

Supplement of J. Micropalaeontol., 37, 519–534, 2018
<https://doi.org/10.5194/jm-37-519-2018-supplement>
© Author(s) 2018. This work is distributed under
the Creative Commons Attribution 4.0 License.



Supplement of

Reproducibility of species recognition in modern planktonic foraminifera and its implications for analyses of community structure

Nadia Al-Sabouni et al.

Correspondence to: Isabel S. Fenton (isabel.fenton@cantab.net)

The copyright of individual parts of the supplement might differ from the CC BY 4.0 License.

Section S1: Species list

The list of the species included in this study. The species list was taken from Hemleben *et al.* (1989) but has been updated to reflect the current taxonomy, for example *Trilobatus* rather than *Globigerinoides*. Two species names (*Globorotalia bermudezi* and *Turborotalita cristata*) were used by participants despite not being in the Hemleben *et al.* (1989) list. Four species from Hemleben were not identified by any of the participants (*Gallitellia vivans*, *Globigerinella adamsi*, *Globoquadrina conglomerata*, *Hastigerina digitata*). The species used for the transfer function ANN analysis are marked with a ‘*’.

**Beella digitata* (Brady, 1879)
Berggrenia pumilio (Parker, 1962)
Candeina nitida d’Orbigny, 1839
Dentigloborotalia anfracta (Parker, 1967)
Globorotalia bermudezi Rögl & Bolli, 1973
**Globigerina bulloides* d’Orbigny, 1826
**Globigerina falconensis* Blow, 1959
**Globigerinella calida* (Parker, 1962)
**Globigerinella siphonifera* (d’Orbigny, 1839)
**Globigerinita glutinata* (Egger, 1893)
Globigerinita minuta (Natland, 1938)
Globigerinita uvula (Ehrenberg, 1861)
**Globigerinoides conglobatus* (Brady, 1879)
**Globigerinoides ruber* (d’Orbigny, 1839)
**Globigerinoides ruber* (pink) (d’Orbigny, 1839) considered a subspecies in Hemleben *et al.* (1989)
Globorotalia cavernula Bé, 1967
**Globorotalia crassaformis* (Galloway & Wissler, 1927)
**Globorotalia hirsuta* (d’Orbigny, 1839)
**Globorotalia inflata* (d’Orbigny, 1839)
**Globorotalia menardii* (Parker, Jones & Brady, 1865 after d’Orbigny, 1826 nomen nudum), combined with *G. tumida* for the ANN analysis
**Globorotalia scitula* (Brady, 1882)
Globorotalia theyeri Fleisher, 1974
**Globorotalia truncatulinoidea* (d’Orbigny, 1839)
**Globorotalia tumida* (Brady, 1877), combined with *G. menardii* for the ANN analysis
Globorotalia ungulata Bermúdez, 1961
Globorotaloidea hexagonus (Natland, 1938)
**Globoturborotalita rubescens* (Hofker, 1956), n.b. in this study this included *G. tenella*; in the ANN analysis, this was treated as *G. tenella*
**Globoturborotalita rubescens* (red) (Hofker, 1956), considered a form of *G. rubescens* in Hemleben *et al.* (1989); in the ANN analysis, this was treated as *G. rubescens*
Hastigerina pelagica (d’Orbigny, 1839)
**Neogloboquadrina dutertrei* (d’Orbigny, 1839)
**Neogloboquadrina incompta* (Cifelli, 1961) described as *N. pachyderma* (dex) in Hemleben *et al.* (1989)
**Neogloboquadrina pachyderma* described as *N. pachyderma* (sin) in Hemleben *et al.* (1989) (Ehrenberg, 1862)
**Orbulina universa* d’Orbigny, 1839
Orcadia riedeli (Rögl & Bolli, 1973)
**Pulleniatina obliquiloculata* (Parker & Jones, 1865)
**Sphaeroidinella dehiscens* (Parker & Jones, 1865)
Tenuitella fleisheri Li, 1987
Tenuitella iota (Parker, 1962)
Tenuitella parkerae (Brönnimann & Resig, 1971)
**Trilobatus sacculifer* (Brady, 1877)
**Trilobatus trilobus* (Brady, 1877), considered a subspecies of *T. sacculifer* in Hemleben *et al.* (1989)
Turborotalita clarkei (Rögl & Bolli, 1973)
Turborotalita cristata (Heron-Allen & Earland, 1929)
Turborotalita humilis (Brady, 1884)
**Turborotalita quinqueloba* (Natland, 1938)

Section S2: Figures

Slide analysis

Digital analysis

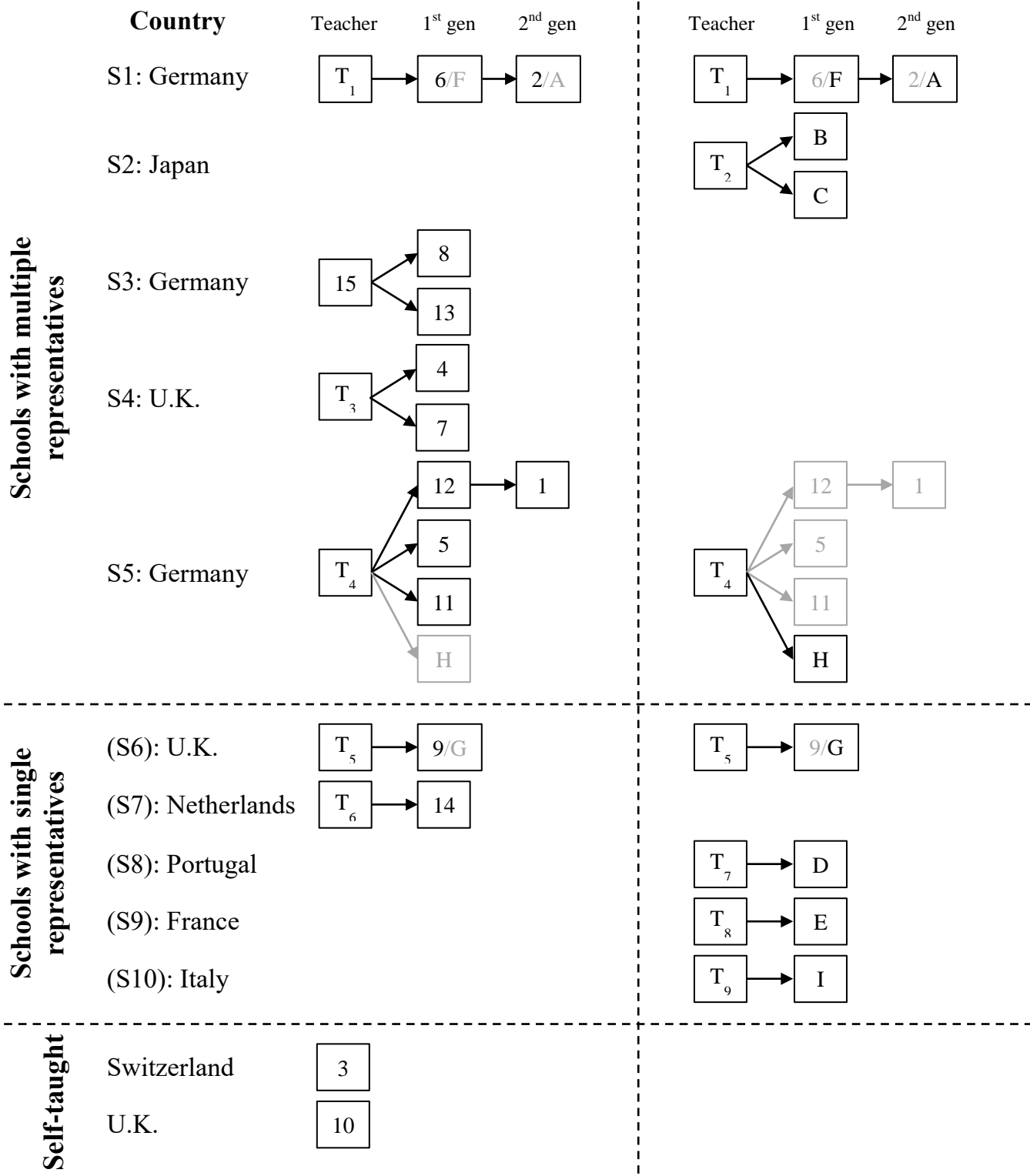


Figure S1: The structure of the taxonomic schools used in this analysis. Schools are relatively stand-alone entities, with little overlap of taxonomic teaching. Some schools (S1-S5; those which are numbered in Table 1) have multiple representatives in this analysis; others (labelled ‘-’ in Table 1) have only a single representative. A few individuals are self-taught. The country gives the dominant location of the representatives of that school at the time of the analysis. In most cases, the original teacher has not been involved in this analysis (T1-T9). However 15 is self-taught, but has taught other representatives, so is classified as a teacher.

