Variations in pelagic bacterial communities in the North Atlantic Ocean coincide with water bodies

Richard L. Hahnke[1], Christina Probian[1], Bernhard M. Fuchs[2], Jens Harder[1,*]

[1]Department of Microbiology and [2]Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Celsiusstr. 1, 28359 Bremen, Germany

*Corresponding author. Email: jharder@mpi-bremen.de

Supplement 1. The supplement provides a map of the sampling area, the density of the samples in different water depths, *k*-dominance curves of T-RFLP (terminal restriction fragment length polymorphism) patterns, graphs of nMDS (non-metric multidimensional scaling) and PCA (principal component) analyses, and a documentation of TRFs (terminal restriction fragments) relative fluorescence intensity along the sampling transect
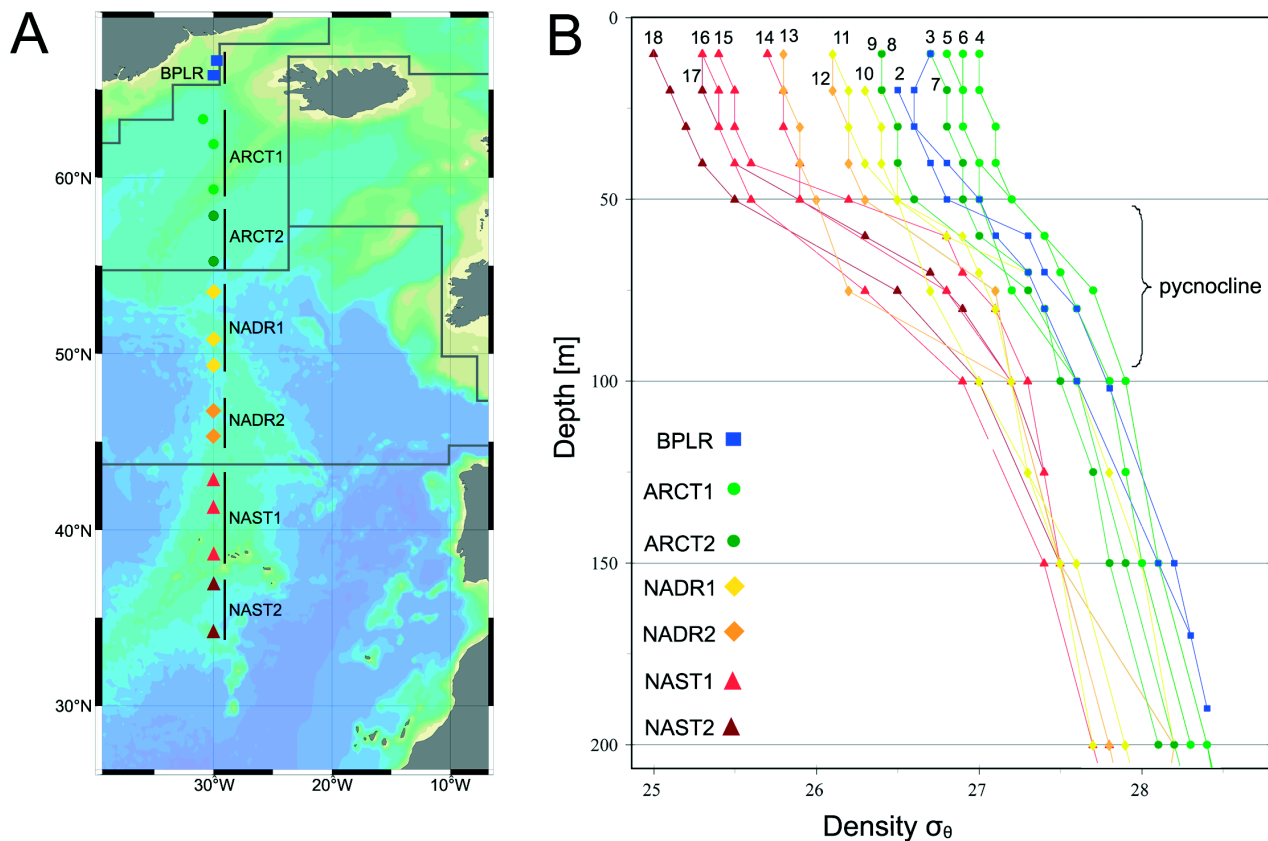

