Supplemental Material *for Ideas and Perspectives: Value and efficiency of ocean sample return and laboratory analysis: Key analytes are orders of magnitude rarer than a needle in a haystack by Mak A. Saito, Matthew R. McIlvin, Annette F. Govindarajan, John A. Breier*

Supplemental Methods

Estimate of the relative mass of metal needles versus a haystack

Modern hand-sewing needles (Singer) of assorted sizes were weighed on an analytical balance and their length measured with a metric ruler. Five sizes of needles were present, each of the five sizes was averaged and then those five groups were averaged to yield an average needle weight of 0.1095 $\pm$ 0.0020 g (Supplemental Table 1).

The weight of a haystack was obtained from the Omeo Standard and Mining Gazette (Anonymous, 1914) 500 cu ft ton^{-1} of hay, and the volume of a haystack was assumed to be 500 cu ft, having dimensions of 10'x10'x5' for this calculation. One imperial ton is equivalent to 1016 kg. The ratio of a needle to haystack is then 1.07 e-7 (0.1095 g / 1016000 g) or 1 in 10 million.

Comparison of selected biogeochemical analytes versus needle within haystack

Analyte concentrations were converted to a mass basis, where the molar concentrations (mol L^{-1}) were converted to mass concentrations (g L^{-1}) using their respective atomic weights, then converted to parts per million (ppm), using a value of 1.03 kg per 1 L of seawater. Analytes were then compared (by calculating the ratio to) with the mass fraction of a needle in a haystack, calculated above, and the inverse also provided to yield scarcity relative to the needle in a haystack (Table S2). Concentrations of macronutrients were obtained from the WOCE dataset. Example nitrite and nitrous oxide values were obtained from Saito et al. (2020). DMSP, biotin and methionine values were obtained from Johnson et al. (2023).

For example, the concentration of dissolved iron in seawater is 0.15 ± 0.08 nM (n = 141) in the Pacific Sector of Antarctica, an iron-limited region (Tagliabue et al., 2012). Using the atomic weight of iron, and a seawater density of 1.03 kg L^{-1}, this yields 8.13×10^{-6} ppm by mass. Relative to the needle-in-haystack abundance of 0.1078 ppm, dissolved iron is therefore approximately 13,300-fold scarcer by mass than a needle in a haystack.

Total protein calculations (Table S2) were based on vertical profiles collected on the METZYME expedition using biomass in 0.2 – 3.0 mm size range (Saito et al., 2020; their Figure S12), using 9 and 2 mg L^{-1} for euphotic and mesopelagic zones, respectively.

Bacterioplankton mRNA abundance was estimated using measured concentration of mRNA molecules in coastal seawater (1e12 molecules L^{-1}, converted to 1.66×10^{-12} mol L^{-1} using Avogadro's number, value from Gifford et al., 2011; internal standard method). A transcript length of 924 nucleotides was used to convert to mass concentration (as used by Gifford et al., 2011) resulting in an approximate molecular weight of 305,000 g mol^{-1} using 330 g mol^{-1} $nucleotide^{-1}$, and converted to ppm using a seawater density of 1.03 kg L^{-1}. This is an approximate seawater value, recognizing that transcript lengths vary and can include transcription of larger gene neighborhoods.

The eDNA calculation used an estimate of 50,000 copies of DNA per liter of seawater (Salter et al., 2019), converted to molarity using Avogadro's number. Salter et al. targeted mitochondrial DNA, and a 1:1 copy number is assumed here, so the mass of mitochondrial genome was estimated using a 16,500 bp mitochondrial genome, yielding a molecular weight of 1.09e7 g mol^{-1} based on a 660 g mol^{-1} bp^{-1}. This was followed by a unit conversion to ppm using a seawater density of 1.03 kg L^{-1} (Equation 1).

(50000 copies / L seawater) x 1 mol / 6.022e23 x (1 mitochondrial genome / 1 gene copy) x (1.09e7 g / mol genome) x 1000 mg / 1 g x (1 L / 1.03 kg) x (1ppm / (1 mg / kg)) = 8.79e-10 ppm (1)

Table S1. Measurement of needles used for needle-in-a-haystack calculation. Needles were grouped by length. Average, standard deviation, and relative standard deviation were calculated for each size grouping. Overall average, used for calculations in Figure 1 and Table S2, represents the mean needle mass across five size classes, and the standard deviation across those size-class means.

