

Supplementary Material

1 Supplementary Data

OTU definition at a threshold of 3% tended to separate the OTUs by location (North ATZ, South ATZ, and North BUS) and by O₂ (oxic, oxycline, OMZ and suboxic) (Fig. 3) which indicates differences based on regional and O₂ features within South/North ATZ and North BUS. Combining all sequences (958), we detected 82 OTUs, which was less when analyzing each region (94) and O₂ feature (103) separately (Table S2). Thus, there was some overlap between region/feature (12 OTUs between regions and 21 OTUs between O₂ feature) (Table S2). Among the 15 most abundant OTUs (>ten sequences), six OTUs (1, 4, 9, 10, 11 and 12) contained sequences from more than one region (>5% sequences from each region). OTUs 3, 5, 6, 7, 8, 13, 14 and 15 contained fewer or no sequences (<5%) from another region (Fig. 3A, Table S4). Grouping by O₂ feature we see an overlap between sequences (Fig. 3B). The overlap seems to be within the North BUS itself perhaps because it is a dynamic shallow system. Still, we see a general trend of separation between OMZ/suboxic and oxic/oxycline features (Fig. 3B). This pattern of overlap between regions suggests a dominance of *nifH* assemblages rather than evenness, which also has been observed and reported before (Cheung et al., 2016; Jayakumar and Ward, 2020). Based on this pattern of dominance, deeper sequencing would most likely not change the observed OTUs dominance but rather lead to the discovery of rare OTUs.

Rarefaction curves (Fig. S1) indicate that we did not achieve sequencing saturation for either region, oxygen range or all sequences grouped. Moreover, Chao statistics (Table S2) reveal that there remains much diversity to be explored. Based on the shape of the rarefaction curves and the diversity indices (Table S2 and Fig. S1) we can see that the greatest *nifH* diversity was present in waters off the North BUS (Simpson: 0.05, Shannon: 3.4, Chao: 98, ACE: 115) compared to North ATZ (Simpson: 0.28, Shannon: 1.66, Chao: 13, ACE: 15) and South (Simpson: 0.4, Shannon: 1.2, Chao: 16, ACE: 23). The number of sequences obtained from Namibia (505) compared to Congo (133) and Angola (320) most likely play a part in the observed statistic. Although different numbers of sequences were obtained, the OTUs per number of sequences were analyzed (Namibia: 0.13, Congo: 0.09, Angola: 0.04). This combined with statistics (Table S2) could indicate we did not reach saturation for North ATZ and North BUS and may be limited by the low sequencing depth associated with the clone library. In comparison, South ATZ seems to be strikingly less diverse compared to North BUS and North ATZ, despite the low OTU number from North ATZ (OTUs: 12) but high Shannon and low Simpson (Table S2). Grouping sequences based on the O₂ feature we have had a high number of sequences from the oxic (435) and oxycline (295) compared to OMZ (159) and suboxic (69) (Table S2). Grouping OMZ and suboxic resulted in 228 sequences with only three OTUs overlap (Table S2). The greatest *nifH* diversity occurred in the OMZ/suboxic waters (Simpson: 0.08, Shannon: 2.95, Chao: 71, ACE: 77) compared to oxic (Simpson: 0.34, Shannon: 1.68, Chao: 29, ACE: 31) and Oxycline (Simpson: 0.08, Shannon: 2.92, Chao: 42, ACE: 43) (Table S2). Roughly 36% of the diversity was in singletons (Table S3 and S4), not shown in the 15 most abundant OTUs

(>ten sequences), which represent 760 sequences (79%) of the total *nifH* sequences analyzed (958 seq). Most of the diversity was observed from the North BUS coast, oxycline and OMZ waters (Fig. S1, Fig. S2).

1.1 Supplementary Tables

Table S1 Sampling overview of individual stations used for molecular experiment.

<i>Region</i>	<i>CTD</i>	<i>Lon (°E)</i>	<i>Lat (°S)</i>	<i>Depth (m)</i>	<i>Oxygen (μM)</i>
<i>Waters off Angola</i>	42	7.0001	-11.746	97	42.2097
				57.7	38.75898
				17.4	216.3581
	46	11.634	-11.75	98.7	65.37882
				57.1	43.50372
				9.5	222.592
	50	12.5	-11.166	97.2	50.72353
				47.4	69.89762
				34.1	147.7237
				9.1	214.8073
	80	13.15	-10.734	689.1	77.49742
				337.7	22.56319
				143.8	44.30478
				47.7	69.1171
				18.3	234.9776
	74	11.5	-6.402	343.2	39.52923

