Estimating microbial populations by flow cytometry: Comparison between instruments

Catherine Gérikas Ribeiro, Dominique Marie, Adriana Lopes dos Santos, Frederico Pereira Brandini, Daniel Vaulot

Estimating microbial populations by flow cytometry: comparison between instruments

Catherine Gérikas Ribeiro[1*], Dominique Marie[2], Adriana Lopes dos Santos[2], Frederico Pereira Brandini[1], Daniel Vaulot[2].

[1] Departamento de Oceanografia Biológica, Instituto Oceanográfico, Universidade de São Paulo, São Paulo, Brasil

[2] Sorbonne Universités, UPMC Univ Paris 06, CNRS, UMR 7144, Station Biologique, Place Georges Teissier, 29680 Roscoff, France

* Corresponding author

E-mail: catherine.gerikas@gmail.com (CGR)

Running head: Estimating microbial populations by FCM

Keywords: flow cytometry; microbial oceanography; heterotrophic bacteria; picoplankton; nanoplankton.

Abstract

For almost 3 decades, flow cytometry has allowed researchers to investigate ocean planktonic communities using size and cell fluorescence properties. However, oceanographic applications must face two constraints. First, when dealing with marine microbes, instruments must be sensitive because these organisms are very small and with low fluorescence. Second, instruments must be portable to be used on board ships. We compared the performance of two instruments, the BD FACSCanto™ and BD Accuri™ C6. The former is an expensive laboratory-based instrument which has a very good sensitivity, whilst the latter is less sensitive but presents critical advantages for field studies (easy handling and transportation, relatively low cost). We have analyzed 102 samples from the South Atlantic Ocean from 3 transects off Brazil, within the euphotic zone. We compared cell abundance of heterotrophic bacteria, *Prochlorococcus* and *Synechococcus*, as well as photosynthetic pico- and nano-eukaryotes. Heterotrophic bacteria, pico- and nano-eukaryotes could be easily detected with both cytometers. *Prochlorococcus* and *Synechococcus* populations were severely under-estimated with the BD Accuri™ C6, particularly for samples from the well-lit layers of the water column. Correction of abundance data using previously suggested approaches was not sufficient to fully compensate for the low sensibility. Our data suggest that the BD Accuri™ C6 is suitable for counting marine bacteria and photosynthetic eukaryotes, but not *Prochlorococcus* and *Synechococcus*.

Introduction

Flow cytometry (FCM) is a well-established technique (Marie et al. 1997; Gasol and del Giorgio 2000) used since the 1980s (Trask et al. 1982; Olson et al. 1985) for enumeration and characterization of marine micro-organisms. FCM analysis of planktonic communities fulfills the scientific demands of rapid and accurate cell counting, as it considerably reduces the bias introduced by visual counting (Marie et al. 2005). By simultaneously recording several parameters during analysis, FCM allows the discrimination of pico- and nanoplankton populations and the estimation of their abundance, cell size, and pigment content (Marie et al. 2005), both by natural (chlorophyll, phycoerythrin) or induced (fluorescent dyes) fluorescence (Marie et al. 1997).

The flow cytometer registers events as cells are aligned in a fluid stream and flow through a beam of focused light usually provided by one or several lasers. For each particle, scattered light and emitted fluorescence are converted to digital signals and recorded. A flow cytometer comprises three main systems: fluidics (particle transport), optics (laser beam and optical filters), and electronics (signal conversion into electronic data). Detectors for scattered light located at 180° and 90° from the light source are called forward scatter (FSC) and side scatter (SSC), respectively. Fluorescence at different wavelengths (typically green, orange and red) is also recorded. Signals associated with each parameter are displayed as cytograms, which are used to discriminate and count different populations based on scattering and fluorescence features. Phytoplankton populations can be differentiated by FCM according to specific values of the recorded parameters (FSC, SSC, red or orange fluorescence).

Bacteria are in general detected after staining with a nucleic acid strain such as SYBR Green-I (Marie et al. 1997). Two different groups can be distinguished based on their apparent nucleic acid content (differences in fluorescence intensity) and side scatter signal (SSC): high nucleic acid (HNA) and low nucleic acid (LNA) bacteria. The function and ecological importance of these two groups is far

from being fully understood (Bouvier et al. 2007; Van Wambeke et al. 2011), and several studies have addressed these nucleic acid content differences in terms of ecological traits, such as bacterial activity and production (Morán et al. 2007; Ortega-Retuerta et al. 2008; Van Wambeke et al. 2011).

Two main groups of autotrophic prokaryotes dominate picoplankton: *Prochlorococcus* and *Synechococcus*. *Prochlorococcus* is ubiquitous in the euphotic zone of tropical oceans, being considered the most abundant photosynthetic organisms on the planet (Partensky et al. 1999b), and its discovery was only made possible with the development of flow cytometry (Chisholm et al. 1988). *Prochlorococcus* is discriminated by its small scattering and low red fluorescence (chlorophyll). *Synechococcus* is widely distributed in marine environments, being particularly abundant in well-lit and nutrient rich top layers of the oceans (Partensky et al. 1999a). One of the key parameters that allows *Synechococcus* populations to be discriminated by FCM is the phycoerythrin content (orange fluorescence). Different *Synechococcus* clades can show distinct fluorescence signatures (Olson et al. 1990; Thompson and van den Engh 2016), as a result of different phycobilisome composition (Scanlan et al. 2009). Pico- and nanoeukaryotes are important contributors to global primary productivity (Li 1994), and due to their larger cell size, they often contribute to an important share of autotrophic biomass in the oceans (Zubkov et al. 1998). Picoeukaryotes, which cells range from 0.8 µm to 2-3 µm (Simon et al. 1994), present well-defined cytometric signatures by FCM, while nano-eukaryotes populations are less well defined.

FCM analysis has led to numerous advances in marine microbial ecology, although cost and maintenance expenses were prohibitive for many laboratories until recently (Gasol and del Giorgio 2000; Vives-Rego et al. 2000). Since the first cytometry-based field study made by Olson, Vaulot & Chisholm (1985), on-board flow cytometry has become a crucial tool in the investigation of both autotrophic and heterotrophic picoplanktonic communities (Legendre et al. 2001). The manufacturing of low cost compact benchtop flow cytometers such as the BD Accuri™ C6, the Millipore Guava® or the Applied Biosystems Attune® has facilitated the use of FCM to study of phytoplankton communities

around the world, due to easy handling, automatic sampling and easy transportation (a critical quality for field measurements). However, these low cost instruments can be less sensitive than laboratory based flow cytometers, due to less sophisticated optical and/or electronic systems.

A lower sensitivity is usually not a problem for bacteria which are detected after staining with strongly fluorescing dyes such as SYBR Green, or for small eukaryotes whose pigment content is relatively high. However, this is not the case for cyanobacteria like *Prochlorococcus*, for which the concentration of photosynthetic pigments per cell is as much as 50-100 fold lower in cells exposed to high light as a result of photo-acclimation (Sosik et al. 1989; Olson et al. 1990), creating 'dim' populations in the surface layers. Such low fluorescence explains why *Prochlorococcus* escaped detection by researchers using epifluorescence microscopy or even during the first use of FCM on board oceanographic ships (Olson et al. 1985). To overcome the problem of low sensitivity flow cytometers, both direct and indirect approaches to infer *Prochlorococcus* abundance have been developed, such as changes in cytometer optical set up to improve excitation energy or fluorescence detection (Dusenberry and Frankel 1994; Partensky et al. 1999b) and the use of mathematical corrections (Zubkov et al. 1998; Crosbie and Furnas 2001).

In this paper, we compare data obtained on marine microbial communities with two flow cytometers, the FACSCanto™ and the Accuri™ C6 (hereafter named as CANTO and C6). Although manufactured by the same company (BD Biosciences, San Jose, CA, USA), these cytometers present distinct fluorescence excitation/detection technical features (Table 1). Differences in laser, optics and electronic systems can potentially affect sensitivity and resolution, influencing the accuracy of field measurements. We analyzed heterotrophic marine bacteria, photosynthetic eukaryotes and cyanobacteria on a set of marine samples from the South Atlantic Ocean (displaying both nutrient and light gradients within the water column). While both instruments produced equivalent data for bacteria and eukaryotes, cyanobacteria, especially *Prochlorococcus*, were severely under-estimated with the C6 instrument, and procedures previously suggested to correct the data proved ineffective.

Materials and methods

Sampling

Water samples were collected onboard the R/V "Alpha Crucis", between 31/10/2013 and 23/11/2013. The surveyed area was located between latitude 23°11'S - 30°52'S and longitude 39°22”W - 49°09”W, along 3 transects (TR1, TR2 and TR3), in the South West Atlantic off Brazil, reaching the 3510 meters isobath (Fig. 1). The sampling strategy comprised cross-shelf transects with 5 depths per station within the euphotic zone for TR1 and TR2, as well as 12 surface samplings for TR3, for a total of 102 samples. Three water masses were sampled during the cruise: the warm and oligotrophic Tropical Water, the cold and nutrient rich South Atlantic Central Water and the Coastal Water, with highly variable features (Castro et al. 2006). Except for TR3 samples, which were collected using a polycarbonate bucket, seawater samples were collected with 12 L Niskin bottles attached to a Seabird® CTD-rosette system (Sea-Bird Electronics, Bellevue, WA, USA), divided into cryotubes, preserved with 0.1% glutaraldehyde, final concentration (modified from Vaulot et al., 1989), incubated for 10 minutes in the dark, flash-frozen in liquid nitrogen and stored at -80°C until analysis.