Fig. S1. Sampling stations and water density of the RV 'Maria S. Merian' cruise 03/1 VISION in 2006. (A) Water samples were obtained from the North Atlantic Ocean along the 30° W meridian from the productive cold Greenland current (Boreal Polar [BPLR], 66° 39' N) across the cold north (Atlantic Arctic [ARCT]) and warm south (North Atlantic Drift [NADR]) of the North Atlantic Current to the oligotrophic central Atlantic Ocean (North Atlantic Subtropical Gyre [NAST], 34° 24' N). (B) The density of the seawater (σ_θ) from depths between 20 m and 200 m indicated the pycnocline between 50 m and 100 m
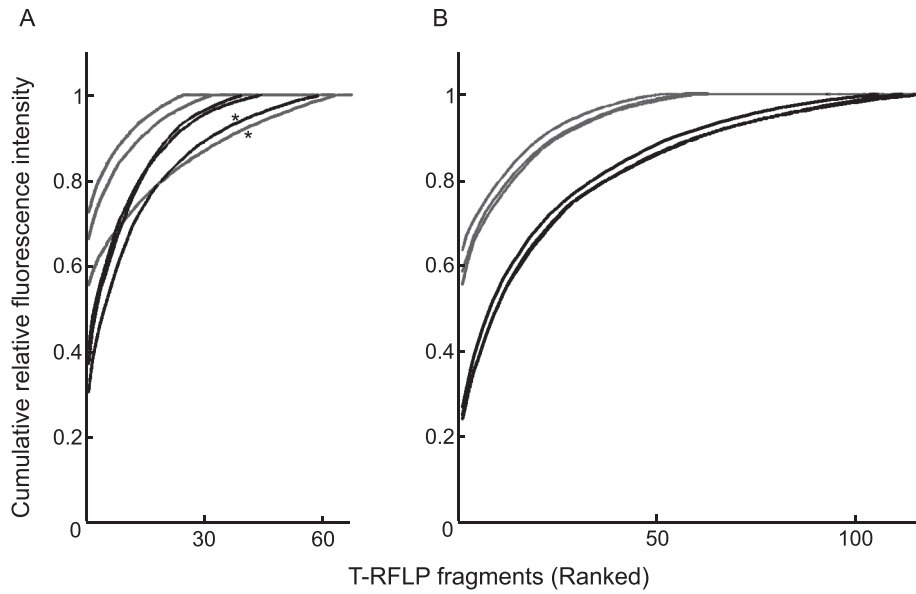
Fig. S2. Identification of T-RFLP patterns with low phylogenetic information was possible with the *k*-dominance plot of T-RFLP patterns. The normalized relative fluorescence intensities (RFI) were visualized in rank versus cumulated abundance curves. Two exemplary sampling sites are shown (A,B), in which each line represents the cumulative RFI of forward (black) or reverse (gray) terminal restriction fragments of one T-RFLP analysis. The species rank