<i>Waters off Congo</i>				87.6	108.3793
				13	206.0881
	76	11.251	-6.483	336	27.77008
				29.7	143.4924
				9.2	211.5004
<i>Waters off South Angola</i>	87	11.42	-15.033	591.9	81.25624
				254.5	22.11131
				97.6	39.68328
				47.3	61.48649
				21.8	127.2658
				9	219.8294
	92	11.838	-15.256	82.6	42.13781
				52.3	48.42305
				13.3	154.8613
<i>Waters off Namibia</i>	98	11.616	-17.499	92.5	35.11313
				37.3	101.2211
				7.6	103.2238
	99	12.284	-19	100	23.02534
				57.9	35.62663
			37.6	78.28821	
			7.8	142.9892	

101	13.316	-21	92.5	6.70631
			47.4	8.57545
			27.5	27.74954
			17.3	123.8357
102	13.783	-21.999	95.8	3.32748
			71	10.44459
			37	158.2402
			2.3	209.4567
103	14.216	-22.665	87.1	2.72155
			28.1	100.4611
			12.2	215.9062

Table S2 Summary of OTU statistics for each location and feature. Shannon and Simpson indicate the richness and diversity while Chao and Ace are nonparametric estimations of the total number of OTU in the original sample. * Water off Angola includes both the southern and the northern part of Angola. ** ACE stands for Abundance-based Coverage Estimator.

<i>Feature</i>	<i>No. seq</i>	<i>No. Unique seq</i>	<i>No. of OTU</i>	<i>OTU/seq</i>	<i>Shannon index</i>	<i>Simpson index</i>	<i>Chao1 index</i>	<i>ACE**</i>
<i>South ATZ (Waters off Angola*)</i>	320	162	13	0.04	1.22	0.41	16	23
<i>North ATZ (Waters off Congo)</i>	133	66	12	0.09	1.66	0.28	13	15
<i>nBUS (Waters off Namibia)</i>	505	275	69	0.13	3.42	0.05	98	115

<i>ATZ and nBUS</i>	958	438	82	0.09	3.05	0.11	115	134
<i>Oxic (>100 $\mu\text{M O}_2$)</i>	435	223	23	0.05	1.68	0.32	29	31
<i>Oxycline (20-100 $\mu\text{M O}_2$)</i>	295	157	36	0.12	2.92	0.08	42	43
<i>OMZ (<5 $\mu\text{M O}_2$)</i>	159	95	33	0.21	2.90	0.08	49	46
<i>Suboxic (5-20 $\mu\text{M O}_2$)</i>	69	37	11	0.16	1.42	0.37	14	16
<i>OMZ + Suboxic conditions</i>	228	127	41	0.18	2.95	0.08	71	77