Flow cytometry analysis

Two flow cytometers were used in this study: a BD FACSCanto II™ and a BD Accuri™ C6 (Table 1). Samples were counted simultaneously on both cytometers located in the same room, in order to avoid any possible bias by manipulation or time span between measurements. The tubing of the C6 was new and fluidics were calibrated for precise volume measurements as recommended by the manufacturer (Section 4.13 of manual).

Samples were first analyzed unstained to enumerate phototrophs. Fluorescent beads (0.95 G Fluoresbrite® Polysciences, Warrington, PA) were added in each sample in order to normalize parameters (Marie et al. 1997). A second analysis was performed to enumerate heterotrophic bacteria

after staining with SYBR Green® (1:10000, final concentration) (Ref-S7585, Life Technologies, Eugene, Oregon).

On the C6, for enumerating phytoplankton, 200 µL of sample were analyzed at the "high" rate (66 $\mu L.min^{-1}$) with a threshold set at 700 on red fluorescence (FL3-H). To enumerate heterotrophs, 60 µL of SYBR Green stained samples were run at "medium" rate (35 $\mu L.min^{-1}$) and the threshold was set at 700 on green fluorescence (FL1-H). In both cases, thresholds were determined by running 0.2 µm filtered sea water sample and lowering the values until electrical or optical noise appears.

On the CANTO for enumerating phytoplankton, samples were run for 3 min with a rate of 72 $\mu L.min^{-1}$ and with the discriminator set on red fluorescence at 200. For bacterial enumeration, SYBR Green stained samples were run for 2 min at a rate of 60 $\mu L.min^{-1}$ and the threshold was set on green fluorescence at 500. Flow rate was determined by the method described by Marie et al. (1997). A known volume of seawater was injected on the CANTO for at least 10 min. Then the remaining volume is measured and the rate is determined by dividing the difference between initial and final volumes by the injection time.

Data were analyzed with the Flowing Software® 2.5 (http://www.flowingsoftware.com). Each population was identified on the cytograms on the basis of its scatter and fluorescence signals (Fig. S1). Each parameter was normalized to that of the reference beads (0.95 µm). Cell counts for each red fluorescence value were exported from the single parameter histogram. The resulting spreadsheet (File S1) was used in subsequent analysis with the R software (R Development Core Team, 2013).

Near the surface the red chlorophyll fluorescence of the picophytoplankton decreases due to photoacclimation (Partensky et al. 1993; Dusenberry et al. 2001; Kulk et al. 2011). Therefore, for a fraction or even all of the *Prochlorococcus* and *Synechococcus* populations, fluorescence can fall below the detection threshold (Fig. 2). For the case where only a part of the population was in the noise, we modified the correction procedure described by Crosbie and Furnas (2001) and implemented it as an R routine (File S2). This correction assumes that the red fluorescence distribution of these

populations has a log-normal shape (Crosbie and Furnas 2001; Shapiro 2003) and that, when the left part of the distribution is partially in the noise, the left part can be extrapolated from the right part. The R routine takes as input histograms produced by the Flowing Software® (but can be adapted to other data formats) and outputs uncorrected and corrected cell abundance data (output data examples can be found in Files S2 and S3). Three cases can occur for a given population.

1. The mode of the histogram is not visible (e.g. Fig. 2G). In this case, the population is considered to be mostly within noise, without the possibility of counting or correction. Hence, this population is removed from the dataset and labeled as 'cells in noise' by the R routine.
2. The mode is visible, but the left part of the distribution is partly below the noise level (e.g. Fig. 2A). In this case, the abundance of each population is calculated as the double of the right part of the histogram, from the mode to the maximum, and samples are labeled as 'correction'
3. The mode is visible and the distribution is totally out of the noise (e.g. Fig. 2C). In this case no correction is performed and the initial output value is kept ('no correction' samples).

In some cases, the automatic correction needs some degree of visual confirmation, especially for deeper samples with low cell numbers resulting in noisy histograms (e.g. Fig. 2Q or 2W). Therefore, the R routine provides a graphical output of the histogram for each sample (File S3), allowing the user to visually confirm whether the automatic labeling (case 1, 2, or 3 above) is correct. Statistical analyses were performed with the PRISM® 7 software (http://www.graphpad.com/scientific-software/prism).

Results

Bacterial populations were well resolved for both CANTO and C6 for all samples (Fig. S1). High Nucleic Acid and Low Nucleic Acid bacterial populations were consistently distinguished with both cytometers and there was a very good correlation for both HNA and LNA (R^2 = 0.85 and 0.89, respectively) between the data obtained on the different instruments (Fig. 3A, B). For HNA slope was

183 statistically different from 1 ($p < 0.0001$) and abundances estimated by the C6 were consistently lower
184 by 10-15% compared to the CANTO.

185 The chlorophyll fluorescence of *Prochlorococcus* and *Synechococcus* decreases from the deeper
186 layers to the surface in response to photoacclimation. For samples near the surface, cells from both
187 populations can be partly or totally in the noise depending on the instrument sensitivity. Fig. 2
188 demonstrates clearly that the C6 is less sensitive than the CANTO by at least a factor of 10 (note for
189 example Fig. 2C and 2I for *Prochlorococcus* at 110 m for the CANTO and C6, the mode of the
190 histogram is in the noise for the C6 and about 10 times higher than the noise level for the CANTO).
191 With the C6, *Prochlorococcus* populations were completely in the noise above 100 m (55 out of 102
192 samples, Fig. 2 and 4, Table 2) and for *Synechococcus* partly or completely in the noise above 50 m (15
193 out of 102 samples, Fig. 2 and 4, Table 2). With the CANTO, only surface *Prochlorococcus* were
194 partly in the noise and *Synechococcus* was always fully resolved (Fig. 2 and 4, Table 2). When cells
195 were only partly in the noise (i.e. when the histogram mode was clearly visible, e.g. Fig. 2U), we
196 estimated the part of the population that was in the noise using the approach proposed by Crosbie and
197 Furnas (2001) (see Material and Methods section). The comparison between the C6 data which
198 required correction and the CANTO data that did not require correction allowed us to assess the
199 validity of this approach (Fig. 3C and D).

200 Clearly some data points that are corrected (grey squares) appear as outliers and are severely
201 underestimated with the C6, even after correction (Fig. 3C and D). The slopes for non-corrected
202 samples (solid circles) are significantly different from 1 ($p < 0.0001$), being respectively 0.75
203 (*Prochlorococcus*) and 0.77 (*Synechococcus*) (Fig. 3C and D), which corresponds to 25% lower
204 abundance on average with the C6. Vertical cross sections of two transects (Fig. 4) illustrate that, while
205 the CANTO provides fully resolved vertical profiles for both *Prochlorococcus* and *Synechococcus*
206 population, the data from the C6 cannot be used in the upper layer (roughly from 100 m to the surface).

Pico- and Nano-eukaryotes were always above the detection limit for both instruments with excellent correlation between the two instruments ($R^2 = 0.94$ and 0.69, respectively, Fig. 3E and F). As for the other populations, the slopes were significantly different from 1 ($p < 0.0001$). While picoeukaryotes were about 15% more abundant with the C6, it was the reverse for nanoeukaryotes which were slightly underestimated by the C6.

Discussion

The analysis of planktonic communities by flow cytometry is complex because the distinctive cell features of each population may change with depth, diel cycle and nutrient conditions (e.g. Vaulot and Marie 1999). From the six planktonic groups studied here, four were well resolved by both the CANTO and C6 flow cytometers: HNA and LNA heterotrophic bacteria, autotrophic pico-eukaryotes and nano-eukaryotes. Abundance measured by both instruments were tightly correlated. Abundance was always slightly lower with the C6 than with the CANTO, except for picoeukaryotes. This could have resulted from an imperfect calibration of the analyzed volume for one of the instruments. While on the CANTO, the analyzed volume is manually calibrated, the C6 relies on a calibration every time tubing is changed following the manufacturer's recommendation. However, we have recently observed that the actual volume analyzed varies, even during the course of a day, and needs to be re-calibrated at fixed intervals using a procedure similar to the one used for the CANTO (D. Marie, unpublished data). The slightly lower correlation coefficient observed for nanoeukaryotes could result from the difficulty to clearly distinguish the limits of the pico and nano-eukaryote populations.