Needle Group	Length (cm)	Weight (g)	Average	Standard Deviation	Relative Standard Deviation (%)
1	3.3	0.0556	0.05525	0.00049	0.90
1	3.3	0.0549			
2	3.55	0.0712	0.07170	0.00095	1.33
2	3.65	0.0728			
2	3.5	0.0711			
3	3.85	0.0783	0.08096	0.00393	4.85
3	3.85	0.078			
3	3.85	0.0766			
3	3.85	0.0786			
3	3.85	0.086			
3	3.9	0.085			
3	3.9	0.0842			
4	4.25	0.1178	0.11745	0.00237	2.02
4	4.1	0.1142			
4	4.1	0.1179			
4	4.2	0.1199			
5	5.15	0.2204	0.22225	0.00527	2.37
5	5.1	0.2231			
5	5.1	0.2167			
5	5.1	0.2283			
5	5.1	0.2167			
5	5.1	0.2283			
		Overall Average	0.1095	0.0020	2.29

Table S2. Abundance relative to a needle in a haystack for ocean analytes. See supplemental methods for data sources and explanation of calculations.

Analyte	Example Concentration (M)	Molecular/atomic weight (g mol^{-1})	Mass concentration (g L^{-1}; g kg^{-1} for haystack)	ppm	Needle-in-haystack equivalents	Times harder than finding a needle in a haystack
Water	55.5	18.015	999.83	970711	9004742	0.0000001
Salinity	NA	NA	35	35000	324675	0.0000031
Chloride	0.546	35.453	19.3573	18794	174337	0.0000057
Sodium	0.468	22.989	10.7589	10445	96897	0.0000103
Sulfate	0.028	96.06	2.6897	2611	24224	0.0000413
Magnesium	0.053	24.305	1.2882	1251	11602	0.0000862
Calcium	0.0103	40.078	0.4128	400.8	3718	0.000269
Potassium	0.0102	39.098	0.3988	387.2	3592	0.000278
Bicarbonate	0.00172	61.0168	0.1048	101.8	944	0.001059
Silicic acid, deep	0.00017	96.11	0.0163	15.9	147	0.007
Nitrate (deep)	0.000043	62.005	0.0027	2.589	24	0.042
Dissolved oxygen	0.000020	31.998	0.00064	0.621	5.8	0.174
Dissolved carbon dioxide	0.0000096	44.01	0.00042	0.410	3.8	0.263
Phosphate, deep	0.000003	94.9714	0.00029	0.286	2.7	0.377
Needle in haystack	NA	NA	1.078E-04	0.1078	1	1
Glycine betaine	3.90E-07	117.148	4.56877E-05	4.44E-02	0.41148	2
Total Euphotic Zone Protein	NA	NA	9.00E-06	0.009	0.08106	12.3
Nitrite	1.00E-07	46.005	4.6005E-06	4.47E-03	0.04143	24
Nitrous oxide	8.00E-08	44.013	3.52104E-06	3.42E-03	0.03171	32
Total Mesopelagic Protein	NA	NA	2.00E-06	0.002	0.01801	55.5
Dissolved zinc (deep)	1.00E-08	65.38	6.538E-07	6.35E-04	0.00589	170
Bacterioplankton mRNA	1.66E-12	3.05E+05	5.06E-07	4.92E-04	0.00456	219
Dimethylsulfoniopropionate (DMSP)	1.20E-09	134.2	1.61E-07	1.56E-04	0.00145	689
Dissolved cadmium (deep)	1.00E-09	112.411	1.12E-07	1.09E-04	0.00101	988
Dissolved zinc (surface)	2.00E-10	65.38	1.31E-08	1.27E-05	0.00012	8491
Dissolved manganese (deep)	2.00E-10	54.94	1.10E-08	1.07E-05	9.90E-05	10105
Dissolved iron (surface)	1.50E-10	55.845	8.38E-09	8.13E-06	7.54E-05	13255
Biotin (vitamin B7)	1.30E-11	244.31	3.18E-09	3.08E-06	2.86E-05	34960
Dissolved cobalt (deep)	5.00E-11	58.93	2.95E-09	2.86E-06	2.65E-05	37683
Methionine	1.30E-11	149.21	1.94E-09	1.88E-06	1.75E-05	57242
Vitamin B12	5.00E-13	1355.38	6.78E-10	6.58E-07	6.10E-06	163842
Tryptophan	3.00E-12	204.23	6.13E-10	5.95E-07	5.52E-06	181224
Pyridoxine (B6)	1.00E-12	169.18	1.69E-10	1.64E-07	1.52E-06	656307
Atlantic Cod eDNA	8.30E-20	1.09E+07	9.05E-13	8.79E-10	8.15E-09	122687477