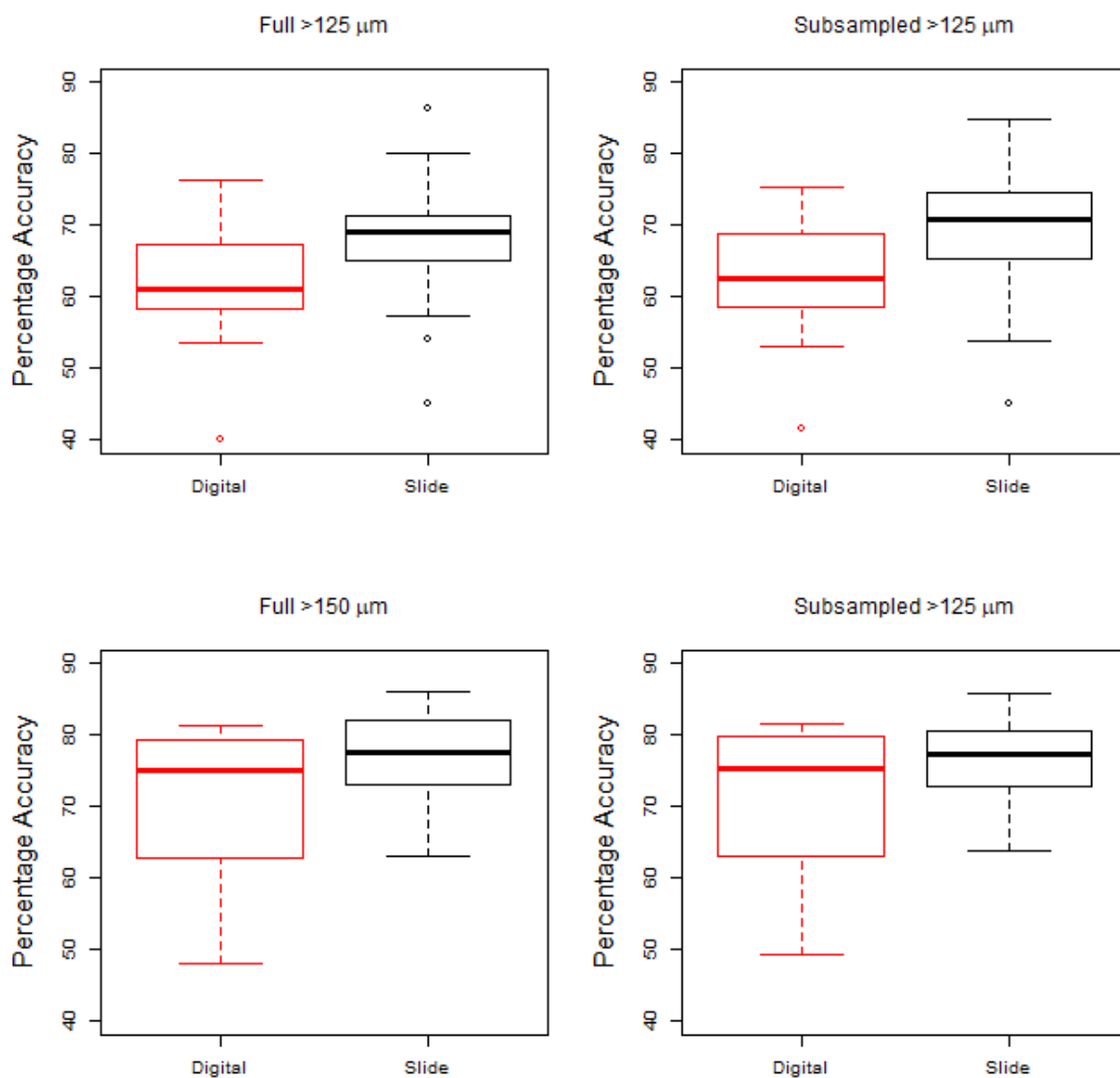


Figure S2: The influence of the number of participants on the percentage accuracy of the slide and digital analyses. The left column shows the full analysis. The right column shows the subsampled results of the recalculated consensus estimates based on equal numbers of slide and digital participants.

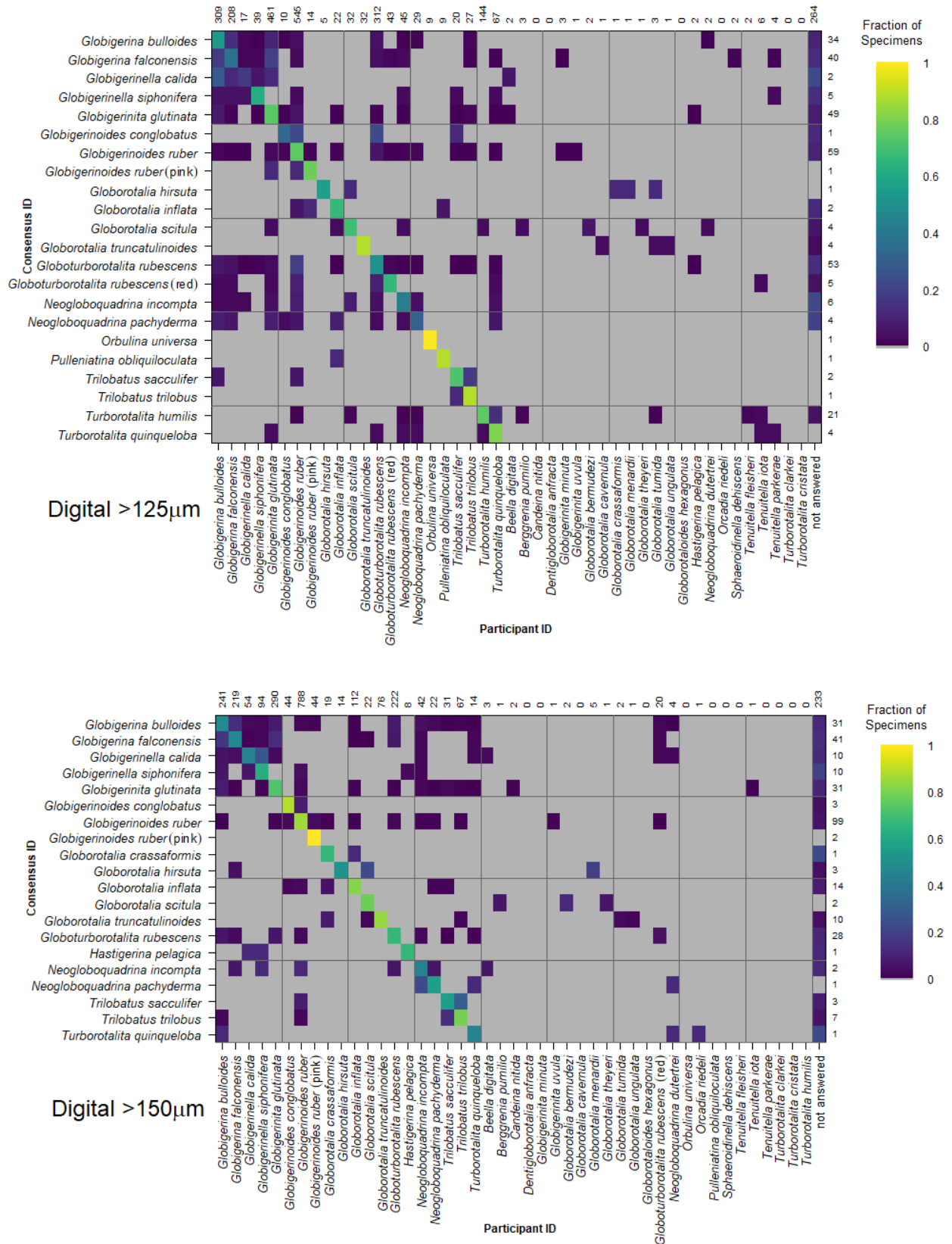


Figure S3: Confusion matrices for taxonomic identifications in the two digital analyses. The y-axis shows the consensus ID, and the x-axis shows the names given by the participants to the individual specimens. Where a specimen was always identified correctly, only one square in that row would be filled, indicating a fraction of 1. The numbers along the top indicate the number of specimens given that name in the analysis. The numbers at the right indicate the number of specimens of each species in the consensus ID. Numerical versions of these confusion matrices are available in the supplementary dataset.

