

A point-by-point response to the reviewer#2 comments

Review by Ed Sobel of Tamer et al. "The Need for Fission Track Data Transparency"

This study presents an inter-operator study on the quality of fission track analyses; the influence of the operator's skill and judgement on the results are the main focus. After revisions (English usage and more precise descriptions), this will be a very useful study for the fission track community.

I have read Ray Donelick's comments and generally agree with his points. However, I think that the required revisions are primarily textural and therefore are minor rather than major/reject. I would modify his list of 6+ essential steps: 7 (collect photos) has to come before 6b (LA-ICP-MS). I agree with his wish to have more precise statements. Yes, the number of analysts is small, but quantitative statements would be more useful than phrases such as 'show similarities'. The criteria for suitable, unsuitable, or borderline needs to be defined - the text is presently quite short, so there is space to elaborate. Fig. 3 shows data histograms. A more quantitative way to compare histograms would be helpful.

Our Reply: We thank the reviewer for the comments and suggestions. We implemented all the suggestions, including reorienting the figures to portrait format, inclusion of Table S1 to the main text and other textural edits.

47 Track density can vary by up to 35% if the grain is oriented without the c-axis in the viewing plane

Rephrase: if the grain is not oriented with the c-axis in the viewing plane

Our Reply: Edited.

58-59 and suffer edge effects from sampling a 4pi region that is variable.

rephrase-perhaps: that hosts variable U concentration

Our Reply: Edited.

A follow-on conclusion would be that software packages should be modified to provide the option of automatically drawing an ROI that is 10 micron inside the edge of the grain.

Our Reply: Next version of FastTracks (Ling et al. in prep) will include this feature.

~80 Please insert a sentence here stating where the images used in this study can be viewed. Presently one has to reach line 326 to get this information.

Our Reply: Edited.

Table 1 - I don't understand the suitable grain selection rate. Analyst 1 selected 22 grains as suitable and has a 100% rate; analyst 8: 35 grains, 100%. I find the description of the selection criteria to be confusing.

Our Reply: We implemented new tables and improved the text in section 3.5 for further clarifications.

Rate is the wrong term. A number of grains were selected - this is a scalar unit. Rate implies a speed. This error occurs throughout the ms.

Our Reply: Edited.

Figs. 1, 3 - rearrange so that these are not in landscape format.

Our Reply: Edited.

The figures and tables are well drawn and appropriate.

Our Reply: Thank you for the comment.

Fig. S1 could be included in the main ms. Table S2 needs a title.

Our Reply: Edited.

154-5 Although calibration is an essential step before performing an analysis,

This needs justification. How large were the differences in lengths between groups that did and did not perform calibrations? I note that my (old) microscope is quite stable. The change from one calibration to the next is too small to measurably affect lengths - on the order of 0.02-0.03 microns for a 15 micron length and less for a shorter length, which is within measurement error.

Our Reply: This is an important question, however, the experimental design of our study cannot provide a satisfactory answer. We cannot directly compare mean track lengths because different analysts measured different lengths and different numbers of lengths, inclusion or exclusion of a

few tracks into a data set may drastically change average mean, stdev, min and max values. The only suitable way is to make a comparison on measurements of single track lengths. There is no confined track length measured by all the analysts in this study. In the figure below we show a case comparison of measurements on a single confined track (Figure a,b) by 11 analysts. For this particular track, the apparent lengths (2D) range from ~11.5 to 12.1 μm , where most of the measurements are within 1 sigma of the mean value (red circle) (Figure c). The difference in true lengths (3D) is slightly higher (Figure d) than 2D length due to different considerations on dip angle (Figure b,e). Based on one single track, there is no correlation between those who measured graticule and those who did not. However, we cannot draw a relationship on a single length. There are several other confined track lengths measured by lesser number of analysts that we can compare and not all of them are measured by the analysts who measured graticules. Few length measurements repeated by lesser number of analysts does not constitute a solid ground to answer this question properly.

