1   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.

2   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 (Chessman 2018). 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 1).

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 1: 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
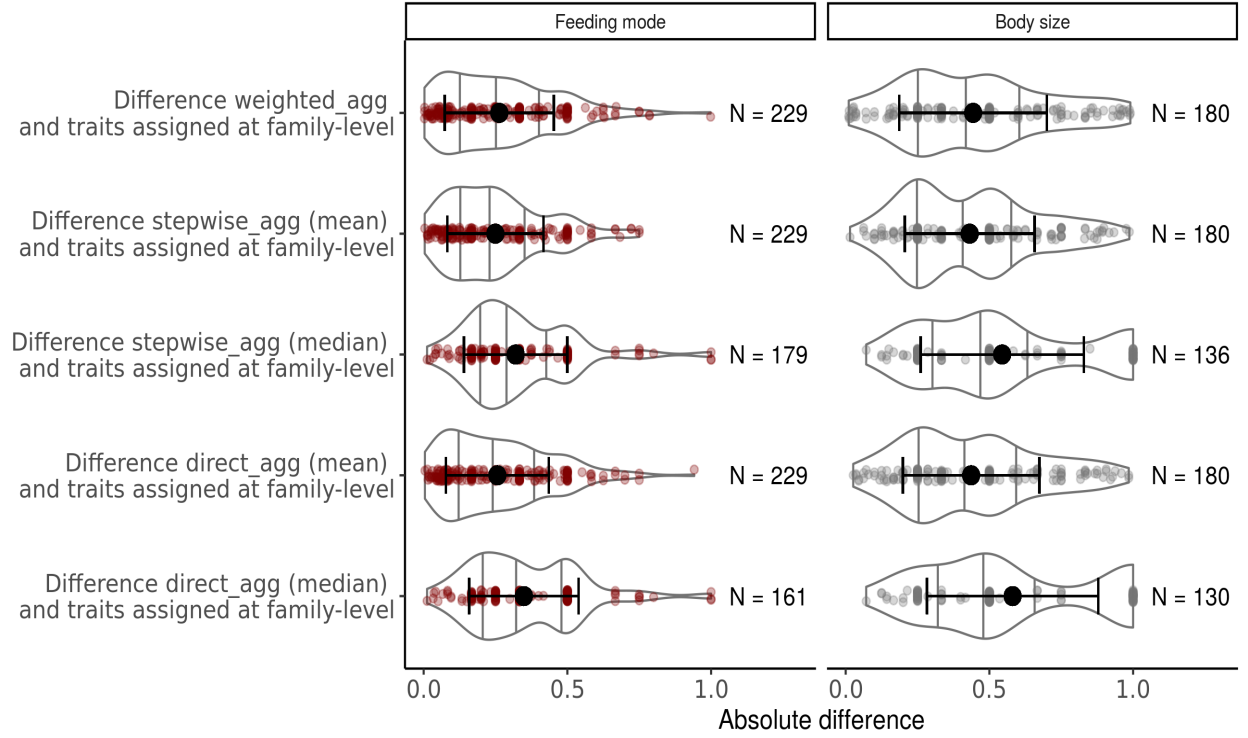
Figure 1: Absolute differences in trait affinities between aggregated traits and traits assigned at family-level by Chessman 2018 for the grouping features feeding mode and body size. N denotes the number of cases for each comparison. The black dot indicates the mean absolute difference, the error bars the standard deviation. The gray horizontal lines show the 25th, 50th and 75th quantile of the density estimate.
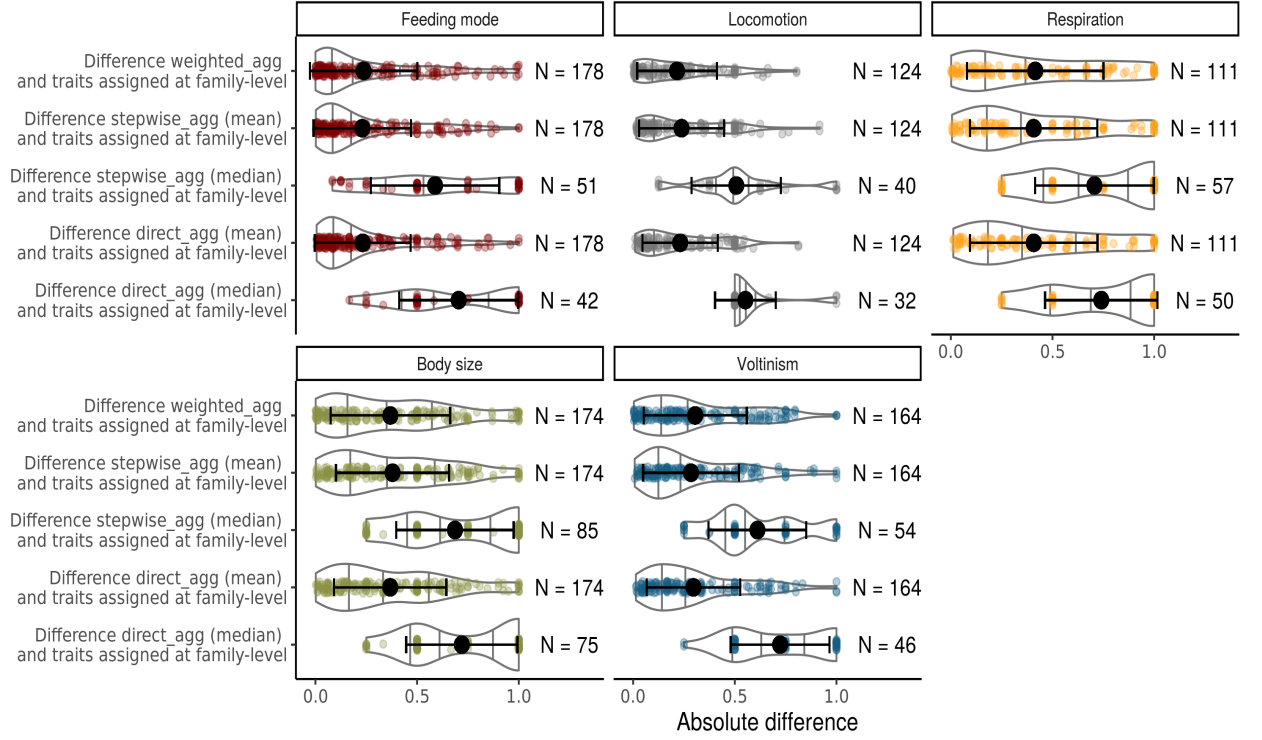
Figure 2: Absolute differences in trait affinities between aggregated traits and traits assigned at family-level by Pyne et al. for the grouping features feeding mode, locomotion, respiration, body size and voltinism. N denotes the number of cases for each comparison. The black dot indicates the mean absolute difference, the error bars the standard deviation. The gray horizontal lines show the 25th, 50th and 75th quantile of the density estimate.