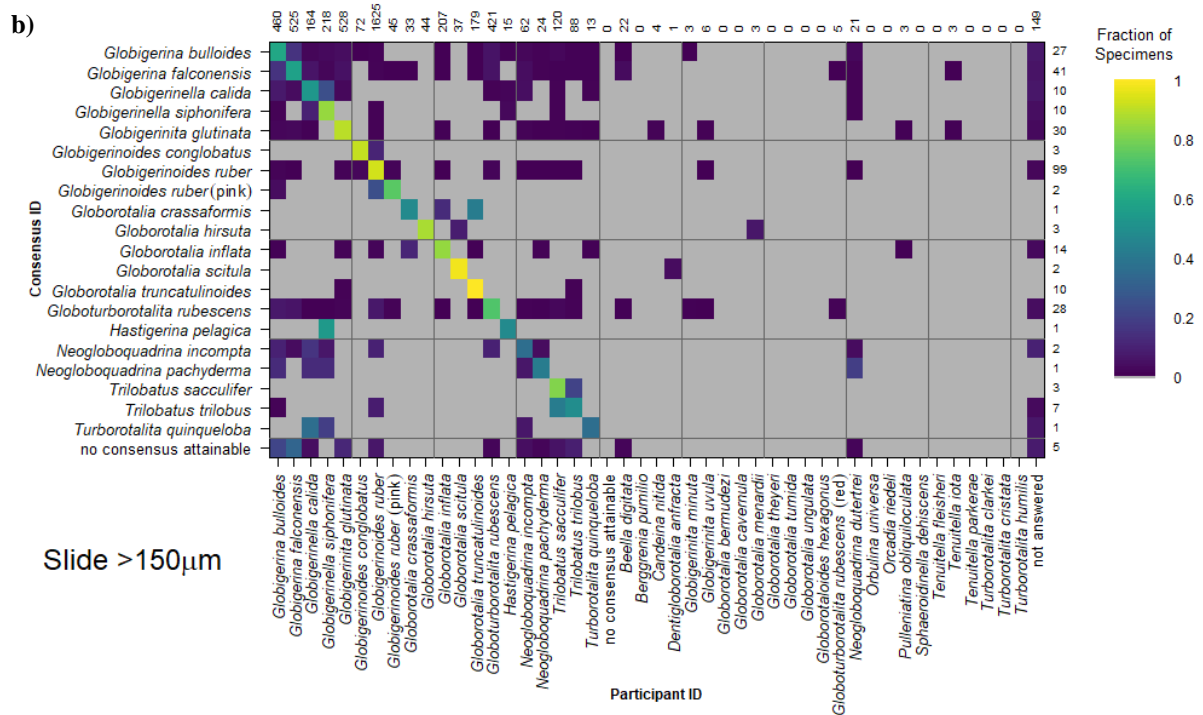
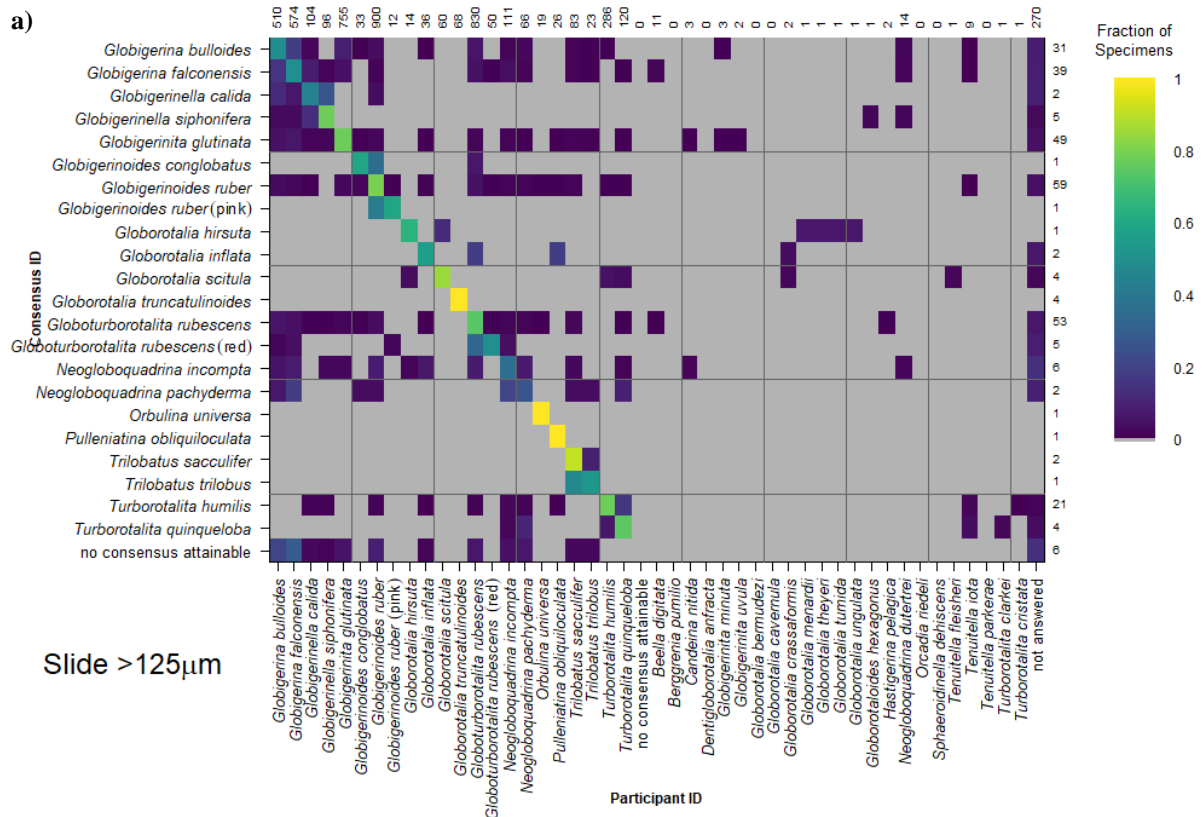


Figure S4: Sensitivity analysis of the method used to split ties in the consensus ID. (a/b) for the slide analysis, (c/d) for the digital analysis. In these confusion matrices ties were classified as “no consensus attainable”. The y-axis shows the consensus ID, and the x-axis shows the names given by the participants to the individual specimens. Where a specimen was always identified correctly, only one square in that row would be filled, indicating a fraction of 1. The numbers along the top indicate the number of specimens given that name in the analysis. The numbers at the right indicate the number of specimens of each species in the consensus ID. Numerical versions of these confusion matrices are available in the supplementary dataset.

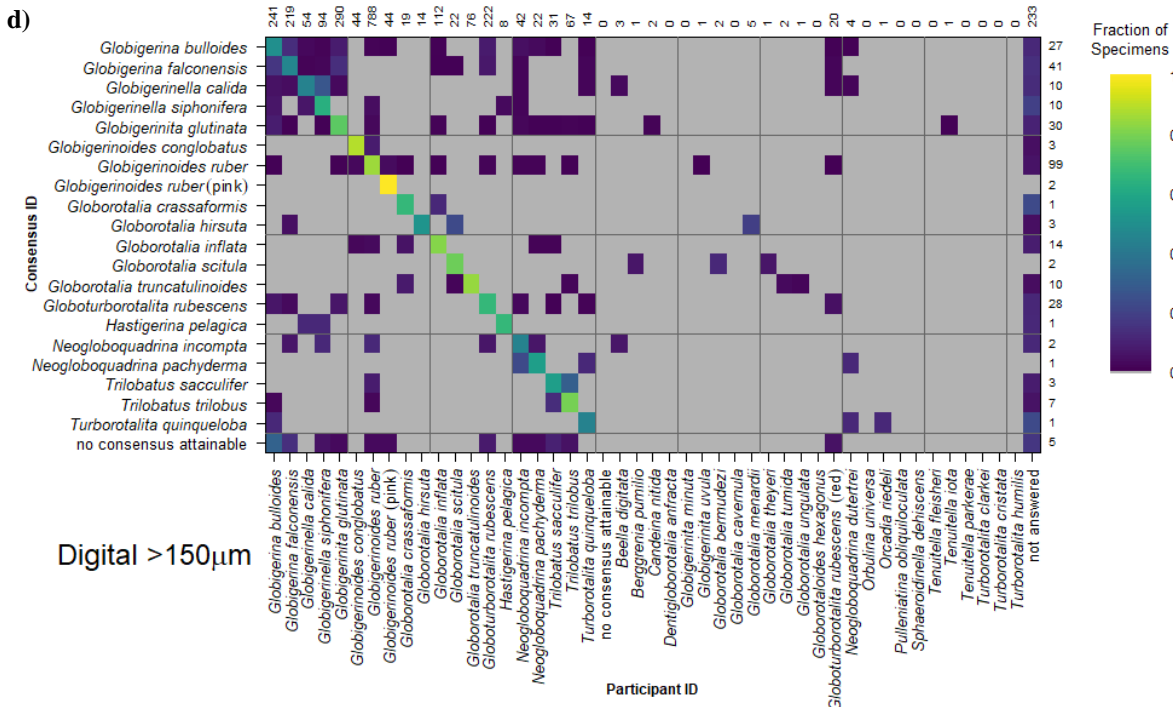
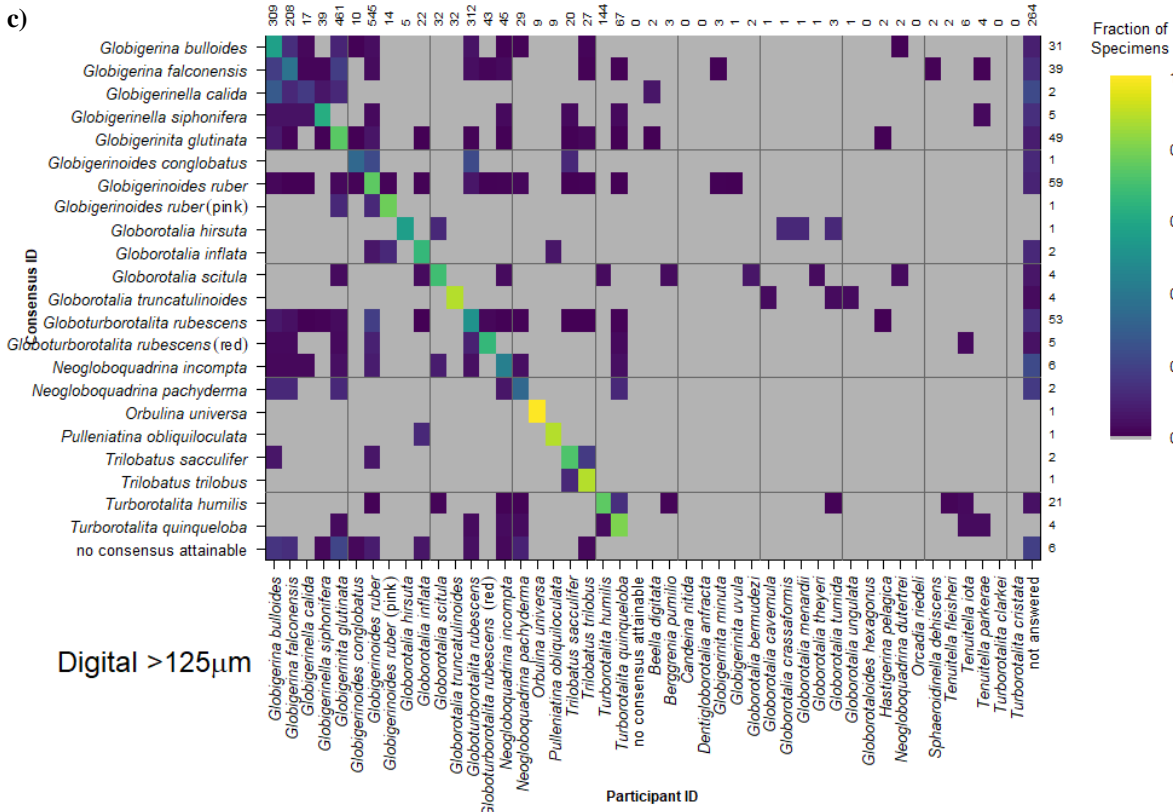