In contrast, the lower sensitivity of the C6 led a drastic underestimation or even non-detection of the cyanobacteria in the upper 100 m of the water column, especially for *Prochlorococcus.* But the phenomenon was also present for *Synechococcus* although it had not been recognized previously. For both populations, corrections, such as those recommended previously (Zubkov et al. 1998; Crosbie and Furnas 2001), did not completely solve the problem, since in some samples the corrected counts were

still much lower than those obtained with the CANTO (Fig. 3C, D). The data obtained with the C6 for *Prochlorococcus* but also *Synechococcus,* should be considered with great caution even when only a part of the population is in the noise.

The oceanographic transects (Fig. 4) demonstrate that the C6 instrument would result in very serious underestimates of the cyanobacteria abundance and therefore contribution to the carbon biomass, at least in oligotrophic to mesotrophic waters. For *Synechococcus,* one solution could be to trigger acquisition on PE fluorescence (FL2) which is general quite strong, but this will require to run every sample twice, with the trigger on FL2 and FL3 respectively. Unfortunately this was not tested on our samples. For *Prochlorococcus*, the only solution to overcome the problem of low sensitivity is to increase either the excitation energy or the fluorescence detection through optical solutions (Dusenberry and Frankel 1994; Partensky et al. 1999b).

Conclusion

The increasing affordability of benchtop flow cytometers comes with limitations in some of the equipment features, such as lower detection limits. The comparison between studies of phytoplankton communities by flow cytometry should take into account the equipment used, particularly in approaches involving populations of *Prochlorococcus* and *Synechococcus*, in which low chlorophyll concentration per cell can lead to the underestimation of their abundance in the top euphotic zone. Still these benchtop flow cytometers provide reliable data for other populations such as heterotrophic bacteria and photosynthetic eukaryotes. Our study highlight the need for careful comparison between instruments before using them for large scale oceanographic surveys, using as reference the most sensitive laboratory instruments available.

253

Acknowledgments

254 We thank the officers and crew of the R/V *Alpha Crucis* for logistical support. Financial
255 support for this work was provided by the CNRS Groupement de Recherche International (GDRI)
256 "Diversity, Evolution and Biotechnology of Marine Algae", the COFECUB-CAPES project "Pico
257 Brás" (TE 871-15) and FAPESP - *Fundação de Amparo à Pesquisa do Estado de São Paulo*
258 (2012/04800-9 and 2014/15242-2).

259

260

References

Bouvier, T., P. A. del Giorgio, and J. M. Gasol. 2007. A comparative study of the cytometric characteristics of High and Low nucleic-acid bacterioplankton cells from different aquatic ecosystems. Environ. Microbiol. **9**: 2050–2066. doi:10.1111/j.1462-2920.2007.01321.x

Castro, B. M., F. P. Brandini, A. M. S. Pires-Vanin, and L. B. Miranda. 2006. Physical oceanography of the western Atlantic continental shelf located between 4 N and 34 S, p. 209–251. *In* A.R. Robinson and K.H. Brink [eds.], The Sea, Volume 14A: The Global Coastal Ocean. Harvard University Press.

Chisholm, S. W., R. J. Olson, E. R. Zettler, R. Goericke, J. B. Waterbury, and N. A. Welschmeyer. 1988. A novel free-living prochlorophyte abundant in the oceanic euphotic zone. Nature **334**: 340–343. doi:10.1038/334340a0

Crosbie, N., and M. Furnas. 2001. Abundance , distribution and flow- cytometric characterization of picophytoprokaryote populations in central (17ºS) and southern (20ºS shelf waters of the Great Barrier Reef. J. Plankton Res. **23**: 809–828. doi:809–828. doi:10.1093/plankt/23.8.809

Dusenberry, J. A., and S. L. Frankel. 1994. Increasing the sensitivity of a FACScan flow cytometer to study oceanic picoplankton. Limnol. Oceanogr. **39**: 206–209. doi:10.4319/lo.1994.39.1.0206

Dusenberry, J. A., R. J. Olson, and S. W. Chisholm. 2001. Photoacclimation kinetics of single-cell fluorescence in laboratory and field populations of *Prochlorococcus*. Deep. Res. Part I Oceanogr. Res. Pap. **48**: 1443–1458. doi:10.1016/S0967-0637(00)00096-0

Gasol, J. M., and P. a del Giorgio. 2000. Using flow cytometry for counting natural planktonic bacteria and understanding the structure of planktonic bacterial communities. Sci. Mar. **64**: 197–224. doi:10.3989/scimar.2000.64n2197

Kulk, G., W. H. van de Poll, R. J. W. Visser, and A. G. J. Buma. 2011. Distinct differences in

photoacclimation potential between prokaryotic and eukaryotic oceanic phytoplankton. J. Exp. Mar. Bio. Ecol. **398**: 63–72. doi:10.1016/j.jembe.2010.12.011

Legendre, L., C. Courties, and M. Troussellier. 2001. Flow cytometry in oceanography 1989-1999: Environmental challenges and research trends. Cytometry **44**: 164–172. doi:10.1002/1097-0320(20010701)44:3<164::AID-CYTO1108>3.0.CO;2-6

Li, W. K. W. 1994. Primary production of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: Measurements from flow cytometric sorting. Limnol. Oceanogr. **39**: 169–175. doi:10.4319/lo.1994.39.1.0169

Marie, D., F. Partensky, S. Jacquet, and D. Vaulot. 1997. Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. Appl. Environ. Microbiol. **63**: 186–193.

Marie, D., N. Simon, and D. Vaulot. 2005. Phytoplankton Cell Counting by Flow Cytometry, p. 253–556. *In* R.A. Andersen [ed.], Algal culturing techniques. Academic Press.

Morán, X., A. Bode, L. Suárez, and E. Nogueira. 2007. Assessing the relevance of nucleic acid content as an indicator of marine bacterial activity. Aquat. Microb. Ecol. **46**: 141–152. doi:10.3354/ame046141

Olson, R. J., S. W. Chisholm, and E. R. Zettler. 1990. Pigments, size, and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. Limnol. Oceanogr. **35**: 45–58.

Olson, R. J., D. Vaulot, and S. W. Chisholm. 1985. Marine phytoplankton distributions measured using shipboard flow cytometry. Deep Sea Res. Part A. Oceanogr. Res. Pap. **32**: 1273–1280. doi:10.1016/0198-0149(85)90009-3

Ortega-Retuerta, E., I. Reche, E. Pulido-Villena, S. Agustí, and C. Duarte. 2008. Exploring the relationship between active bacterioplankton and phytoplankton in the Southern Ocean. Aquat. Microb. Ecol. **52**: 99–106. doi:10.3354/ame01216

Partensky, F., J. Blanchot, and D. Vaulot. 1999a. Differential distribution and ecology of

Prochlorococcus and *Synechococcus* in oceanic waters : a review. Bull. l'Institut océanographique.

Partensky, F., W. R. Hess, and D. Vaulot. 1999b. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. Rev. **63**: 106–127.

Partensky, F., N. Hoepffner, K. W, O. Ulloa, and D. Vaulot. 1993. Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) Strains isolated from the northern Atlantic and the Mediterranean sea. Plant Physiol. **101**: 285–296.

Scanlan, D. J., M. Ostrowski, S. Mazard, and others. 2009. Ecological genomics of marine picocyanobacteria. Microbiol. Mol. Biol. Rev. **73**: 249–299. doi:10.1128/MMBR.00035-08

Shapiro, H. M. 2003. Practical flow cytometry, 3rd Edition, by Howard M. Shapiro, M.D., Wiley-Liss, Inc., New York, 1995, 542 pages, 4th Ed.

Simon, N., R. G. Barlow, D. Marie, F. Partensky, D. Vaulot, and Qd. 1994. Characterization of oceanic photosynthetic picoeukaryotes by flow-cytometry. J. Phycol. **30**: 922–935. doi:10.1111/j.0022-3646.1994.00922.x

Sosik, H. M., S. W. Chisholm, and R. J. Olson. 1989. Chlorophyll fluorescence from single cells: Interpretation of flow cytometric signals. Limnol. Oceanogr. **34**: 1749–1761. doi:10.4319/lo.1989.34.8.1749

Team R Development Core. 2013. R: A language and environment for statistical computing. R Found. Stat. Comput. 1–3464.