at 100% cumulative abundance (RFI) represents the richness of TRFs of the sampling site. (A) The T-RFLP patterns with a high amount of false positive signals (indicated by a star) originated from a fixed fluorescence intensity threshold and overall low fluorescence intensity in this T-RFLP pattern. These T-RFLP patterns were identified as outliers and excluded from further analysis. (B) In contrast, the cumulated abundance curves of T-RFLP patterns of comparable good quality had only slightly different fluorescence intensities between T-RFLP patterns
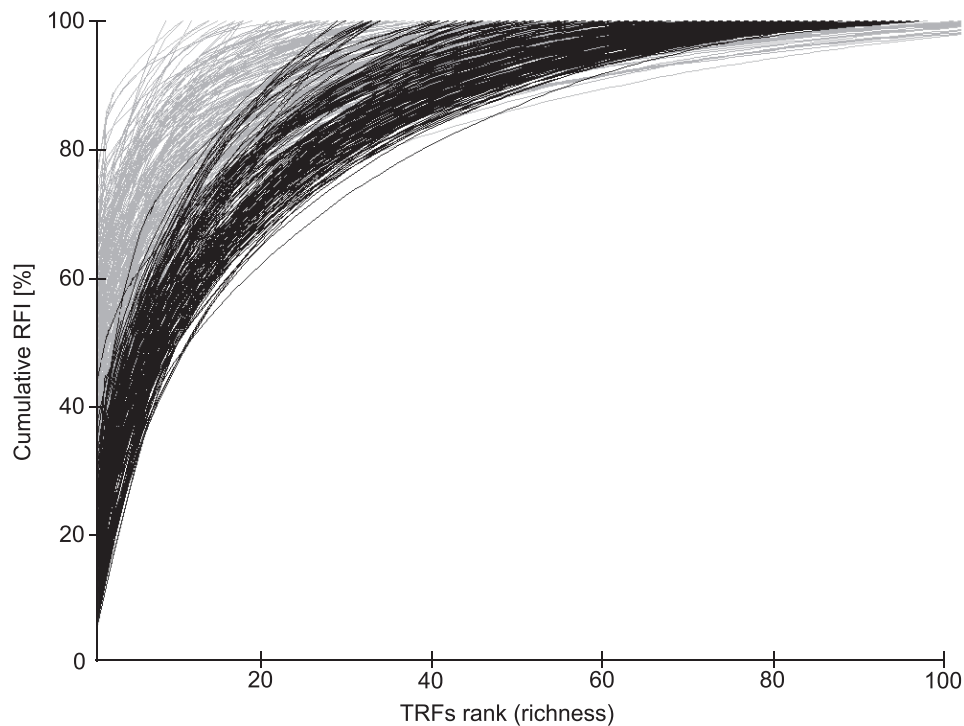

Fig. S3. The diversity richness differed between terminal restriction fragments derived from the forward (black) and from the reverse (gray) primer. Each black and gray line represents the cumulative relative fluorescence intensity of the terminal restriction fragments of 1 T-RFLP analysis. The species rank at 100% cumulative abundance (RFI) represents the richness