Figure S4 cont.

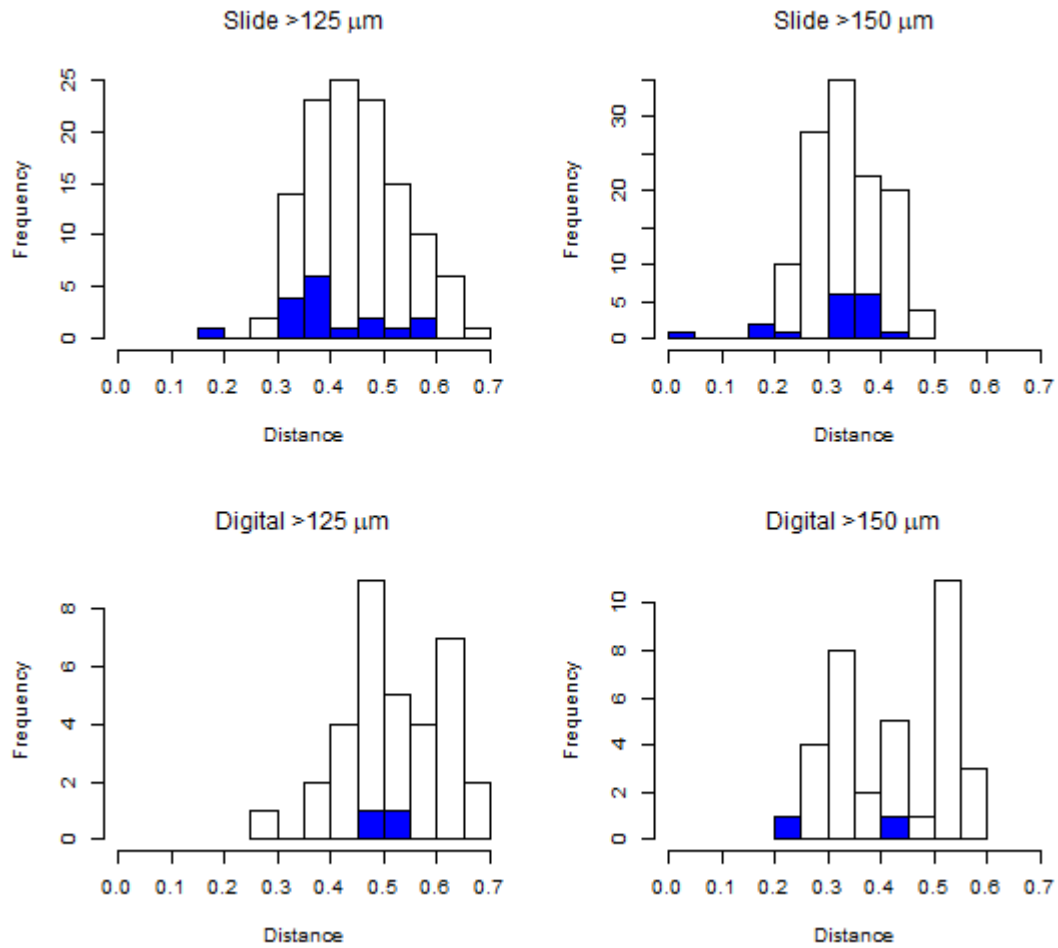


Figure S5: The distances in NMDS space between pairs of points, split by size fraction and slide / digital analyses. Blue bars indicate distances between points within the same school, whilst white bars show the distances for all points not within the same school.

>125 μm

(a) The transfer function temperature reconstructions are calibrated based on the abundances of the >150 μm size fraction. Therefore they were not calculated for the >125 μm size fraction.

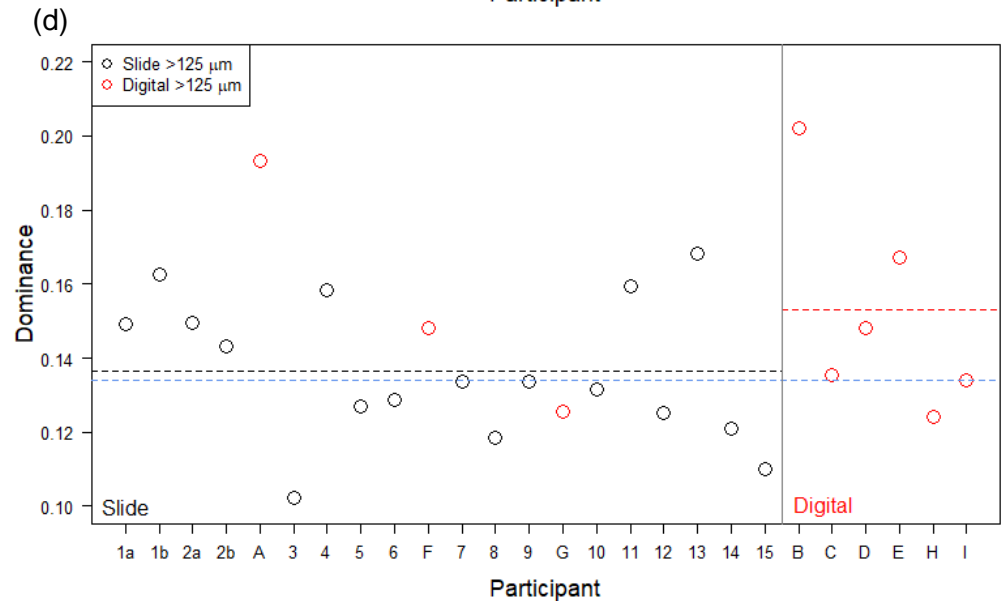
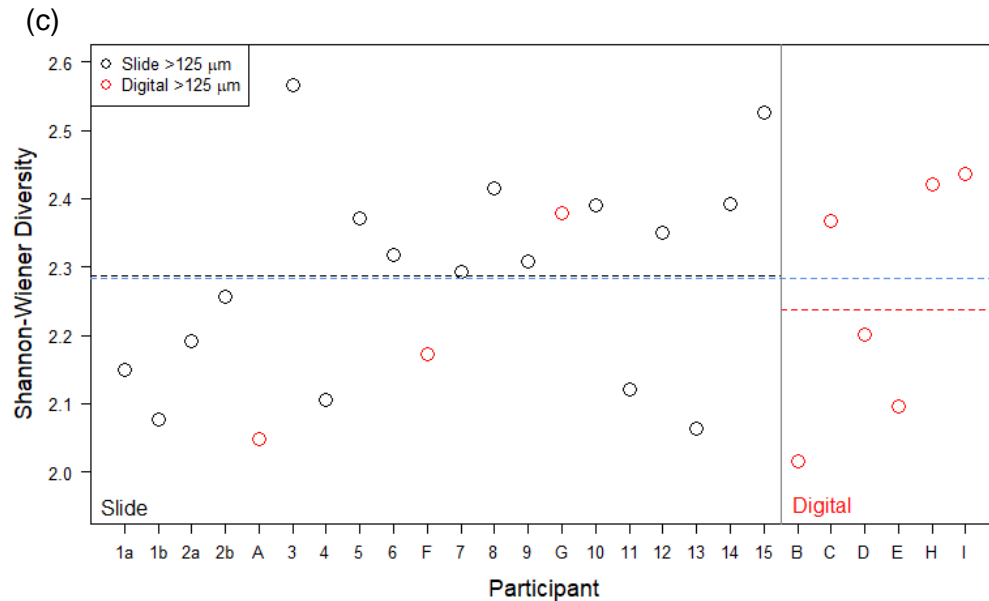
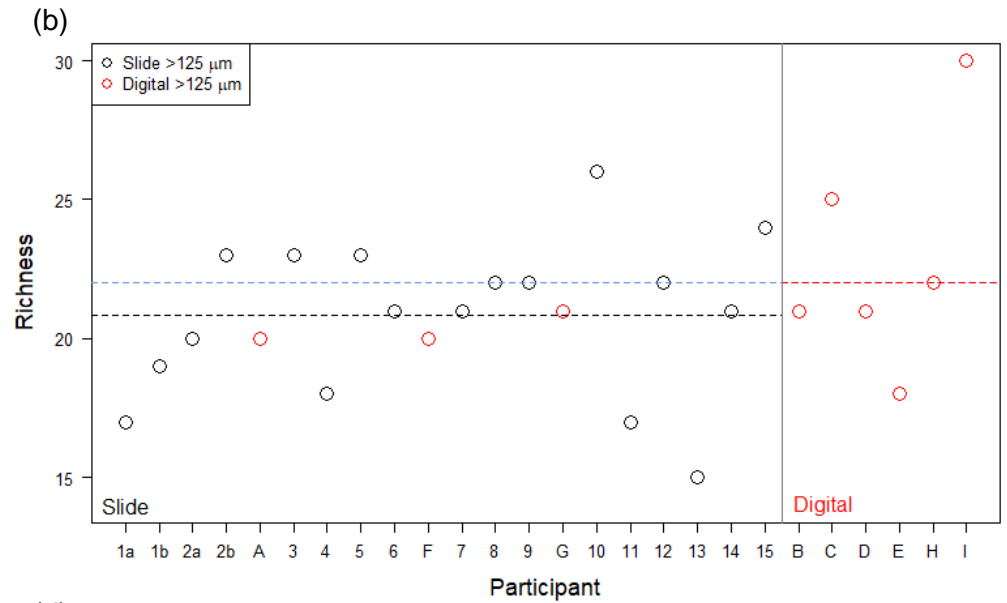