Thompson, A. W., and G. van den Engh. 2016. A multi-laser flow cytometry method to measure single cell and population-level relative fluorescence action spectra for the targeted study and isolation of phytoplankton in complex assemblages. Limnol. Oceanogr. Methods **14**: 39–49. doi:10.1002/lom3.10068

Trask, B. J., G. J. van den Engh, and J. H. Elgershuizen. 1982. Analysis of phytoplankton by flow cytometry. Cytometry **2**: 258–264. doi:10.1002/cyto.990020410

Vaulot, D., C. Courties, and F. Partensky. 1989. A simple method to preserve oceanic phytoplankton for flow cytometric analyses. Cytometry **10**: 629–635. doi:10.1002/cyto.990100519

Vaulot, D., and D. Marie. 1999. Diel variability of photosynthetic picoplankton in the equatorial Pacific. J. Geophys. Res. Ocean. **104**: 3297–3310. doi:10.1029/98JC01333

Vives-Rego, J., P. Lebaron, and G. Nebe-von Caron. 2000. Current and future applications of flow cytometry in aquatic microbiology. FEMS Microbiol. Rev. **24**: 429–448. doi:10.1111/j.1574-6976.2000.tb00549.x

Van Wambeke, F., P. Catala, M. Pujo-Pay, and P. Lebaron. 2011. Vertical and longitudinal gradients in HNA-LNA cell abundances and cytometric characteristics in the Mediterranean Sea. Biogeosciences **8**: 1853–1863. doi:10.5194/bg-8-1853-2011

Zubkov, M. V, M. A. Sleigh, G. A. Tarran, P. H. Burkill, and R. J. . Leakey. 1998. Picoplanktonic community structure on an Atlantic transect from 50°N to 50°S. Deep Sea Res. Part I Oceanogr. Res. Pap. **45**: 1339–1355. doi:10.1016/S0967-0637(98)00015-6

Tables

Table 1. Technical features of BD FACSCanto™ and BD Accuri™ C6 flow cytometers according to the manufacturer.

Features	BD FACSCanto™	BD Accuri™ C6
Weight	149.7 kg	13.6 kg
Acquisition Software	BD FACSDiva	BD CSampler™
Signal Processing	Digital	Digital
Number of lasers	2	2
Total PMT[a] for fluorescence	8	4
Laser configuration	Blue/red	Blue/red
Laser wavelength	488 nm, 20 mW solid state 633 nm, 17 mW HeNe	488nm; 50mW solid state 640nm; 30mW diode
Excitation light	Optic fiber	Direct
Florescence sensitivity	FITC[b] < 100 MESF[c] PE[d] < 50 MESF	FITC < 150 MESF PE < 100 MESF
Optical alignment	Fixed alignment	Fixed alignment
Fluidics	Positive-pressure pump	Peristaltic pump
Sample acquisition	18 bits / 5 decades	24 bits / 7 decades
Sample processing	Tubes	Tubes/96-well plates

a) PMT: photomultiplier; b) FITC: fluorescein isothiocyanate; c) MESF: molecules of equivalent soluble fluorochrome; d) PE: phycoerythrin.

Table 2. Number of samples assigned as 'correction', 'no correction' and 'cells in noise' for each pico-cyanobacteria group and equipment tested.

	No correction	**Correction**	**Cells in noise**
CANTO - *Prochlorococcus*	74	27	1
C6 - *Prochlorococcus*	27	20	55
CANTO - *Synechococcus*	101	1	0
C6 - *Synechococcus*	58	29	15

362

363 ## Figure legends

364 **Fig. 1.** Stations sampled in the South Atlantic Ocean off Brazil during the CARBOM cruise in 2013.
365 Profiles: transect 1 (TR1, shaded circles) ; transect 2 (TR2, shaded triangles) and surface sampling,
366 transect 3 (TR3, shaded squares). The grey scale on the right indicates bottom depths.
367
368 **Figure 2.** Examples of depth profiles (St. 100 and St. 114) of normalized red fluorescence distribution
369 (relative cell number *versus* chlorophyll fluorescence) and cell abundance for *Prochlorococcus* (**a-l**)
370 and *Synechococcus* (**m-x**) on BD FACSCanto™ and BD Accuri™ C6. For each distribution, it is
371 indicated whether the cells were in the noise (Noise) or whether a correction was needed (Corr.). In the
372 depth profiles (**f, l, r, x**), solid symbols represent samples for which no correction was needed; grey
373 symbols indicate samples for which we applied a correction (see Materials and Methods); samples
374 within noise were removed.
375
376 **Figure 3.** Relationship between abundance measurements performed with BD Accuri™ C6 and BD
377 FACSCanto™ (in cells.mL^{-1}): (**a**) HNA bacteria; (**b**) LNA bacteria; (**c**) *Prochlorococcus*; (**d**)
378 *Synechococcus*; (**e**) picoeukaryotes and (**f**) nanoeukaryotes. For *Prochlorococcus* and *Synechococcus*:
379 'no correction': solid circles, 'correction': grey squares; 'cells in noise': open triangles. The coefficient of
380 determination and the equation are indicated on each graphic. The regression line calculated from 'no
381 correction' samples is marked in black. All the slopes differed significantly from 1 ($p < 0.0001$), except
382 for LNA bacteria ($p = 0.049$).
383
384 **Figure 4.** Vertical abundance distribution (cells.mL^{-1}) for measurements with BD FACSCanto™ (left
385 column) and BD Accuri™ C6 (right column): *Prochlorococcus* (**a, b, e, f**) and *Synechococcus* (**c, d, g,**
386 **h**). Top labels correspond to station number. Sampled points are marked as: 'no correction' (solid
387 circles), 'correction' (grey squares), or 'cells in noise' (open triangles). Figures were drawn with the
388 Ocean Data View software (https://odv.awi.de/).
389

Supplementary material is available at https://figshare.com/s/a9499d9ab4f4740eb576 .

Supplementary figure legend

Figure S1. Cytograms of phycoerythrin *versus* chlorophyll fluorescence (**a, b**), side scatter *versus* chlorophyll fluorescence (**c, d**) and side scatter *versus* DNA fluorescence (**e, f**) for sample 137 (St. 100, 110 meters depth) for BD FACSCanto™ and BD Accuri™ C6 analyses, showing the gating windows: *Prochlorococcus* (pink), *Synechococcus* (green), picoeukaryotes (blue) nanoeukaryotes (orange), HNA bacteria (yellow) and LNA bacteria (red). Calibrations beads are marked in black.

Supplementary material

File S1. Example of input file for R routine.

channel	sample135_C6_PRO_5m	sample136_C6_PRO_50m	sample137_C6_PRO_110m	sample138_C6_PRO_130m	sample139_C6_PRO_170m
1	0	0	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0
4	0	0	0	0	0
5	0	0	0	0	0
6	0	0	0	0	0
7	0	0	0	0	0
8	0	0	0	0	0
9	0	0	0	0	0
10	0	0	0	0	0
11	0	0	0	0	0
12	0	0	0	0	0
13	0	0	0	0	0
14	0	0	0	0	0
15	0	0	0	0	0
16	0	0	0	0	0
17	0	0	0	0	0
18	0	0	0	0	0
19	0	0	0	0	0
20	0	0	0	0	0
21	0	0	0	0	0
22	0	0	0	0	0
23	0	0	0	0	0
24	0	0	0	0	0
25	0	0	0	0	0
26	0	0	0	0	0
27	0	0	0	0	0
28	0	0	0	0	0
29	0	0	0	0	0
30	0	0	0	0	0
31	0	0	0	0	0
32	0	0	0	0	0
33	0	0	0	0	0
34	0	0	0	0	0
35	0	0	0	0	0

36	0	0	0	0	0
37	0	0	0	0	0
38	0	0	0	0	0
39	0	0	0	0	0
40	0	0	0	0	0
41	0	0	0	0	0
42	0	0	0	0	0
43	0	0	0	0	0
44	0	0	0	0	0
45	0	0	0	0	0
46	0	0	0	0	0
47	0	0	0	0	0
48	0	0	0	0	0
49	0	0	0	0	0
50	0	0	0	0	0
51	0	0	0	0	0
52	0	0	0	0	0
53	0	0	0	0	0
54	0	0	0	0	0
55	0	0	0	0	0
56	0	0	0	0	0
57	0	0	0	0	0
58	0	0	0	0	0
59	0	0	0	0	0
60	0	0	0	0	0
61	0	0	0	0	0
62	0	0	0	0	0
63	0	0	0	0	0
64	0	0	0	0	0
65	0	0	0	0	0
66	0	0	0	0	0
67	0	0	0	0	0
68	0	0	0	0	0
69	0	0	0	0	0
70	0	0	0	0	0
71	0	0	0	0	0
72	0	0	0	0	0
73	0	0	0	0	0
74	0	0	0	0	0
75	0	0	0	0	0
76	0	0	0	0	0
77	0	0	0	0	0
78	0	0	0	0	0

79	0	0	0	0	0
80	0	0	0	0	0
81	0	0	0	0	0
82	0	0	0	0	0
83	0	0	0	0	0
84	0	0	0	0	0
85	0	0	0	0	0
86	0	0	0	0	0
87	0	0	0	0	0
88	0	0	0	0	0
89	0	0	0	0	0
90	0	0	0	0	0
91	0	0	0	0	0
92	0	0	0	0	0
93	0	0	0	0	0
94	0	0	0	0	0
95	0	0	0	0	0
96	0	0	0	0	0
97	0	0	0	0	0
98	0	0	0	0	0
99	0	0	0	0	0
100	0	0	0	0	0
101	0	0	0	0	0
102	0	0	0	0	0
103	0	0	0	0	0
104	0	0	0	0	0
105	0	0	0	0	0
106	0	0	0	0	0
107	0	0	0	0	0
108	0	0	0	0	0
109	0	0	0	0	0
110	0	0	0	0	0
111	0	0	0	0	0
112	0	0	0	0	0
113	0	0	0	0	0
114	0	0	0	0	0
115	0	0	0	0	0
116	0	0	0	0	0
117	0	0	0	0	0
118	0	0	0	0	0
119	0	0	0	0	0
120	0	0	0	0	0
121	0	0	0	0	0