of TRFs of the sampling site. TRFs of the reverse primer had a lower species rank compared to the TRFs of the forward primer at the same cumulative abundance. Thus, the richness of the reverse primer TRFs was less than that of the forward primer TRFs
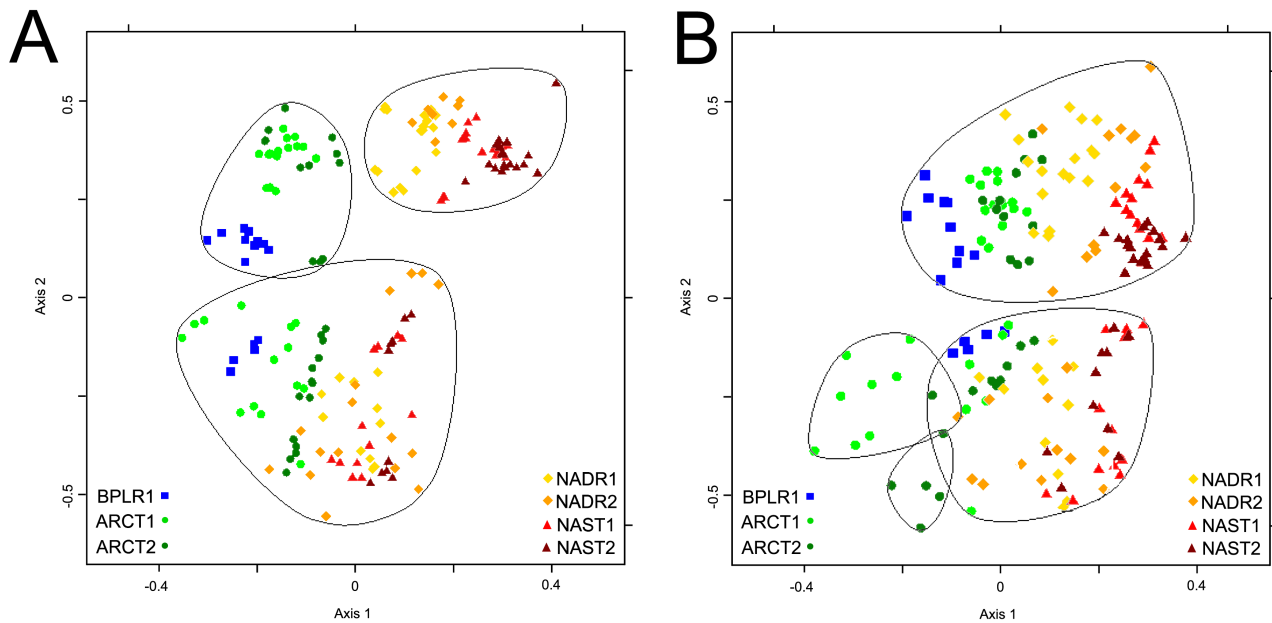
Fig. S4. Nonmetric multidimensional scaling (nMDS) was applied to visualize (A) the Bray-Curtis similarity based on relative fluorescence intensities and (B) the Sørensen index based on presence/absence of terminal restriction fragments (TRFs) in TRF patterns of each sampling site (β diversity). Both biplots had a low stress value of 0.12, indicating a meaningful 2-dimensional visualization. The presentation of the differences revealed a clustering of bacterial communities from the individual water bodies (square: BPLR, dot: ARCT, diamond: NADR, triangle: NAST) along a latitudinal gradient.. A hierarchical clustering defined groups of bacterial communities at 48% Bray-Curtis similarity and 55% Sørensen index (indicated by solid gray lines)