Figure S6: This figure is the equivalent of Figure 6 in the main text, but the >125 μm and >150 μm analyses are split over separate pages. Transfer function temperature reconstructions (a, but not calculated for >125 μm) and community structure (b: Richness, c: Shannon-Wiener diversity, d: Dominance) estimates for assemblage counts from the four analyses. The black / red lines indicate the mean values for the slide / digital analyses and the (thin) blue line is the consensus estimate: solid line >150 μm ; dashed line >125 μm . The order of participants is the same as in Fig. 3. The error bars in (a) show 1 standard deviation among the ten temperature estimates derived from the ANN technique (see Kučera et al., 2005); the 1 standard deviation for the consensus is indicated by shading. The World Ocean Atlas value (thick blue line) is added for reference.

>150 μm

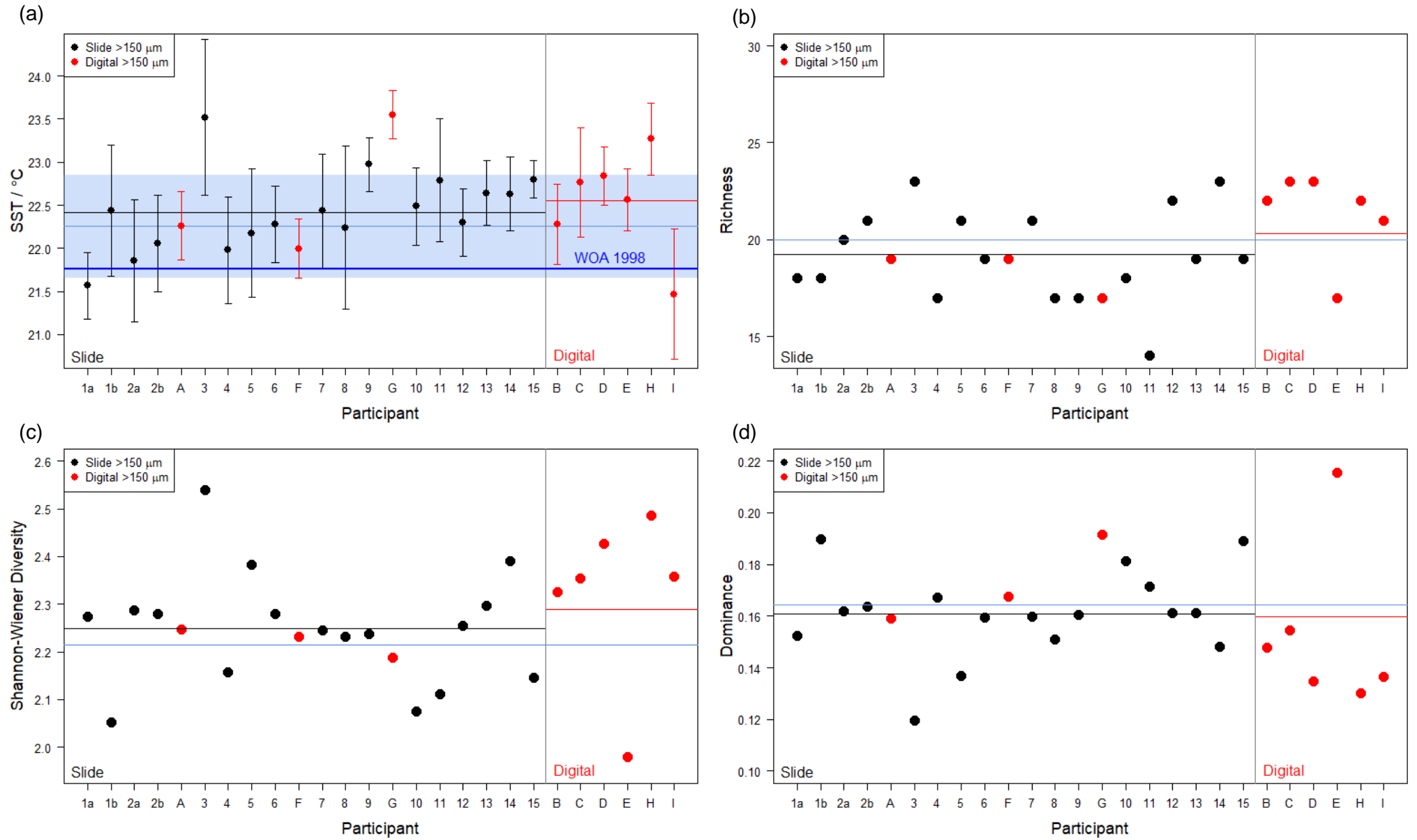


Figure S6 cont.