122	0	0	0	0	0
123	0	0	0	0	0
124	0	0	0	0	0
125	0	0	0	0	0
126	0	0	0	0	0
127	0	0	0	0	0
128	0	0	0	0	0
129	0	0	0	0	0
130	0	0	0	0	0
131	0	0	0	0	0
132	0	0	0	0	0
133	0	0	0	0	0
134	0	0	0	0	0
135	0	0	0	0	0
136	0	0	0	0	0
137	0	0	0	0	0
138	0	0	0	0	0
139	0	0	0	0	0
140	0	0	0	0	0
141	0	0	0	0	0
142	0	0	0	0	0
143	0	0	0	0	0
144	0	0	0	0	0
145	0	0	0	0	0
146	0	0	0	0	0
147	0	0	0	0	0
148	0	0	0	0	0
149	0	0	0	0	0
150	0	0	0	0	0
151	0	0	0	0	0
152	0	0	0	0	0
153	0	0	0	0	0
154	0	0	0	0	0
155	0	0	0	0	0
156	0	0	0	0	0
157	0	0	0	0	0
158	0	0	0	0	0
159	0	0	0	0	0
160	0	0	0	0	0
161	0	0	0	0	0
162	0	0	0	0	0
163	0	0	0	0	0
164	0	0	0	0	0

165	0	0	0	0	0
166	0	0	0	0	0
167	0	0	0	0	0
168	0	0	0	0	0
169	0	0	0	0	0
170	0	0	0	0	0
171	0	0	0	0	0
172	0	0	0	0	0
173	0	0	0	0	0
174	0	0	0	0	0
175	0	0	0	0	0
176	0	0	0	0	0
177	0	0	0	0	0
178	0	0	0	0	0
179	0	0	0	0	0
180	0	0	0	0	0
181	0	0	0	0	0
182	0	0	0	0	0
183	0	0	0	0	0
184	0	0	0	0	0
185	0	0	0	0	0
186	0	0	0	0	0
187	0	0	0	0	0
188	0	0	0	0	0
189	0	0	0	0	0
190	0	0	0	0	0
191	0	0	0	0	0
192	0	0	0	0	0
193	0	0	0	0	0
194	0	0	0	0	0
195	0	0	0	0	0
196	442	440	303	0	0
197	1104	1055	963	0	0
198	1647	1615	1519	9	0
199	1664	1604	1696	203	0
200	1385	1362	1440	545	0
201	1158	1123	1241	851	0
202	933	939	1048	882	0
203	748	760	898	716	0
204	590	591	745	543	0
205	479	474	661	421	0
206	376	362	566	322	0
207	286	283	529	239	0

208	203	198	494	184	0
209	147	149	488	135	0
210	103	102	478	100	0
211	80	77	463	71	0
212	56	56	475	53	0
213	43	41	486	36	0
214	35	30	491	22	0
215	24	21	489	16	0
216	16	12	492	13	1
217	10	7	488	12	6
218	9	5	501	12	9
219	7	5	491	9	11
220	6	6	509	7	10
221	3	5	507	6	10
222	2	4	511	7	9
223	2	3	491	7	8
224	3	3	472	5	8
225	2	3	450	5	9
226	2	2	435	5	8
227	1	1	408	8	9
228	0	0	384	9	8
229	0	0	357	11	9
230	1	0	343	10	7
231	2	1	330	11	8
232	2	1	302	11	8
233	1	1	271	14	8
234	1	0	237	16	6
235	1	1	224	18	6
236	1	0	205	17	6
237	0	0	190	20	7
238	1	0	180	22	7
239	1	0	175	31	7
240	1	0	166	34	7
241	0	0	160	35	8
242	0	0	156	43	11
243	0	0	157	48	14
244	0	0	149	58	16
245	0	0	140	59	20
246	0	0	143	66	22
247	0	0	148	71	24
248	0	0	148	84	25
249	0	0	142	91	33
250	0	0	134	94	40

251	0	0	136	101	46
252	0	0	131	112	48
253	0	0	124	126	54
254	0	0	112	130	56
255	0	0	107	126	67
256	0	0	107	127	68
257	0	0	106	127	77
258	0	0	97	125	75
259	0	0	85	121	74
260	0	0	68	114	75
261	0	0	56	120	73
262	0	0	51	117	79
263	0	0	48	123	75
264	0	0	46	108	73
265	0	0	39	107	71
266	0	0	34	93	68
267	0	0	27	89	64
268	0	0	19	75	61
269	0	0	13	73	60
270	0	0	10	73	59
271	0	0	11	65	55
272	0	0	10	56	48
273	0	0	8	43	43
274	0	0	6	39	36
275	0	0	5	35	36
276	0	0	3	27	32
277	0	0	3	26	28
278	0	0	3	18	22
279	0	0	3	17	18
280	0	0	3	13	20
281	0	0	2	11	18
282	0	0	3	9	18
283	0	0	3	6	13
284	0	0	4	5	10
285	0	0	3	5	7
286	0	0	2	5	9
287	0	0	2	5	10
288	0	0	1	3	12
289	0	0	2	2	8
290	0	0	1	2	6
291	0	0	1	3	4
292	0	0	1	3	4
293	0	0	0	2	3

294	0	0	0	1	2
295	0	0	0	1	1
296	0	0	0	1	2
297	0	0	0	1	2
298	0	0	0	0	1
299	0	0	0	0	2
300	0	0	0	0	3
301	0	0	0	0	3
302	0	0	0	0	2
303	0	0	0	0	1
304	0	0	0	0	1
305	0	0	0	0	2
306	0	0	0	0	1
307	0	0	0	0	0
308	0	0	0	0	0
309	0	0	0	0	0
310	0	0	0	0	0
311	0	0	0	0	0
312	0	0	0	0	0
313	0	0	0	0	0
314	0	0	0	0	0
315	0	0	0	0	0
316	0	0	0	0	0
317	0	0	0	0	0
318	0	0	0	0	0
319	0	0	0	0	0
320	0	0	0	0	0
321	0	0	0	0	0
322	0	0	0	0	0
323	0	0	0	0	0
324	0	0	0	0	0
325	0	0	0	0	0
326	0	0	0	0	0
327	0	0	0	0	0
328	0	0	0	0	0
329	0	0	0	0	0
330	0	0	0	0	0
331	0	0	0	0	0
332	0	0	0	0	0
333	0	0	0	0	0
334	0	0	0	0	0
335	0	0	0	0	0
336	0	0	0	0	0

337	0	0	0	0	0
338	0	0	0	0	0
339	0	0	0	0	0
340	0	0	0	0	0
341	0	0	0	0	0
342	0	0	0	0	0
343	0	0	0	0	0
344	0	0	0	0	0
345	0	0	0	0	0
346	0	0	0	0	0
347	0	0	0	0	0
348	0	0	0	0	0
349	0	0	0	0	0
350	0	0	0	0	0
351	0	0	0	0	0
352	0	0	0	0	0
353	0	0	0	0	0
354	0	0	0	0	0
355	0	0	0	0	0
356	0	0	0	0	0
357	0	0	0	0	0
358	0	0	0	0	0
359	0	0	0	0	0
360	0	0	0	0	0
361	0	0	0	0	0
362	0	0	0	0	0
363	0	0	0	0	0
364	0	0	0	0	0
365	0	0	0	0	0
366	0	0	0	0	0
367	0	0	0	0	0
368	0	0	0	0	0
369	0	0	0	0	0
370	0	0	0	0	0
371	0	0	0	0	0
372	0	0	0	0	0
373	0	0	0	0	0
374	0	0	0	0	0
375	0	0	0	0	0
376	0	0	0	0	0
377	0	0	0	0	0
378	0	0	0	0	0
379	0	0	0	0	0

380	0	0	0	0	0
381	0	0	0	0	0
382	0	0	0	0	0
383	0	0	0	0	0
384	0	0	0	0	0
385	0	0	0	0	0
386	0	0	0	0	0
387	0	0	0	0	0
388	0	0	0	0	0
389	0	0	0	0	0
390	0	0	0	0	0
391	0	0	0	0	0
392	0	0	0	0	0
393	0	0	0	0	0
394	0	0	0	0	0
395	0	0	0	0	0
396	0	0	0	0	0
397	0	0	0	0	0
398	0	0	0	0	0
399	0	0	0	0	0
400	0	0	0	0	0
401	0	0	0	0	0
402	0	0	0	0	0
403	0	0	0	0	0
404	0	0	0	0	0
405	0	0	0	0	0
406	0	0	0	0	0
407	0	0	0	0	0
408	0	0	0	0	0
409	0	0	0	0	0
410	0	0	0	0	0
411	0	0	0	0	0
412	0	0	0	0	0
413	0	0	0	0	0
414	0	0	0	0	0
415	0	0	0	0	0
416	0	0	0	0	0
417	0	0	0	0	0
418	0	0	0	0	0
419	0	0	0	0	0
420	0	0	0	0	0
421	0	0	0	0	0
422	0	0	0	0	0

