonized traits
Voltinism	Semivoltine	Semivoltine	Semivoltine
Voltinism	Univoltine	Univoltine	Univoltine
Voltinism	Bivoltine	Polyvoltine	Bi/Multivoltine
Voltinism	Trivoltine		
Voltinism	Multivoltine		
Voltinism	Flexible		
Feeding Mode	Shredder	Shredder	Shredder
Feeding Mode	Miner	Deposit-feeder	Gatherer
Feeding Mode	Xylophagus	Absorber	Filterer
Feeding Mode	Gatherer	Filter-feeder	Herbivore
Feeding Mode	Active filterer	Scraper	Predator
Feeding Mode	Passive filterer	Predator	Parasite
Feeding Mode	Grazer	Parasite	
Feeding Mode	Predator	<i>Piercer (plants or animals)</i> ^a	
Feeding Mode	Parasite		
Feeding Mode	Other		
Locomotion	Swimming/scating	Surface swimmer	Swimmer
Locomotion	Swimming/diving	Full water swimmer	Burrower
Locomotion	Burrowing/boring	Burrower	Crawler
Locomotion		Interstitial	Sessil
Locomotion	Sprawling/walking	Crawler	
Locomotion	(semi) sessil	Temporarily attached	
Locomotion	Other	Permanently attached	
Locomotion		Flier	
Respiration	Tegument	Tegument	Tegument
Respiration	Gill	Gill	Gills
Respiration	Plastron	Plastron	Plastron, spiracle
Respiration	Spiracle (aerial)	Spiracle (aerial)	
Respiration	Hydrostatic vesicle	Hydrostatic vesicle (aerial)	
Respiration	Tapping (air stores of aq. plants)		
Respiration	Excursion/Extension (to surface)		
Body size		<= 0.25cm	Small (< 1 cm)
Body size		> 0.25 - 0.5cm	Medium (>= 1cm - 2 cm)
Body size		> 0.5- 1cm	Large (>= 2 cm)

Body size		> 1 – 2 cm	
Body size		2 – 4 cm	
Body size		4 – 8 cm	
Body size		> 8 cm	
Reproduction	ovovivipar	ovoviviparity	ovoviviparity
Reproduction	free isolated eggs	isolated eggs, free	aquatic eggs
Reproduction	cemented isolated eggs	isolated eggs, cemented	terrestrial eggs
Reproduction	fixed clutches	clutches, cemented or fixed	
Reproduction	free clutches	clutches, free	
Reproduction	clutches in vegetation	clutches, in vegetation	
Reproduction	terrestrial clutches	clutches, terrestrial	
Reproduction	asexual	asexual reproduction	
Reproduction	parasitic		

a Taxa exhibiting this trait have been assigned to predators or herbivores based on a classification by Philippe Usseglio-Polatera.

Table 9: Comparison of trait definitions between invertebrate trait databases. Only traits that are differently described across databases are listed. The definition is quoted if it enables differences to be identified, otherwise the differences are described. The hyphen indicates a missing trait. Reproduction was captured in multiple grouping features per database. Hence, differences for reproduction have been described in the paper. Body form traits are not different between databases, except that the North America (Vieira) database contains the trait Bluff (blocky) which does not appear in the other databases.

Trait	Freshwaterecology	Tachet	North America (Twardochleb)	North America (Vieira)	Australia	New Zealand
Feeding shredder	"Feed from fallen leaves, plant tis- sues, CPOM"	"Eat coarse detri- tus, plants or <i>ani- mal material</i> "	<ul style="list-style-type: none"> "Shred de- composing vascular plant tissue" Trait herbi- vore includes among others insect that shred <i>liv- ing aquatic plants</i> 	Shredder	<ul style="list-style-type: none"> Detrivore ^a Trait herbi- vore includes among others the trait shredder 	Shredders