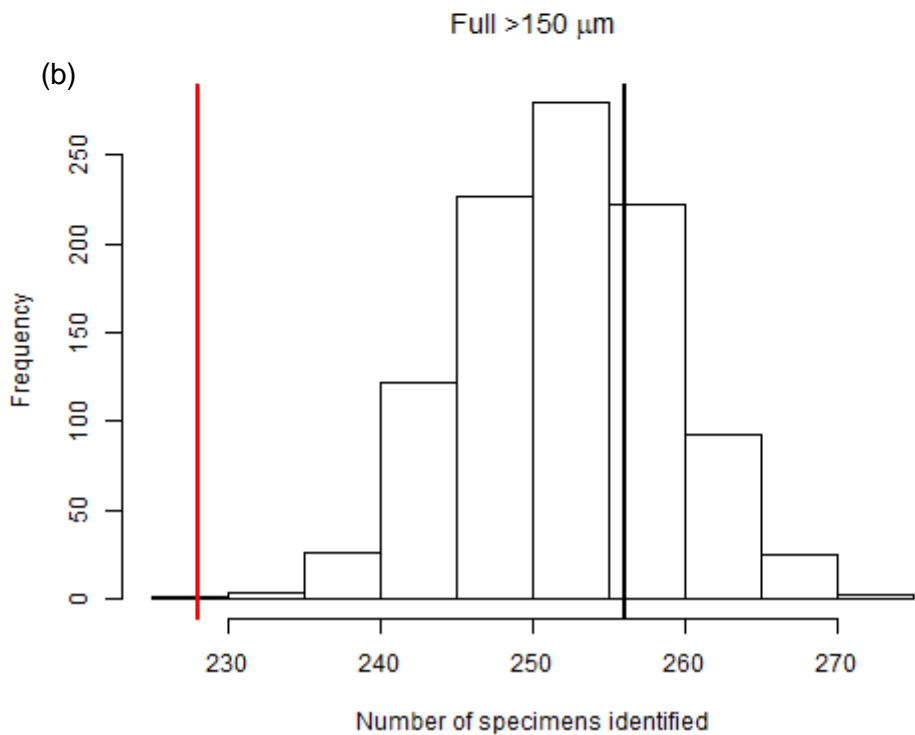
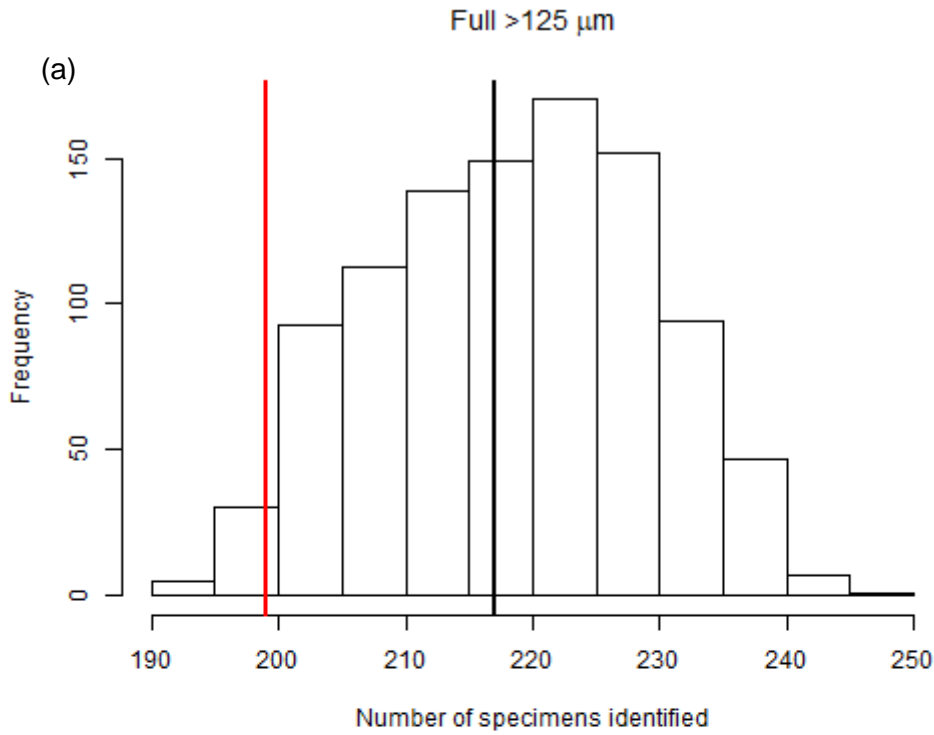


Figure S7: Sensitivity analysis for testing the influence of the number of participants on the levels of agreement in the slide vs. digital separate consensus analyses. These graphs show how the number of specimens identified by the strict consensus (i.e. not classified as “no consensus”) varies between the slide and the digital analyses when the consensus values are obtained separately for these analyses. The black line indicates the number of specimens identified in the full slide analysis; the red line indicates the number in the digital analysis. The histogram shows the estimate of the number of specimens for the 1000 replicates of the subsampled slide analysis when the slide data was subsampled to only contain 9 participants.

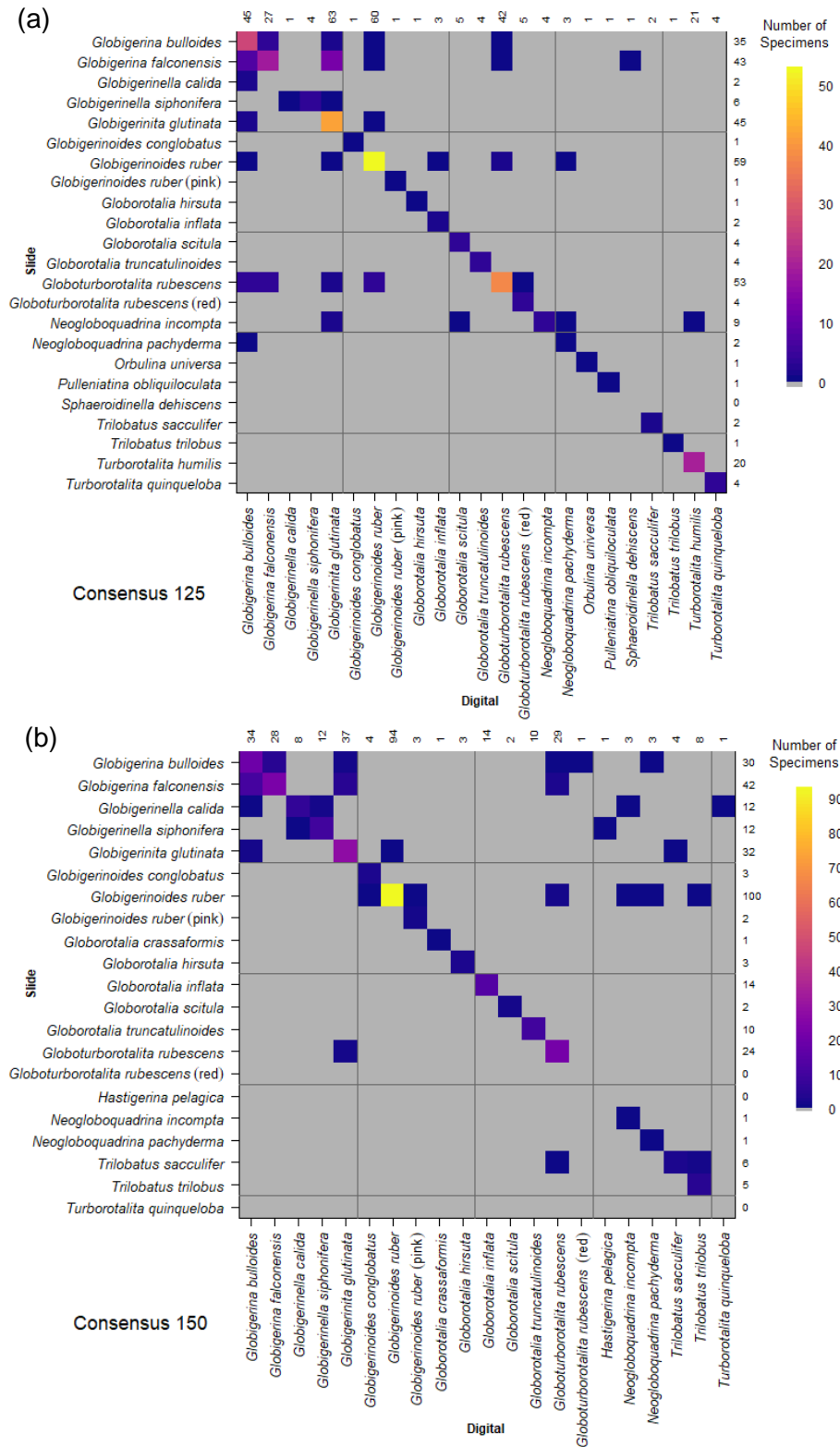
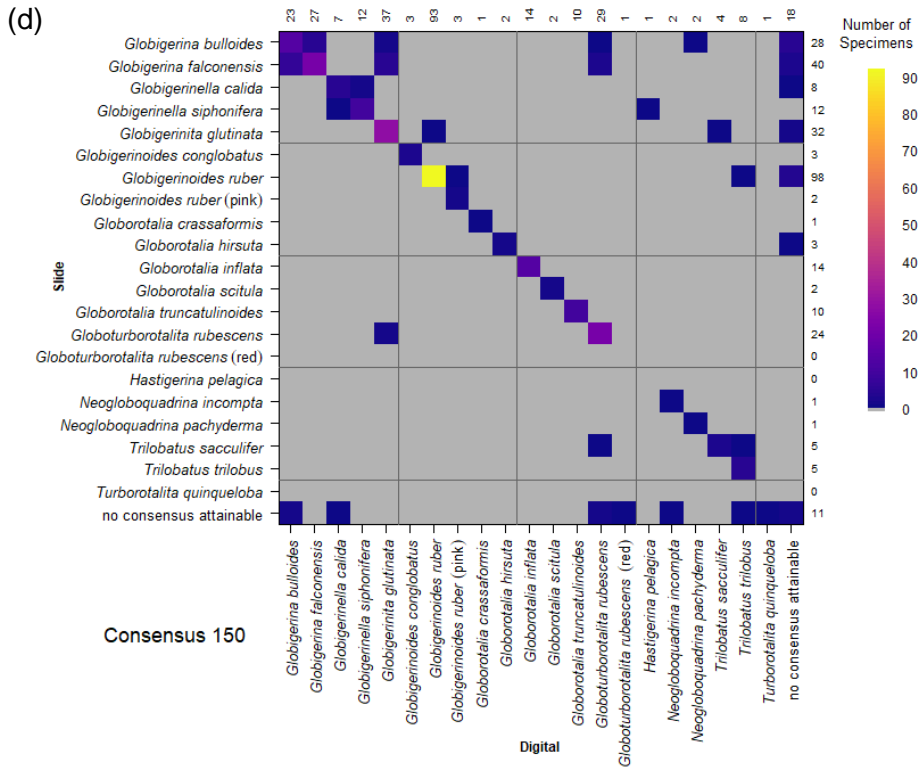
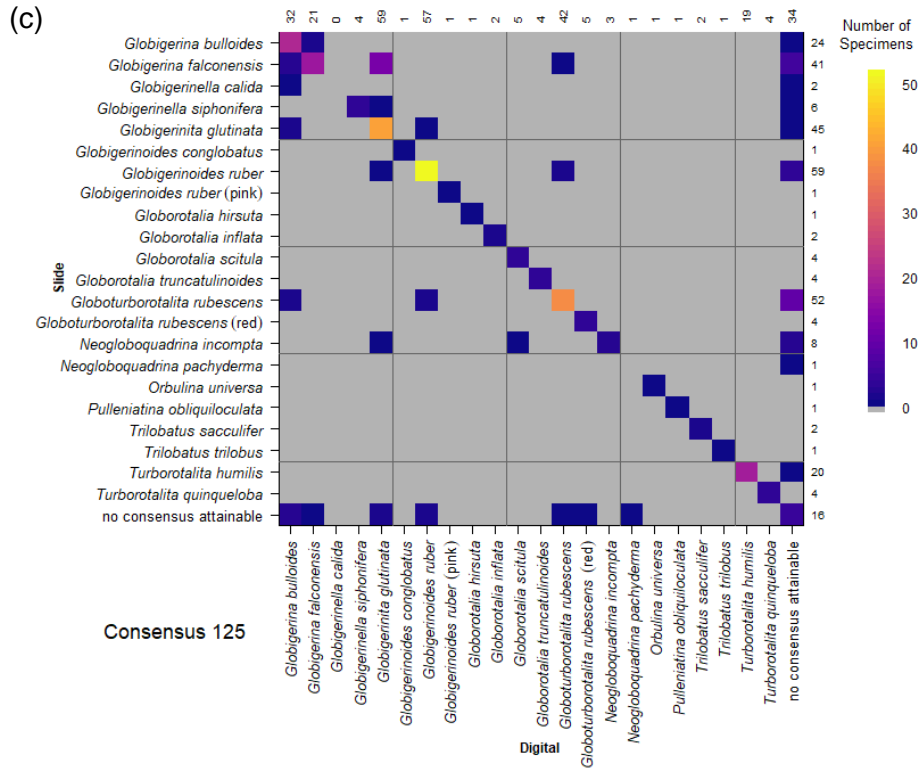