423	0	0	0	0	0
424	0	0	0	0	0
425	0	0	0	0	0
426	0	0	0	0	0
427	0	0	0	0	0
428	0	0	0	0	0
429	0	0	0	0	0
430	0	0	0	0	0
431	0	0	0	0	0
432	0	0	0	0	0
433	0	0	0	0	0
434	0	0	0	0	0
435	0	0	0	0	0
436	0	0	0	0	0
437	0	0	0	0	0
438	0	0	0	0	0
439	0	0	0	0	0
440	0	0	0	0	0
441	0	0	0	0	0
442	0	0	0	0	0
443	0	0	0	0	0
444	0	0	0	0	0
445	0	0	0	0	0
446	0	0	0	0	0
447	0	0	0	0	0
448	0	0	0	0	0
449	0	0	0	0	0
450	0	0	0	0	0
451	0	0	0	0	0
452	0	0	0	0	0
453	0	0	0	0	0
454	0	0	0	0	0
455	0	0	0	0	0
456	0	0	0	0	0
457	0	0	0	0	0
458	0	0	0	0	0
459	0	0	0	0	0
460	0	0	0	0	0
461	0	0	0	0	0
462	0	0	0	0	0
463	0	0	0	0	0
464	0	0	0	0	0
465	0	0	0	0	0

466	0	0	0	0	0
467	0	0	0	0	0
468	0	0	0	0	0
469	0	0	0	0	0
470	0	0	0	0	0
471	0	0	0	0	0
472	0	0	0	0	0
473	0	0	0	0	0
474	0	0	0	0	0
475	0	0	0	0	0
476	0	0	0	0	0
477	0	0	0	0	0
478	0	0	0	0	0
479	0	0	0	0	0
480	0	0	0	0	0
481	0	0	0	0	0
482	0	0	0	0	0
483	0	0	0	0	0
484	0	0	0	0	0
485	0	0	0	0	0
486	0	0	0	0	0
487	0	0	0	0	0
488	0	0	0	0	0
489	0	0	0	0	0
490	0	0	0	0	0
491	0	0	0	0	0
492	0	0	0	0	0
493	0	0	0	0	0
494	0	0	0	0	0
495	0	0	0	0	0
496	0	0	0	0	0
497	0	0	0	0	0
498	0	0	0	0	0
499	0	0	0	0	0
500	0	0	0	0	0

File S2. R routine to correct abundance when populations are partly in noise.

R code

The code below describes how to implement an R routine to correct the abundance of picoplanktonic populations based on their red fluorescence distribution. All libraries used here are freely available from R repositories. The input file used in this examples is named as Pro_C6.txt (See input file example File S1). This file has been created by exporting FL3 (chlorophyll) histogram from the Flowing Software (http://www.flowingsoftware.com) combining different samples into a single file. The first column contains the channel number and each following column corresponds to a different sample with rows corresponding to cell counts in each channel. Such a file could be created with any flow cytometry software. After running the `cyto_plot` function, a pdf output file is created named "Pro_C6.txt 1.0 .pdf" which contains all histograms from the input file (see File S3)and the file statistics (sample, uncorrected and corrected total cell abundance) are available as a data frame in the R session (see example at bottom of this file)

```
# Example of use of cyto_plot function (run first the R code below to define the necessary functions)
stats_Pro_C6<-cyto_plot("Pro_C6.txt", decades_C6, channel_min_C6, xmin_C6, xmax_C6)

# Example of statistics output

  sample                 cell_tot       cell_tot_correc
1 sample135_C6_PRO_5m    134            cells in noise
2 sample136_C6_PRO_50m   111            cells in noise
3 sample137_C6_PRO_110m  13072          20240
4 sample138_C6_PRO_130m  3598           no correction
5 sample139_C6_PRO_170m  2211           no correction
```

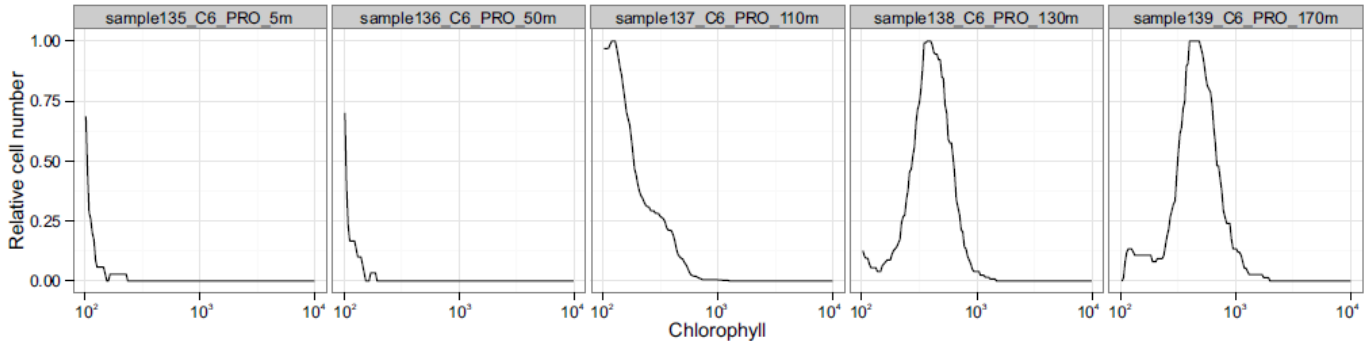
R code

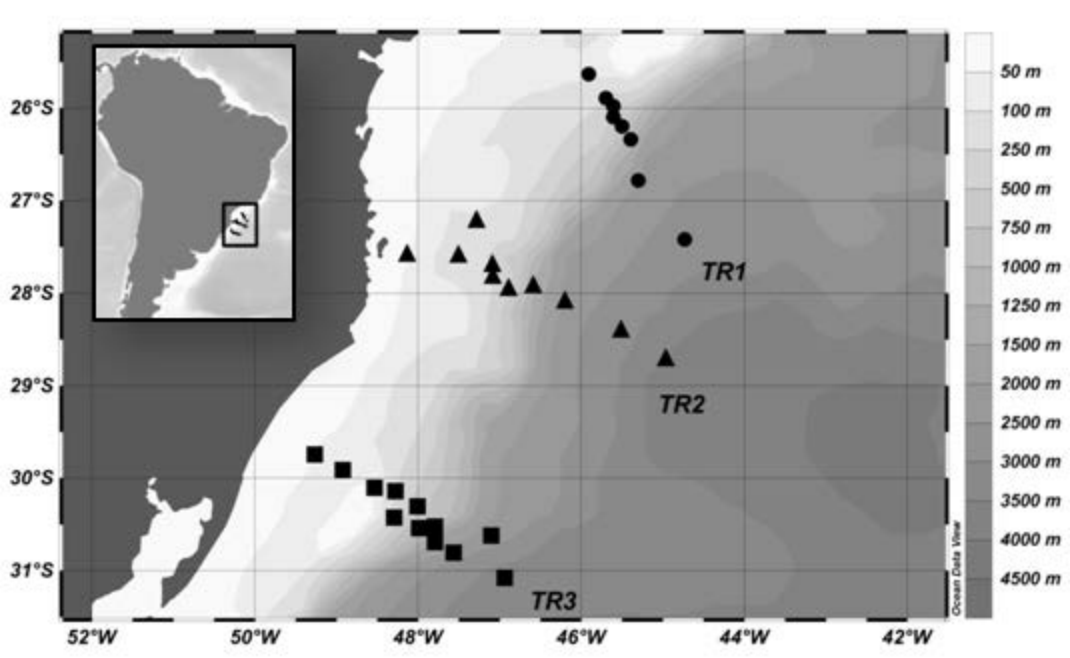
```
# Install libraries
library("ggplot2")
library("reshape2")
library("plyr")
library("scales")
require(grid)

# Set the working directory where the files are located
setwd ("C:/My Documents/cytometry data/")

# Define basic parameters
decades_Canto = 5
decades_C6 = 7
channel_min_Canto = 100
channel_min_C6 = 214
xmin_Canto = 10
xmin_C6 = 1000
xmax_Canto = 10000
xmax_C6 = 100000
channel_max = 500
point <- format_format(big.mark = "", decimal.mark = ".", scientific = TRUE)
# ----------------------------------------------------------
# cell_correct(channel, cell_number, cell_smooth)
# Arguments
#       channel : vector containing the channels (from 1 to 500 in the present case)
#       cell_number : vector containing cell abundance in each channel
#       cell_smooth : vector containing smoothed cell abundance in each channel
# Description
# This function determines in which case we are ("no correction", "cells in noise" or "correction")and
return the corrected cell abundance in the latter case.

cell_correct<-function(channel, cell_number, cell_smooth)
{ df<-data.frame(channel, cell_number, cell_smooth)  # create a data frame
i_min<-which.min(channel)                            # determine the minimum channel
```