1.2 Supplementary Figures

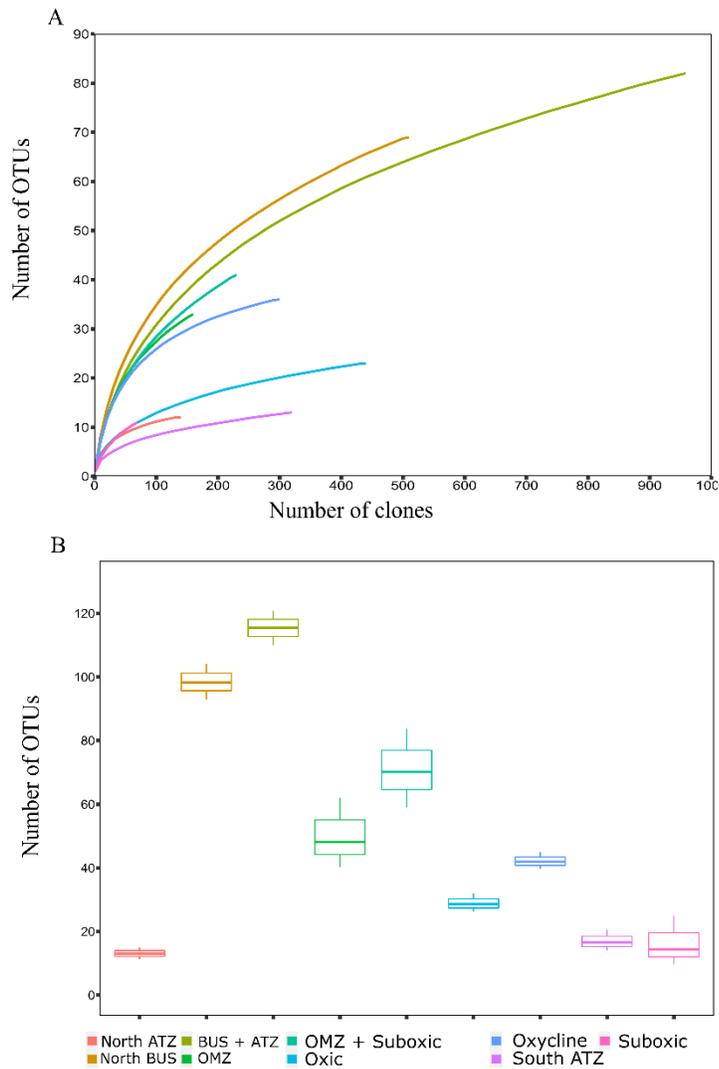


Figure S1 Rarefaction curve showing observed OTU richness versus the number of clones sequenced, from each location and O₂ feature (A). Chao estimates (number of species in the community) for each location/feature (B) indicating that more diversity remains to be sequenced. ATZ stands for Angola tropical zone and BUS for Benguela Upwelling system. “BUS + ATZ” is a compilation of all sequences (958) from this study.

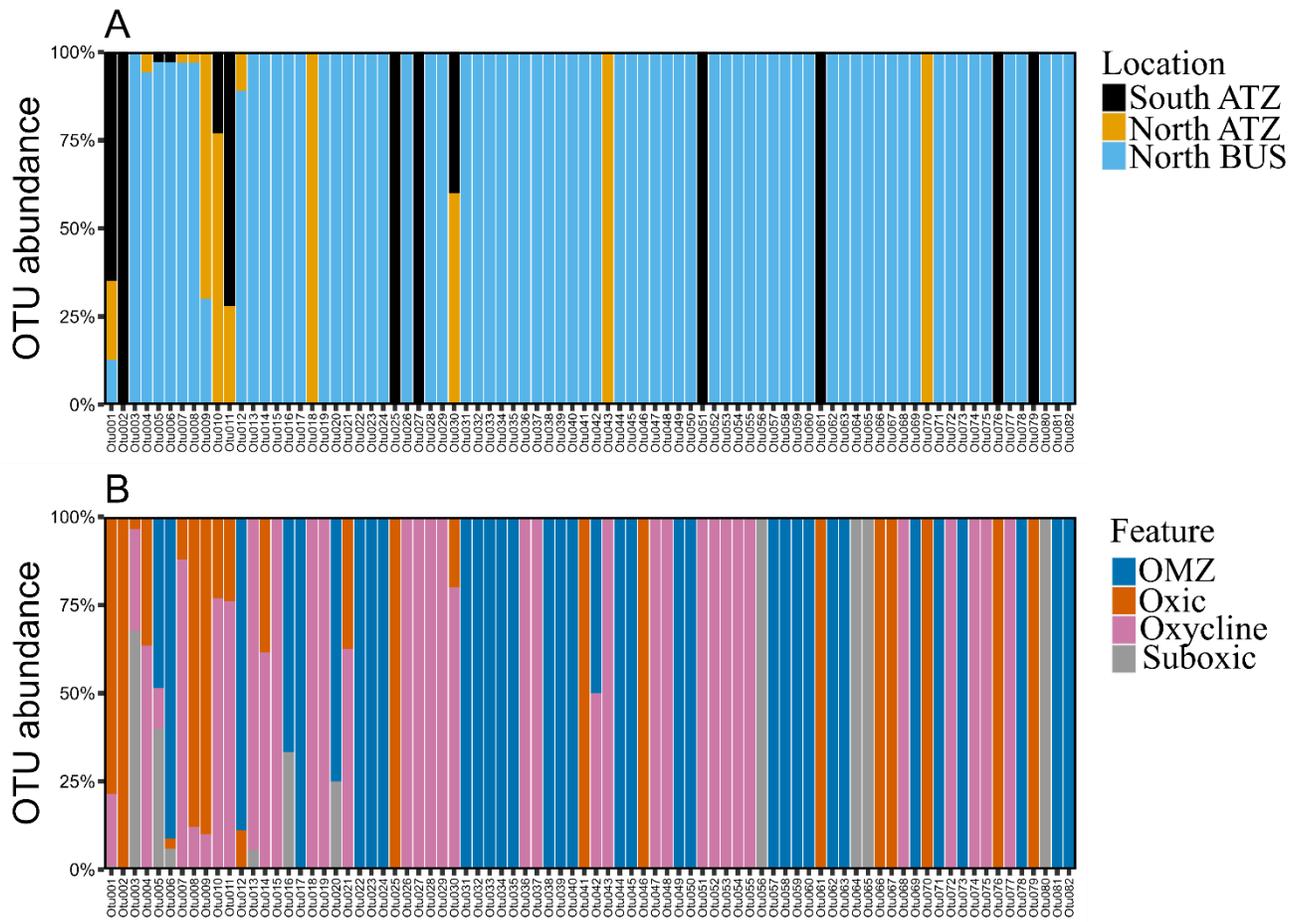


Figure S2 Histogram of all OTUs identified according to location (A) and O₂ feature (B). OTU were defined at a threshold of 3% nucleotide differences and is in descending order, OTU1 being the most abundant.