Figure S8: Confusion matrices showing the comparison between the slide and the digital consensus values, if they are calculated separately, for the two size fractions. (a/b) Shows the soft consensus estimates, (c/d) shows the results if ties are designated “no consensus attainable” rather than being split alphabetically. N.b. the colouring on these indicates the number of specimens rather than the fraction correct. The y-axis shows the slide consensus ID, and the x-axis shows the digital consensus ID. Where both agree on all the specimens of a species, only one square in that row/column would be filled. The values indicate the number of specimens of each species in each consensus ID. Numerical versions of these confusion matrices are available in the supplementary dataset.



Section S3: Tables

Table S1: Consistency scores by participant.

Participant	Analysis	Experience / years	Routine	% Agreement >125 μm	% Agreement >150 μm	% Identified >125 μm	% Identified >150 μm
1a	Slide	0.25	Yes	64.00	73.67	99.67	100.00
1b	Slide	1	Yes	75.67	73.00	100.00	100.00
2a	Slide	2	Yes	68.67	85.67	100.00	100.00
2b	Slide	3	Yes	80.00	85.67	100.00	100.00
3	Slide	6	No	57.33	63.00	100.00	100.00
4	Slide	6	No	70.67	82.67	95.67	98.33
5	Slide	8	No	69.00	74.67	100.00	100.00
6	Slide	11	No	70.67	82.00	100.00	100.00
7	Slide	12	Yes	79.33	86.00	100.00	100.00
8	Slide	12	No	71.33	79.33	100.00	100.00
9	Slide	14	Yes	70.67	72.00	79.33	81.67
10	Slide	14	No	65.67	77.67	100.00	100.00
11	Slide	17	No	65.00	75.67	100.00	100.00
12	Slide	18	Yes	86.33	79.67	100.00	100.00
13	Slide	20	Yes	45.00	65.33	48.00	76.00
14	Slide	22	Yes	67.67	79.67	93.33	98.33
15	Slide	39	No	54.00	69.33	94.00	96.00
A	Digital	3	Yes	58.33	80.33	100.00	100.00
B	Digital	3	Yes	58.33	79.33	100.00	99.67
C	Digital	6	Yes	61.00	62.67	100.00	100.00
D	Digital	6	Yes	70.67	77.67	93.33	99.33
E	Digital	7	Yes	40.00	48.00	65.67	70.33
F	Digital	12	No	76.33	81.33	100.00	100.00
G	Digital	14	Yes	53.33	49.33	59.67	56.33
H	Digital	16	Yes	67.33	71.33	93.67	96.67
I	Digital	18	Yes	67.33	75.00	99.67	100.00

Table S2: The diversity metrics and sea surface temperature estimates for each of the participants and the consensus estimates.


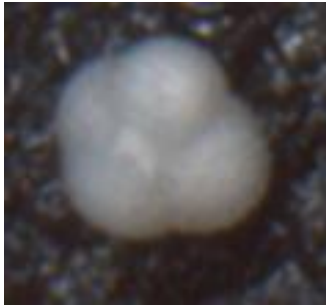

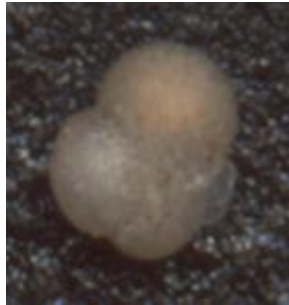

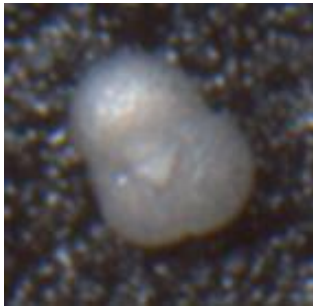

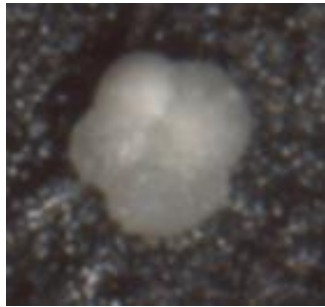


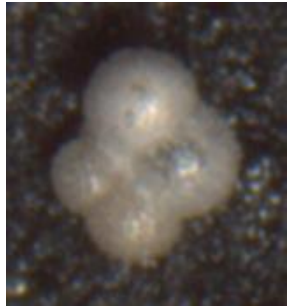
Participant	Analysis	Richness		Shannon Wiener		Dominance		SST 10m / °C	SD SST / °C
		125	150	125	150	125	150		
consensus	-	22	20	2.28	2.21	0.13	0.16	22.25	0.60
1a	Slide	17	18	2.15	2.27	0.15	0.15	21.57	0.38
1b	Slide	19	18	2.08	2.05	0.16	0.19	22.43	0.76
2a	Slide	20	20	2.19	2.29	0.15	0.16	21.85	0.71
2b	Slide	23	21	2.26	2.28	0.14	0.16	22.06	0.56
3	Slide	23	23	2.57	2.54	0.10	0.12	23.52	0.91
4	Slide	18	17	2.10	2.16	0.16	0.17	21.98	0.62
5	Slide	23	21	2.37	2.38	0.13	0.14	22.17	0.75
6	Slide	21	19	2.32	2.28	0.13	0.16	22.28	0.44
7	Slide	21	21	2.29	2.24	0.13	0.16	22.43	0.66
8	Slide	22	17	2.42	2.23	0.12	0.15	22.24	0.94
9	Slide	22	17	2.31	2.24	0.13	0.16	22.97	0.31
10	Slide	26	18	2.39	2.08	0.13	0.18	22.48	0.45
11	Slide	17	14	2.12	2.11	0.16	0.17	22.79	0.71
12	Slide	22	22	2.35	2.25	0.13	0.16	22.30	0.40
13	Slide	15	19	2.06	2.30	0.17	0.16	22.64	0.38
14	Slide	21	23	2.39	2.39	0.12	0.15	22.63	0.43
15	Slide	24	19	2.53	2.14	0.11	0.19	22.80	0.21
A	Digital	20	19	2.05	2.25	0.19	0.16	22.26	0.40
B	Digital	21	22	2.01	2.33	0.20	0.15	22.28	0.47
C	Digital	25	23	2.37	2.35	0.14	0.15	22.76	0.64
D	Digital	21	23	2.20	2.43	0.15	0.13	22.84	0.34
E	Digital	18	17	2.10	1.98	0.17	0.22	22.56	0.36
F	Digital	20	19	2.17	2.23	0.15	0.17	21.99	0.34
G	Digital	21	17	2.38	2.19	0.13	0.19	23.55	0.28
H	Digital	22	22	2.42	2.49	0.12	0.13	23.27	0.42
I	Digital	30	21	2.44	2.36	0.13	0.14	21.47	0.76