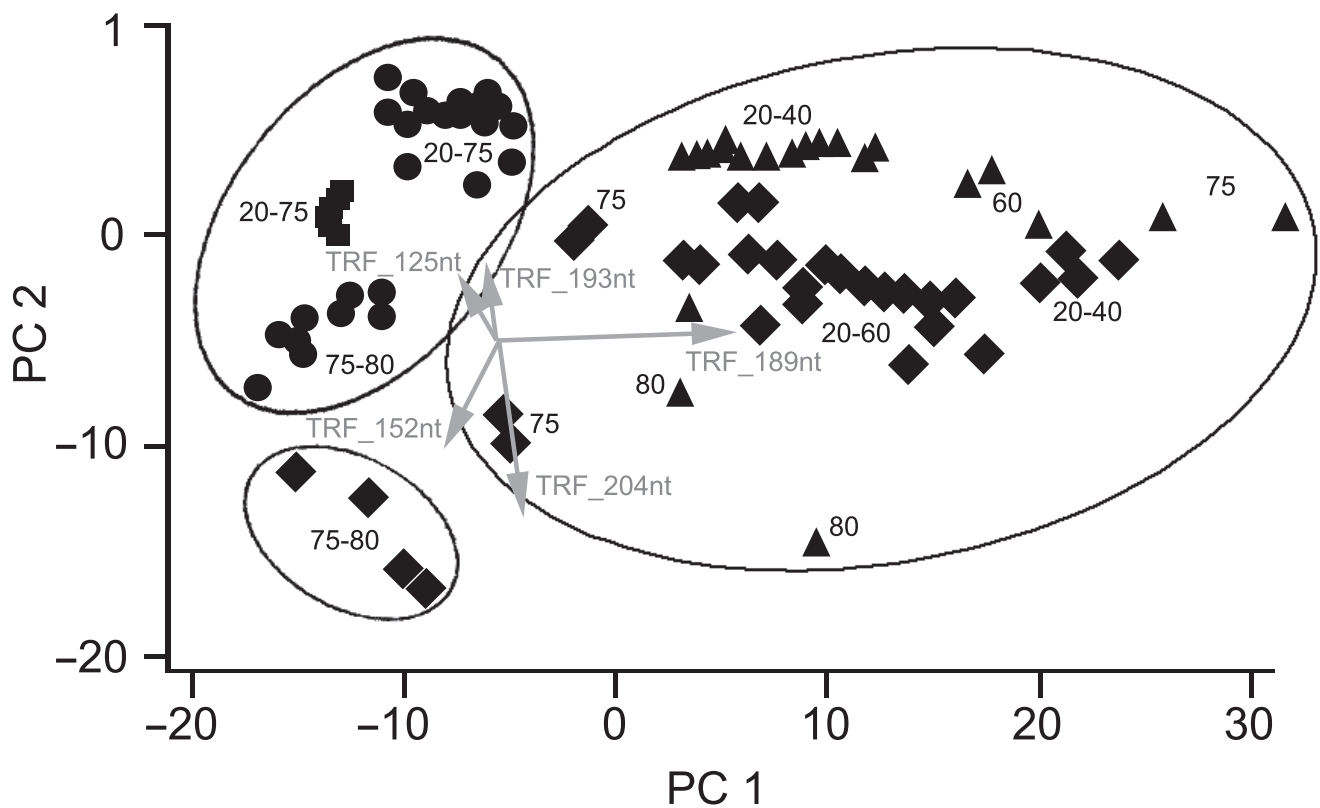

Fig. S5. Principal component analysis was applied to find terminal restriction fragments (TRFs) that cause changes in the overall community structure. The eigenvectors and eigenvalues revealed a differentiation between the northern sampling sites, BPLR (dot) and ARCT (square), above and below 75 m depth, and the southern sampling sites in NADR (diamond) and NAST (triangle), along the first principal component (PC 1). The second principal component (PC 2) distinguished the northern sampling sites (BPLR, ARCT) above and below 75 m depth and

distinguished the southern sampling sites in NADR and NAST. TRF_189nt had the largest eigenvector (gray arrow) parallel to the first principal component, indicating a large contribution to the population in the south of the transect. Numbers in gray represent the depth of the sampling site. A circle represents sampling sites that fall into 1 hierarchical cluster of 50% similarity