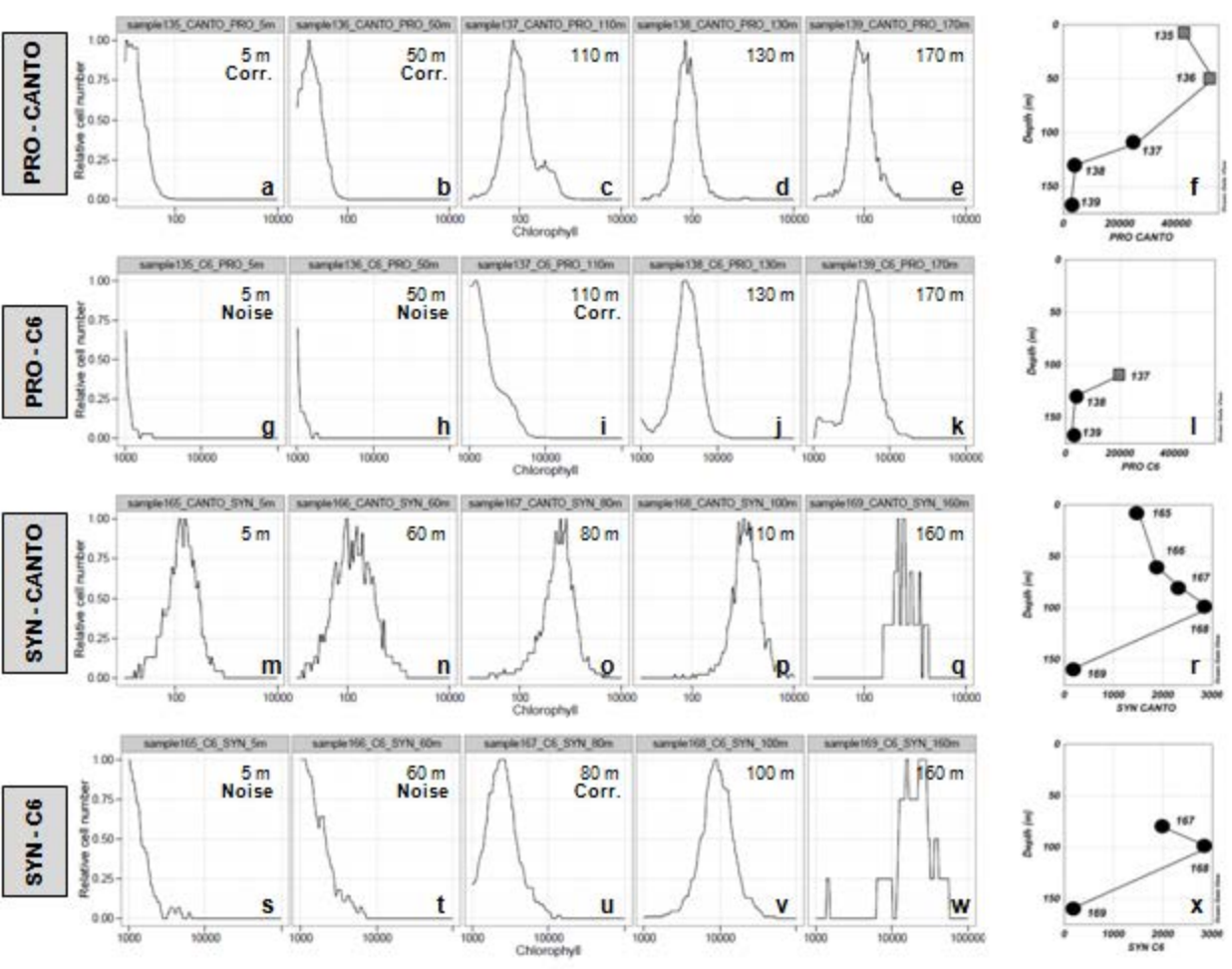
```
i_max<-which.max(channel)                          # determine the maximum channel
i_cell_max<-which.max(cell_smooth)                 # determine in which channel is the histogram mode

# "no correction" : cell abundance in the first channel is 5 times lower than abundance at the maximum
of the histogram
if (cell_smooth[i_cell_max]>5*cell_smooth[i_min])    {cell_correct<-"no correction"}
# "cells in noise" : maximum of cell abundance is in the first channel
  else {if (i_cell_max==i_min)
        {cell_correct<-"cells in noise"}
# "correction" : all the other cases, we then apply a correction by computing the total cell abundance
as twice the number of cells in the channels right of the histogram maximum
  else
        {cell_correct<-2*sum(cell_number[i_cell_max:i_max])}
        }
  return (cell_correct)
}

# ------------------------------------------------------------
#cyto_plot(file_name,decades,channel_min,xmin,xmax)
# Arguments
#       file_name : name of input file containing the different samples (see File S1)
#       decades : number of logarithmic decades of the flow cytometer (e.g. 7 for C6)
#       channel_min : threshold channel for the histogram (depends on fcm acquisition settings)
#       xmin : linear value corresponding to the threshold channel
#       xmax : linear value corresponding to the maximum channel
# Description
# This function plots a set of histograms for the input samples,saves the graphics as a pdf file and
compute the total cell abundance indicating whether a corrections is needed or not.  It returns a
dataframe containing three columns : sample, cell_tot, cell_tot_correc (see top of this file for an
example)

cyto_plot<-function(file_name,decades,channel_min,xmin,xmax)
{       channel_max = 500 # this the number of channels provided as output of the Flowing Software
        histo<- read.delim(file_name)
        histo<- histo[histo$channel>=channel_min,]
        histo_melt<-  melt(histo, id.vars=c("channel"),variable.name = "sample", value.name =
"cell_number")
```

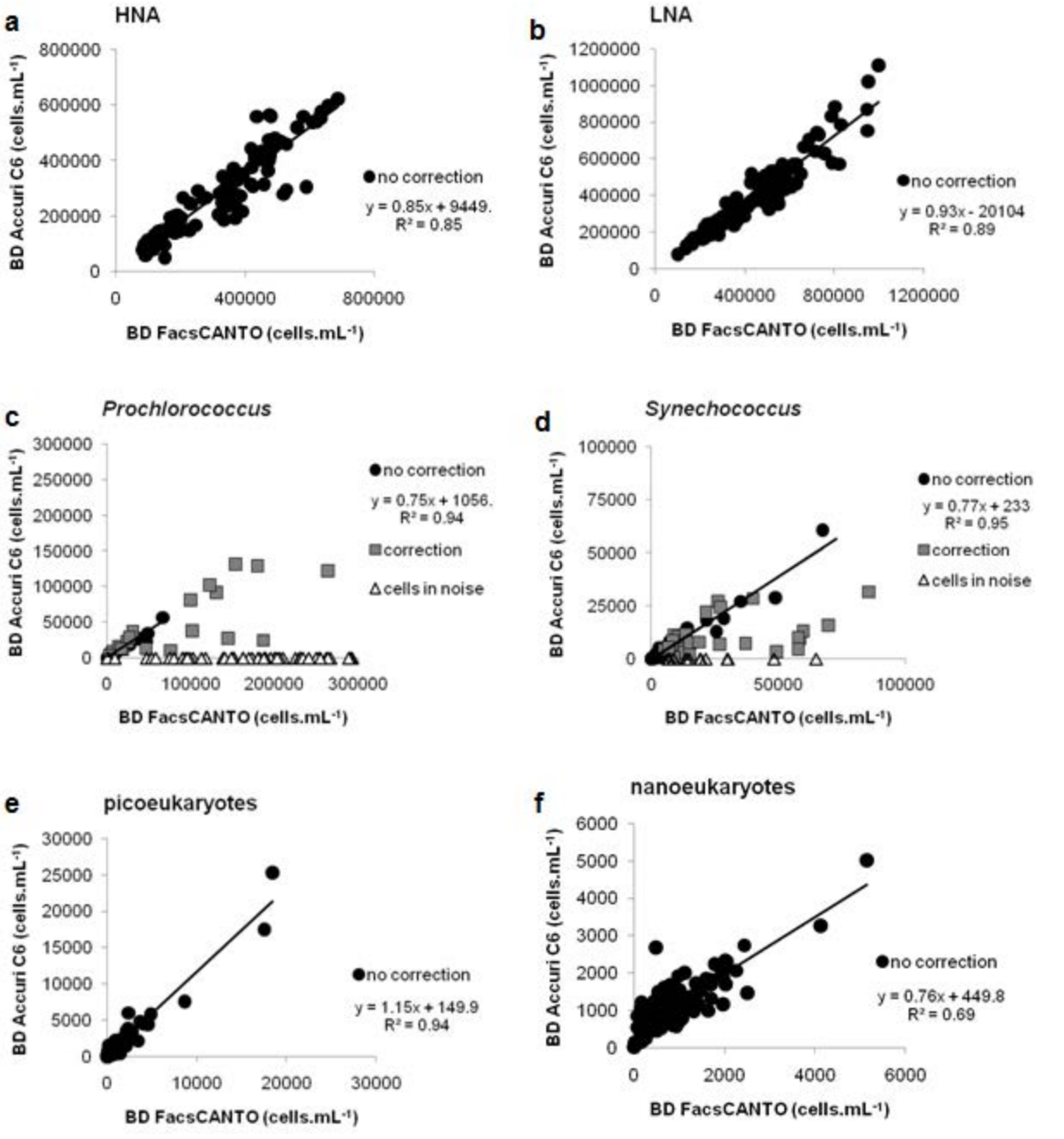
```
# smooth histogram using default R smoothing function
        histo_melt<- ddply(histo_melt,c("sample"), transform,
cell_smooth=as.vector(smooth(cell_number)))
# normalize histogram so that maximum abundance is equal to 1
        histo_melt<- ddply(histo_melt,c("sample"), transform, cell_norm=cell_smooth/max(cell_smooth))
# transform log channel to linear scale for plotting
        if (decades==5)
                {histo_melt<- ddply(histo_melt,c("sample"), transform,
fluo=(10^5)^(channel/channel_max))}
        else
                {histo_melt<- ddply(histo_melt,c("sample"), transform,
fluo=(10^7)^(channel/channel_max))}
# plots histograms using 5 columns
        histo_plot<-ggplot(histo_melt, aes(fluo,cell_norm)) + geom_line() + theme_bw () + facet_wrap(~
sample, nrow=21, ncol=5) + xlab("Chlorophyll")+ylab("Relative cell number") +
scale_x_log10(limits=c(xmin,xmax), labels=point)
# save plots as pdf
ggsave(plot=histo_plot, filename=paste(file_name," 1.0 .pdf",sep=""),width = 15, height = 4, scale=2,
units="cm")
# compute uncorrected and corrected total cell number calling the cell_correct function defined above
        stats<-ddply(histo_melt,c("sample"),summarise,
cell_tot=sum(cell_number),cell_tot_correc=cell_correct(channel,cell_number,cell_smooth))
        print(paste("# of decades:",decades,"minimum channel : ",channel_min, "xmin : ",  xmin, " xmax
: ",    xmax))
        print (paste("File : ",file_name))
        stats
        return (stats)
}
```