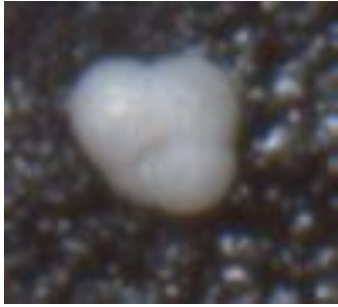








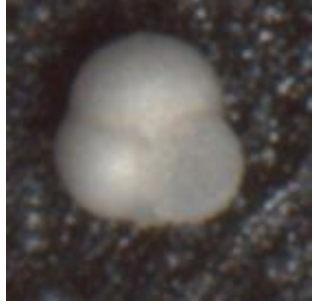
Table S3: The maximum diameter of species based on the specimens, identified by a strict consensus, in this analysis






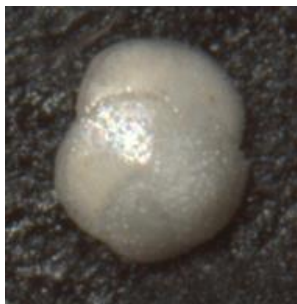
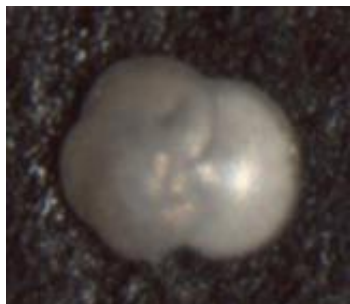
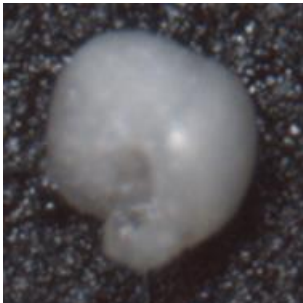



Species	Max. diameter / μm
<i>Neogloboquadrina incompta</i>	164.86
<i>Turborotalita quinqueloba</i>	185.38
<i>Globoturborotalita rubescens</i> (red)	191.28
<i>Turborotalita humilis</i>	239.86
<i>Globorotalia scitula</i>	248.52
<i>Globoturborotalita rubescens</i>	259.64
<i>Globigerina falconensis</i>	333.15
<i>Globigerinita glutinata</i>	362.70
<i>Globigerinoides ruber</i> (pink)	382.11
<i>Globigerinoides ruber</i>	383.6
<i>Globigerina bulloides</i>	390.54
<i>Pulleniatina obliquiloculata</i>	437.85
<i>Globigerinella siphonifera</i>	453.97
<i>Globorotalia inflata</i>	479.12
<i>Trilobatus trilobus</i>	494.72
<i>Globorotalia crassaformis</i>	538.68
<i>Hastigerina pelagica</i>	561.89
<i>Trilobatus sacculifer</i>	564.65
<i>Globigerinella calida</i>	590.02
<i>Globigerinoides conglobatus</i>	598.76
<i>Orbulina universa</i>	612.49
<i>Globorotalia hirsuta</i>	638.76
<i>Globorotalia truncatulinoides</i>	709.96






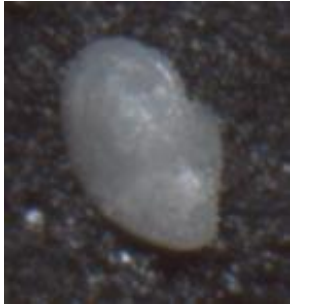


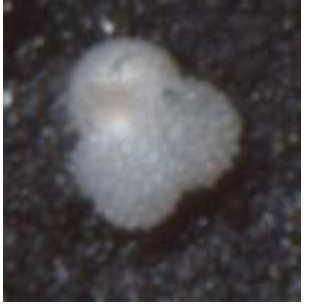


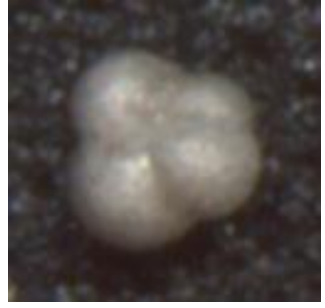


Section S4: Digital examples by species




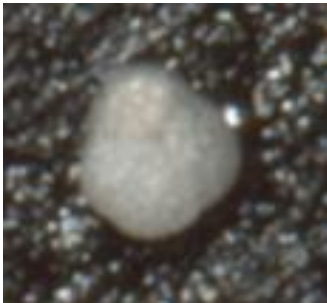


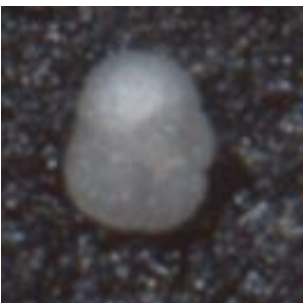
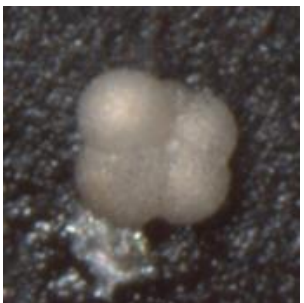

Example images of specimens with the highest and lowest agreement for each species identified in the consensus, for each sample size. The maximum / minimum consensus values indicate the number of participants (out of 26) who agreed on the identification of that specimen.






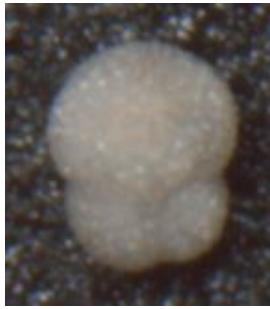



Species name	>125 μm				>150 μm			
	Maximum consensus		Minimum consensus		Maximum consensus		Minimum consensus	
<i>Globigerina bulloides</i>	24		5		25		5	
<i>Globigerina falconensis</i>	19		5		22		7	
<i>Globigerinella calida</i>	9		-	-	18		6	

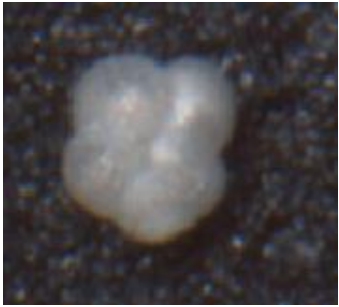
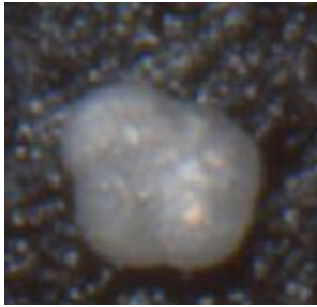
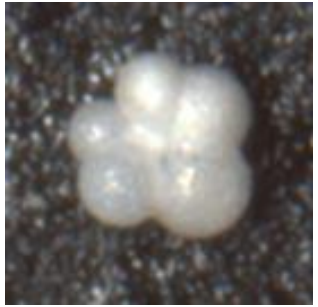
<i>Globigerinella siphonifera</i>	26		11		24		15	
<i>Globigerinita glutinata</i>	26		6		26		7	
<i>Globigerinoides conglobatus</i>	13		-	-	25		22	
<i>Globigerinoides ruber</i>	26		5		26		12	

<i>Globigerinoides ruber</i> (pink)	17		-	-	22		21	
<i>Globorotalia crassaformis</i>	-	-	-	-	14		-	-
<i>Globorotalia hirsuta</i>	16		-	-	21		17	
<i>Globorotalia inflata</i>	18		13		26		14	

<i>Globorotalia scitula</i>	24		16		24		23	
<i>Globorotalia truncatulinoides</i>	26		23		26		22	
<i>Globoturborotalita rubescens</i> inc. <i>G. tenella</i>	24		5		26		7	
<i>Globoturborotalita rubescens</i> (red)	17		10		-	-	-	-

<i>Hastigerina pelagica</i>	-	-	-	-	14		-	-
<i>Neogloboquadrina incompta</i>	16		7		12		8	
<i>Neogloboquadrina pachyderma</i>	8		5		12		-	-
<i>Orbulina universa</i>	26		-	-	-	-	-	-

<i>Pulleniatina obliquiloculata</i>	25		-	-	-	-	-	-
<i>Trilobatus sacculifer</i>	22		-	-	19		18	
<i>Trilobatus trilobus</i>	17		-	-	18		13	
<i>Turborotalita humilis</i>	23		6		-	-	-	-

<i>Turborotalita quineloba</i>	26		12		10		-	-
------------------------------------	----	-----------------------------------------------------------------------------------	----	------------------------------------------------------------------------------------	----	-------------------------------------------------------------------------------------	---	---