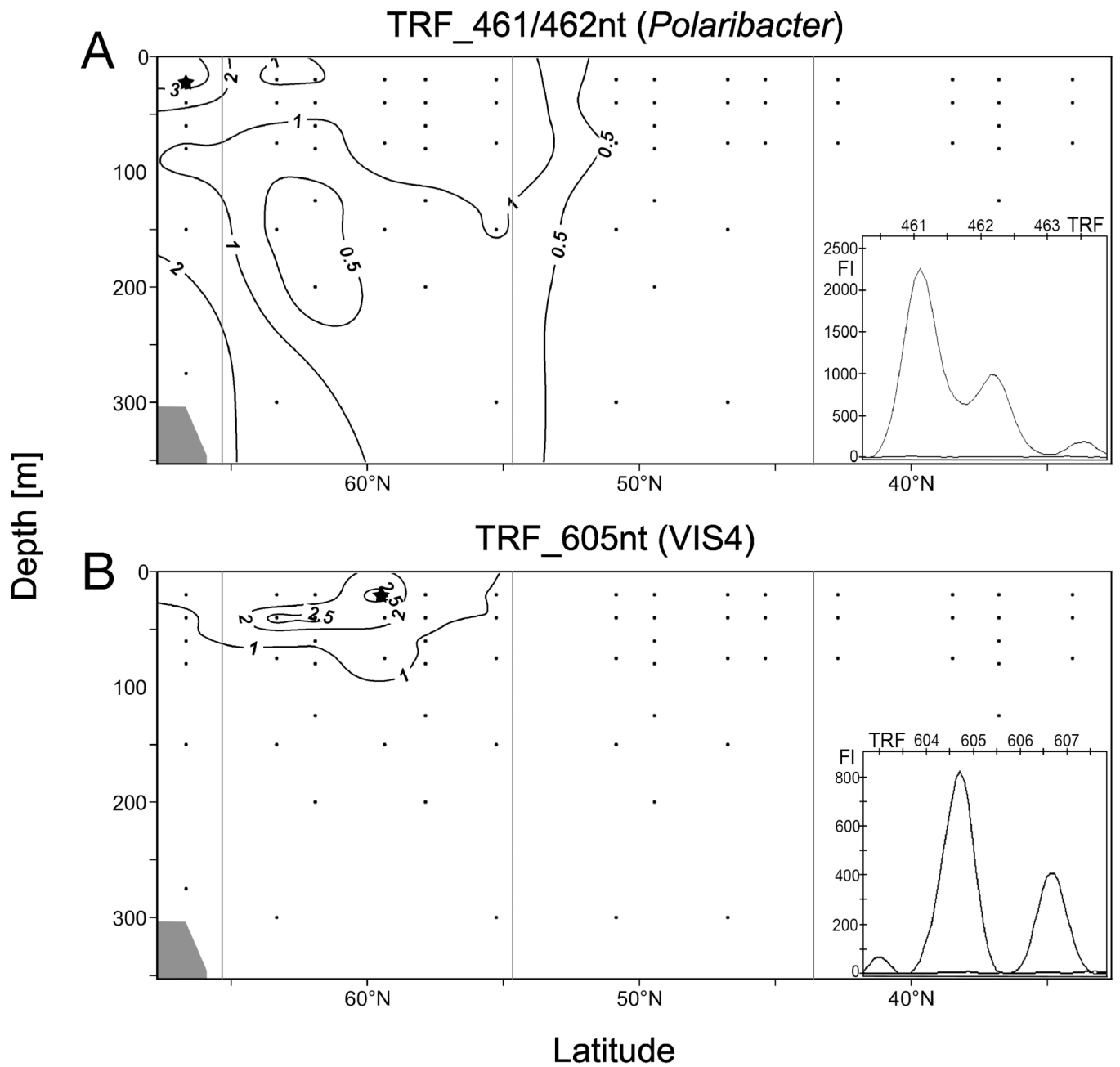
Fig. S6. Relative fluorescence intensity (RFI) pattern of significant terminal restriction fragments (TRFs) affiliated with (A) the *Flavobacteriaceae Polaribacter* and (B) group NS4. Contour lines indicate the relative abundance of TRFs (% RFI). A star indicates the sampling site with the largest RFI and the inset shows the corresponding fluorescence intensity (FI) in the T-RFLP profile
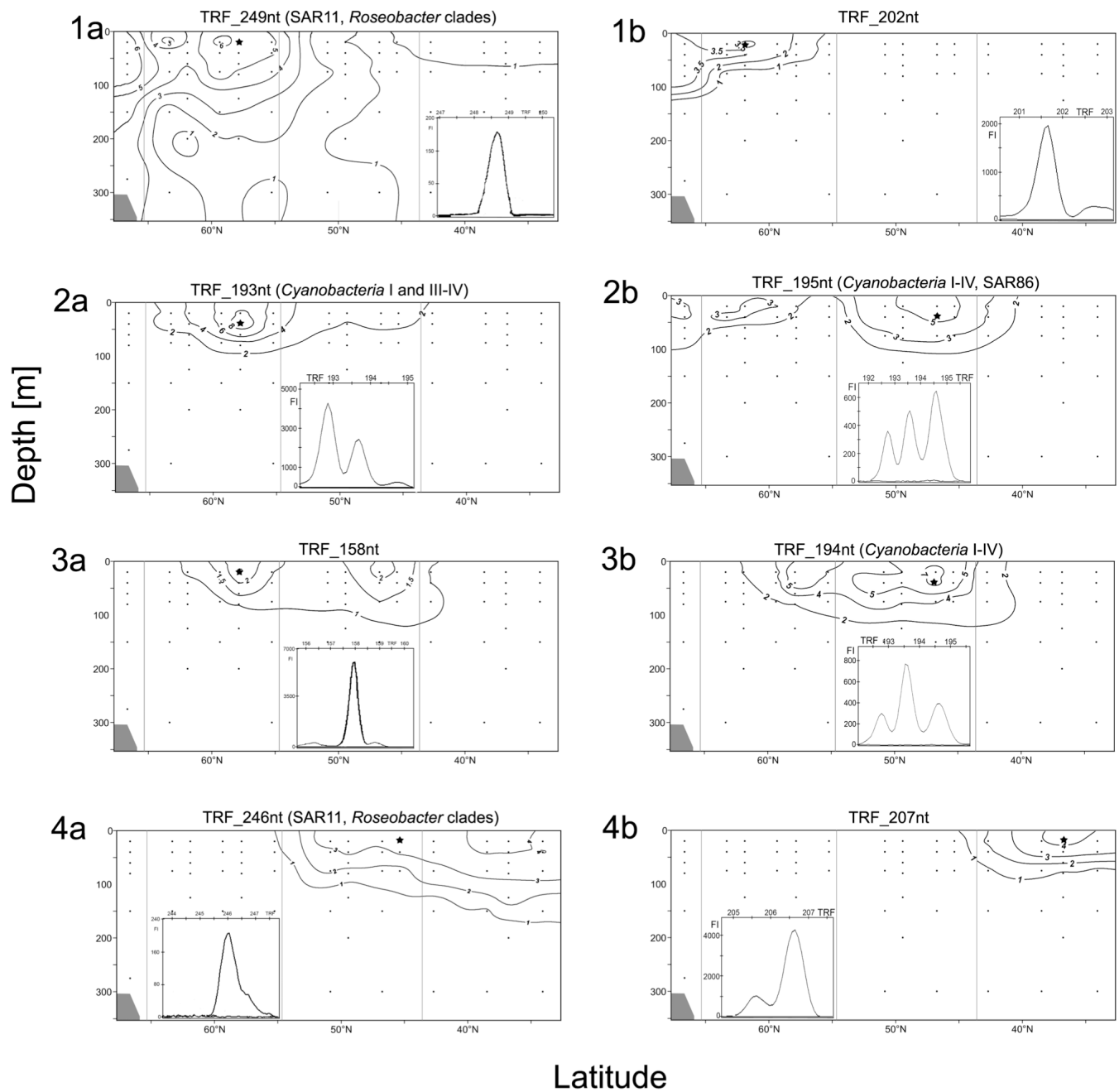
Fig. S7. Relative fluorescence intensity (RFI) pattern of terminal restriction fragments (TRFs) with a regional distribution and no affiliation to a single taxon in the *in silico* (iTRF) calculation. Contour lines indicate the relative abundance of TRFs (% RFI). A star indicates the sampling site with the largest RFI, and the inset shows the corresponding fluorescence intensity (FI) in the T-RFLP profile