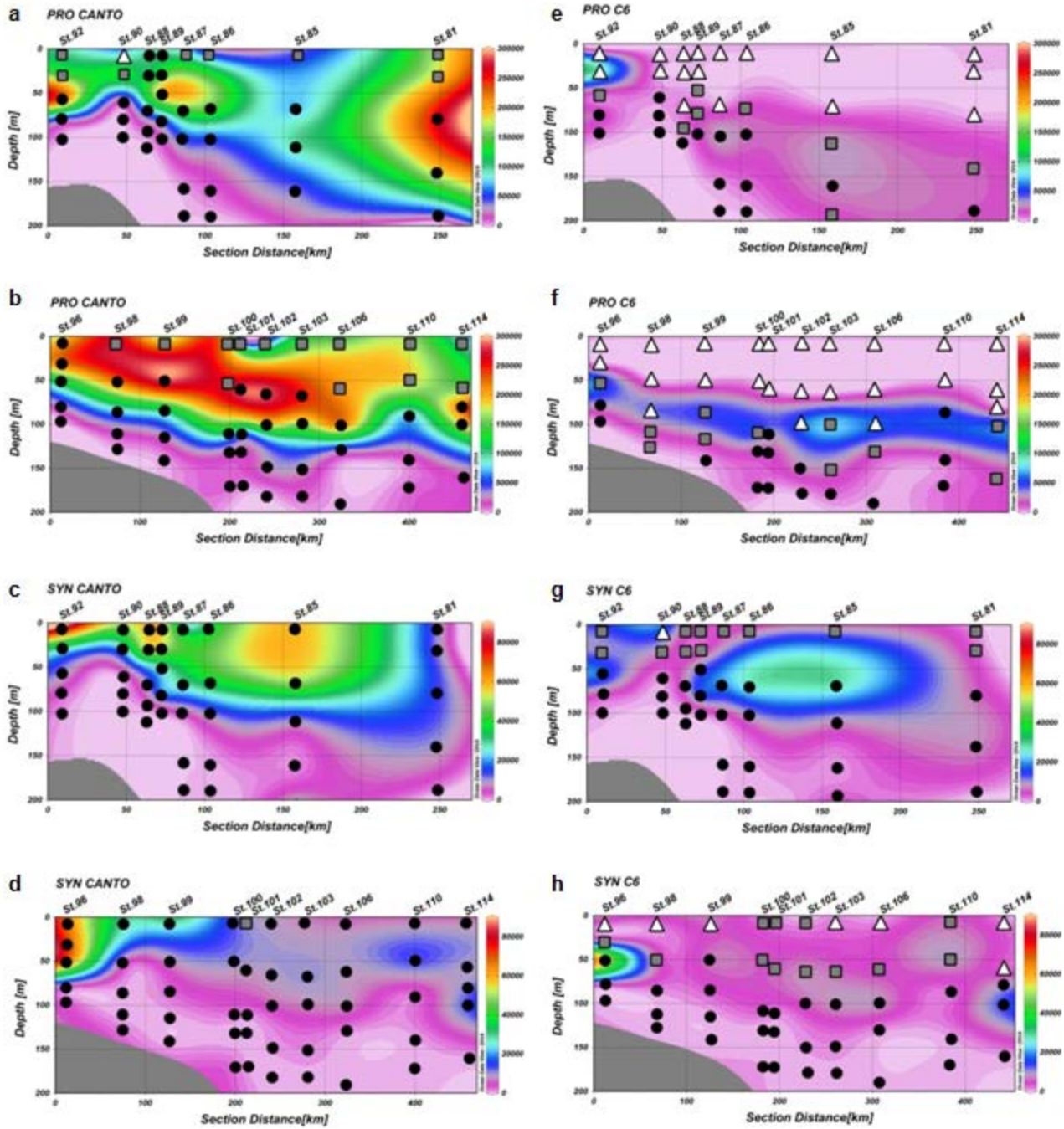
File S3. Example of output file for R routine.




26°S
27°S
28°S
29°S
30°S
31°S
52°W
50°W
48°W
46°W
44°W
42°W
TR1
TR2
TR3
50 m
100 m
250 m
500 m
750 m
1000 m
1250 m
1500 m
2000 m
2500 m
3000 m
3500 m
4000 m
4500 m
Ocean Data View


PRO - CANTO
PRO - C6
SYN - CANTO
SYN - C6
Relative cell number
Chlorophyll
Depth (m)
PRO CANTO
PRO C6
SYN CANTO
SYN C6
5 m Corr.
50 m Corr.
110 m
130 m
170 m
5 m Noise
50 m Noise
110 m Corr.
60 m
80 m
160 m
100 m
a
b
c
d
e
f
g
h
i
j
k
l
m
n
o
p
q
r
s
t
u
v
w
x


a
HNA
BD Accuri C6 (cells.mL-1)
BD FacsCANTO (cells.mL-1)
no correction
y = 0.85x + 9449.
R² = 0.85
b
LNA
BD Accuri C6 (cells.mL-1)
BD FacsCANTO (cells.mL-1)
no correction
y = 0.93x - 20104
R² = 0.89
c
Prochlorococcus
BD Accuri C6 (cells.mL-1)
BD FacsCANTO (cells.mL-1)
no correction
y = 0.75x + 1056.
R² = 0.94
correction
cells in noise
d
Synechococcus
BD Accuri C6 (cells.mL-1)
BD FacsCANTO (cells.mL-1)
no correction
y = 0.77x + 233
R² = 0.95
correction
cells in noise
e
picoeukaryotes
BD Accuri C6 (cells.mL-1)
BD FacsCANTO (cells.mL-1)
no correction
y = 1.15x + 149.9
R² = 0.94
f
nanoeukaryotes
BD Accuri C6 (cells.mL-1)
BD FacsCANTO (cells.mL-1)
no correction
y = 0.76x + 449.8
R² = 0.69


a
PRO CANTO
St.92
St.90
St.88
St.89
St.87
St.86
St.85
St.81
Depth [m]
Section Distance[km]
b
PRO CANTO
St.96
St.98
St.99
St.100
St.101
St.102
St.103
St.106
St.110
St.114
Depth [m]
Section Distance[km]
c
SYN CANTO
Depth [m]
Section Distance[km]
d
SYN CANTO
Depth [m]
Section Distance[km]
e
PRO C6
Depth [m]
Section Distance[km]
f
PRO C6
Depth [m]
Section Distance[km]
g
SYN C6
Depth [m]
Section Distance[km]
h
SYN C6
Depth [m]
Section Distance[km]