The major goal of this paper is to assess the response of two major calcifying taxa (coccolithophorids and foraminifera) to higher (future) atmospheric pCO2 levels. To do so the authors selected two sediment cores and investigated the abundance and some calcification parameters of a few dominant species of coccolithophorids and foraminifera (lith size and (size normalized) weight, respectively) for the period 3.3 through 2.6 Myr BP. pCO2 levels for that period are estimated to have been around 380 to 415 ppm (not so much different from today).

I enjoyed reading the manuscript. It is well written but has a few flaws and, in my view, can be improved substantially by addressing the comments and questions I raise below.
1) The first question that came to my mind was why the authors didn’t choose a period with a higher pCO2, such as, e.g. expected for the end of this century at a business as usual scenario? Although the authors mention some good arguments “continental placement and vegetation similar to modern” and going from the “end of the Pliocene Warm Period into the subsequent interval of global cooling”, the rationale didn’t become completely clear to me. Of course, our world is getting warmer but we can look at the data backwards. However, if you look at fig. 4. I do not see a long term cooling trend. Maybe they can state why they didn’t go for e.g. the PETM. Also, by choosing a period of assumed atmospheric pCO2 levels as today, the study would benefit from some comparison with today’s assemblages or at least with a more thorough discussion of the experiments by Langer et al (several papers) or experimental data on weight changes as a function of carbonate chemistry. For instance, see fig. 8 in Bijma et al. (1999). Reassessing foraminiferal stable isotope geochemistry: Impact of the oceanic carbonate system (experimental results). Use of Proxies in Paleoceanography: Examples from the South Atlantic. G. Fischer and G. Wefer. Berlin, Heidelberg, Springer-Verlag: 489-512.). Interestingly, G. bulloides did not show a trend (not plotted, I’m afraid, sorry).

2) Secondly, what was the rationale for choosing the core sites? Just the difference in temperature or where there additional reasons? The manuscript would benefit if the authors indicate how these two different sites are going to help them to disentangle the different factors that control the “calcification parameters”

3) It is important to say that we do NOT know pH nor the saturation for the period of investigation, but that we infer lower pH based on modeled higher atmospheric pCO2. This is a reasonable assumption but none-the-less an assumption. Furthermore it is important to note that the most important difference between all paleo OA scenarios and the situation today is that the rate of change today is unprecedented and that pH and the saturation state of the ocean with respect to calcium carbonate are declining in concert. pH and saturation state were probably decoupled during the Pliocene and the
INHG when glacial-interglacial (G-IG) changes of ca. 40 ppm in atmospheric pCO2 occurred, meaning that pH could be lower but that the saturation state remains constant. This is important information that should be added.

Method section:

4) on p. 6843; line 2-4: “...Changes in the relative contribution of foraminifers to coccolithophores to sediment production were calculated by dividing foraminiferal MAR by carbonate MAR. ...” But what about pelagic bivalves, ostracods, pteropods, CaCO3 debris,...? Please comment if this can be neglected (or not)? In my view, this might lead to a bias especially at site 999 with 40-60% fragmentation! Hence, the foram contribution may be underestimated....and this would not be the result of in vivo and surface water conditions, but controlled by bottom water saturation state (post mortum).

5) With regard to the size normalized weight (SNW) proxy, the authors are referred to Bijma at al., 2002. Impact of the ocean carbonate chemistry on living foraminiferal shell weight: Comment on “Carbonate ion concentration in glacial-age deep waters of the Caribbean Sea” by W. S. Broecker and E. Clark. G3, v.3, nr. 11. Although, an important part is dependent on the in vivo carbonate chemistry, is a large share(the most?) controlled by bottom water saturation values (post mortum). This was, in fact, the reason why SNW was originally introduced by Pat Lohmann, to reconstruct bottom water [CO32-].

6) on the bottom of p. 6844 line 27: “…Unexpectedly, the deep dweller G. puncticulata shows significantly higher weight variability than the surface dweller”. This observation is left unanswered but may have to do with structural test features. Some species may be more resistant to dissolution than others before they completely break down and disappear from the archive (but contribute to the CaCO3 MAR). Pat Lohmann showed at a conference that on a depth transect at Sierra rise different species dissolve and disintegrate at different depths. Although, all are made up of calcium carbonate, G. ruber disappears first, G. sacculifer is still present in deeper cores and P. obliquiloculata
is the last to disappear from the record. He argued that structural differences are responsible and that every species has its own "lysocline depth". I'm not sure if it ever was published but this can be investigated.

Discussion section.

7) p. 6845; line 19: I would spell INHG out and write Intensification of Northern Hemisphere Glaciation.

8) P. 6847; line 4-6: ...."we would have expected smaller foraminiferal test weight changes in the Pliocene (half the magnitude observed here), if this was the main or even sole driver.".... Here you are assuming that SNW is only driven by surface water carbonate chemistry when the forams were alive....this is not the case. As argued above, in most cases the major driver of SNW is the pore water water saturation state. If you have exceptional preservation and find e.g. the aragonitic shells of pteropods in your cores, SNW may be used as a tracer of in vivo surface water chemistry (at least for sfc. dwellers).

Line 7-9:...."As for the carbonate MAR data, the contrasting records at these two sites suggest that regional drivers are responsible for the changes in test weight and not changes in pCO2.".... Then the authors go into explaining the different temperature regimes.... I'm missing a discussion on pore water [CO32-] and saturation state...(see above).

For the discussion on the coccolithophorids (later on in the msc) the authors use the argument of different genotypes to explain the observed size trends. Why not for the forams? There are many papers out, documenting that all planktonic foraminifera have multiple genotypes (with the notable exception of G. sacculifer, which seems to be genetically identical, no matter where in the world or when it is collected).

On p. 6848; line 5-9, the authors say: “The different size trends seen at the two sites represent different proportions of species/genotypes/ecophenotypes.... their relation-
ship to environmental change is uncertain." Then the authors continue . . . ” It is unlikely that these trends reflect dissolution because fragmentation is relatively stable across the record at Site 607 and improves at Site 999 while large liths disappear which are less prone to dissolution.” But maybe this disappearance of the larger liths is not related to bottom water saturation state but due to the “overall trend towards smaller liths through the Pliocene into the Quaternary”.

Conclusion.

In the last line of the conclusions, the authors say: “that rates of change rather than absolute pCO2 values might drive a calcification response in these groups.” This is a very important conclusion but this was not mentioned or developed at all in the main text (discussion). This should be done.

To conclude: The interval chosen seems to be relatively stable in terms pCO2. Similar to or a bit higher than today, with G-IG shifts in atmospheric pCO2 of ca. 40 ppm. Temperature, while showing large fluctuations, does not show a cooling trend over the time interval from "warm towards the Intensification of Northern Hemisphere Glaciation". The (G-IG) variability of the calcification parameters in the data is larger than any trend over that period. I guess the main message is that calcareous plankton responds to G-IG change very similar as during the late Pleistocene where atmospheric pCO2 was more than 100 ppm lower and the G-IG difference in pCO2 was more than twice as large. This is an interesting conclusion.

Some major revisions are needed and figures 2 and 3 should be scaled up.

Interactive comment on Biogeosciences Discuss., 10, 6839, 2013.