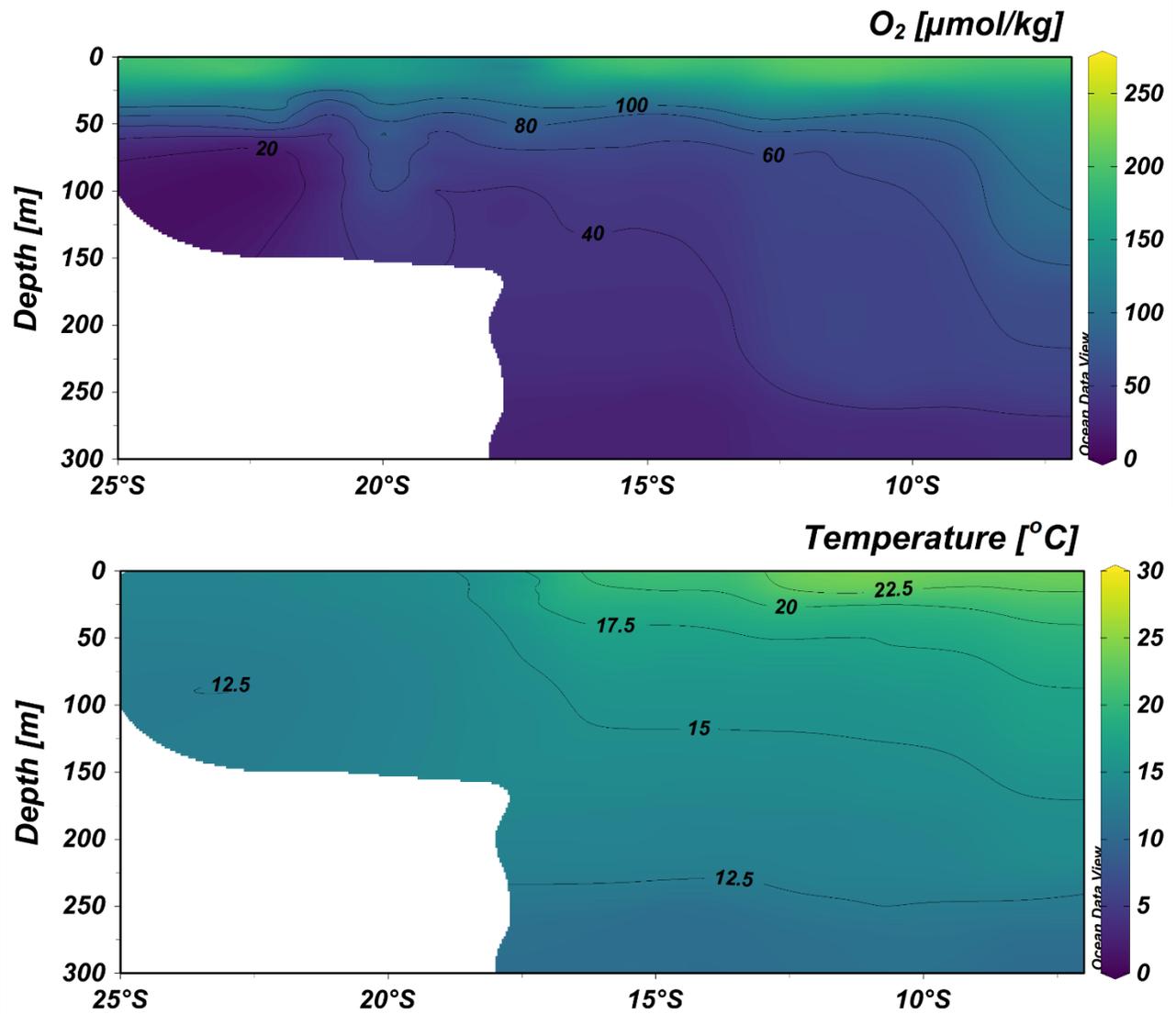


Figure S3 Top figure display O_2 concentration ($\mu\text{mol kg}^{-1}$) in the Northern Benguela upwelling system (17-25°S) and from the Angola tropical zone (14-6°S). Bottom figure showed temperature.