Feeding predator	"Eating from prey"	<ul style="list-style-type: none"> • Carvers, & engulfers & swallowers • Piercers & (plants & animals) are an additional trait 	Engulfers ("ingest prey whole or in parts") & piercers ("prey tissues and suck fluids")	Predator	Piercer & engulfer	Predator
Feeding filter-feeder	Distinguishes between active and passive	No distinction between active and passive	No distinction between active and passive	No distinction between active and passive	No distinction between active and passive	No distinction between active and passive
Semivoltine	"One generation in two years"	"Life cycle lasts <i>at least</i> two years"	"< 1 generation per year"	"< 1 generation per year"	"< 1 generation per year"	"< 1 reproductive cycle per year"
Multivoltine	"More than <i>three</i> generations per year," <i>b</i>	"Able to complete <i>at least</i> two successive generations per year"	"> 1 generations per year"	"> 1 generations per year"	<ul style="list-style-type: none"> • 1-2 generations per year • bi/multivoltine • up to 5 generations per year • up to 10 generations per year 	"> 1 reproductive cycles per year"

Locomotion swimming	<ul style="list-style-type: none"> • Passive movement like floating or drifting (trait swimming/scating) • Active movement (trait swimming/diving) 	<ul style="list-style-type: none"> • Surface swimmers (over and under the water surface) • Full water swimmers (e.g. Baetidae). 	<p>”Adapted for ”fish-like” swimming”</p>	Swimmer	Distinguishes swimmer and skater	Swimmers (water column)
Locomotion burrowing	<p>”Burrowing in <i>soft</i> substrates or boring in <i>hard</i> substrates”</p>	<ul style="list-style-type: none"> • Burrowing ”within the first centimeters of the fine benthic fine sediment” • Differentiates also the trait interstitial (endobenthic) 	<p>”Inhabiting <i>fine</i> sediment of streams and lakes”</p>	Burrower	”Moving deep into the substrate and thus avoiding flow”	Burrowers (in-fauna)
Locomotion sprawling & walking	<p>”Sprawling or walking actively with legs, pseudopods or on a mucus”</p>	-	<p>Sprawling: ”inhabiting the surface of floating leaves of vascular hydrophytes or fine sediments”</p>	Sprawler	-	-

Locomotion crawling	-	"Crawling over the bottom substrate"	Defined as crawling on the surface of floating leaves or fine sediments on the bottom	-	Database contains traits: sprawler, climber and clinger.	Crawlers (epibenthic)
Locomotion sessile	Does not distinguish temporarily and permanently attached	Distinguishes temporarily and permanently attached	Does not distinguish temporarily and permanently attached	Does not distinguish temporarily and permanently attached	Distinguishes temporarily and permanently attached	Does not distinguish temporarily and permanently attached
Respiration & plastron & spiracle	Plastron and spiracle (aerial) are two separate traits	Definition includes respiration using air stores of aquatic plants	Plastron and spiracle combined into one trait	Distinguishes spiracular gills, plastron, atmospheric breathers and plant breathers	Plastron and spiracle (termed aerial) occur as separate and combined traits. Contains also traits: air (plants), atmospheric, and functional spiracles	Distinguishes plastron and spiracle (termed aerial)
Body size small	-	Multiple size classifications ^d	< 9 mm	< 9 mm	< 9 mm ^{a,c}	Multiple size classifications ^e
Body size medium	-		9 - 16 mm	9 - 16 mm	9 - 16 mm	
Body size large	-		> 16 mm	> 16 mm	> 16 mm	

^a Traits from Botwe et al.

^b Contains also bivoltine (two generations per year), trivoltine (three generations per year) and flexible.

^c Contains a size trait with numeric size values. Contains also traits classifying size like Tachet and like the North American trait databases.

^d Size classifications: ≤ 0.25 cm, $> 0.25 - 0.5$ cm, $0.5 - 1$ cm, $1 - 2$ cm, $2 - 4$ cm, $4 - 8$ cm, > 8 cm. No distinction into small, medium and large.

^e Size classifications: $> 0.25 - 0.5$ cm, $0.5 - 1$ cm, $1 - 2$ cm, $2 - 4$ cm, $4 - 8$ cm. No distinction into small, medium and large.