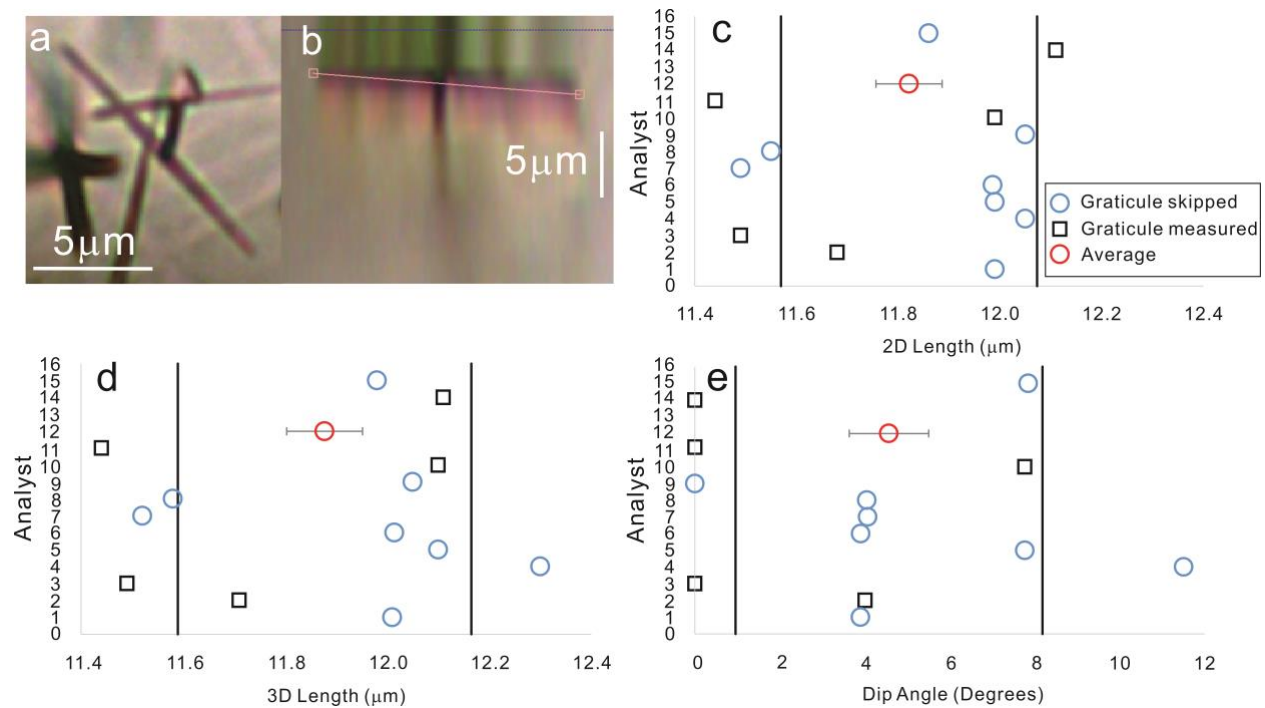


Figure: Comparison of a single confined track length measured by 11 analysts. Confined track length (a), its cross section view (b), comparison of the apparent (2D) length (c), true (3D) length (d) and dip angle (e). Solid lines show 1-sigma intervals from the average value.

185-6: Some of the participants used FastTracks' automatic tools for c-axis orientation and Dpar length measurements.

this sentence doesn't match the section header (3.3 Post-review follow-up and objectivity of the review). It belongs somewhere else.

Our Reply: The sentence is moved to section 3.

189 The density 190 distributions of suitable grains appear to be somewhat more reliable...
more reliable is the wrong phrase. similar? In general, the writing in section 3.4 is rough.

Our Reply: Edited.

section 3.5 new text: "Accordingly, those who admit unsuitable grains at higher rates "
Admit is the wrong word - select?

Our Reply: Edited.

231 ...but inclusion of large defects in the ROI...

a large number? I don't think the size of the individual defects is as relevant as their abundance.

Our Reply: Both number and size of the defects contribute the area that is unable to be analyzed within a given ROI. We edited the sentence.

148 reasulting typo

Our Reply: Edited.

Does the paper address relevant scientific questions within the scope of GChron?

Absolutely

Our Reply: Thank you.

Does the paper present novel concepts, ideas, tools, or data?

The paper presents the novel study in which multiple users analyze the exact same apatite fission track crystals. This leads to observations and conclusions about the causes of variability and the reliability of measurements.

Our Reply: Thank you.

Are the scientific methods and assumptions valid and clearly outlined?

Can be improved - the criteria for grain selection need to be better defined.

Our Reply: Edited.

Are the results sufficient to support the interpretations and conclusions?

It would be better if some qualified observations were quantified, as discussed above.

Our Reply: Edited.

Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

yes

Our Reply: Thank you.

Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

yes

Our Reply: Thank you.

Does the title clearly reflect the contents of the paper?

ok. Donelick's suggested title would be an improvement.

Our Reply: Edited.

Does the abstract provide a concise and complete summary?

ok. Needs to include the word apatite in the 1st sentence.

Our Reply: Edited.

Is the overall presentation well structured and clear?

yes

Our Reply: Thank you.

Is the language fluent and precise?

Some English polishing is needed. The language usage light source preference can be imprecise. An example: "light source preference", which actually seems to mean the decision to use reflected as well as transmitted light. The definition of this phrase needs to be clearly stated.

Our Reply: Edited.

Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Yes - with the exception of the word 'rate', as discussed above.

Our Reply: Edited.

Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

on

Our Reply: We assume the answer is no.

Are the number and quality of references appropriate?

yes

Our Reply: Thank you.

Is the amount and quality of supplementary material appropriate?

yes

Our Reply: Thank you.