DOI: https://doi.org/10.2138/am-2019-7046

Revision 1

Analyst and etching protocol effects on the reproducibility of apatite confined fission-track length measurement, and ambient-temperature annealing at decadal time scales

Murat T. Tamer[1], Ling Chung[2], Richard A. Ketcham[1], Andrew J.W. Gleadow [2]

[1]Jackson School of Geosciences, The University of Texas at Austin, 78712 Austin, TX, USA

[2]School of Earth Sciences, University of Melbourne, 3010 Melbourne, VIC, Australia

1. Abstract

Previous inter-laboratory experiments on confined fission-track length measurements in apatite have consistently reported variation substantially in excess of statistical expectation. There are two primary causes for this variation: (1) differences in laboratory procedures and instrumentation, and (2) personal differences in perception and assessment between analysts. In this study, we narrow these elements down to two categories, etching procedure and analyst bias. We assembled a set of eight samples with induced tracks from four apatite varieties, initially irradiated between 2 and 43 years prior to etching. Two mounts were made containing aliquots of each sample to ensure identical etching conditions for all apatites on a mount. We employed two widely used etching protocols, 5.0M HNO_3 at 20°C for 20s and 5.5M HNO_3 at 21°C for 20s. Sets of track images were then captured by an automated system and exchanged between two analysts, so that measurements could be carried out on the same tracks and etch figures, in the same image data, allowing us to isolate and examine the effects of analyst bias. An additional 5 seconds of etching was then used to evaluate etching behavior at track tips. In total, 8391 confined fission-track length measurements were performed; along with 1480 etch figure length measurements. When the analysts evaluated each other's track selections within the same images for suitability for measurement, the average rejection rate was ~14%. For tracks judged as suitable by both analysts, measurements of 2D and 3D length, dip, and c-axis angle were in excellent agreement, with slightly less dispersion when using the 5.5M etch. Lengths were shorter in the 5.0M-etched mount than the 5.5M-etched one, which we interpret to be caused by more prevalent under-etching in the former, at least for some apatite compositions. After an additional 5s of etching, 5.0M tracks saw greater lengthening and more reduction in dispersion than 5.5M tracks, additional evidence that they were more likely to be under-etched after the initial etching step. Systematic differences between analysts were minimal, with the main exception being likelihood of observing tracks near perpendicular to the crystallographic c axis, which may reflect different use of transmitted versus reflected light when scanning for tracks. Etch figure measurements were more consistent between analysts for the 5.5M etch, though one apatite variety showed high dispersion for both. Within a given etching protocol, each sample reflected a decrease of mean track length with time since irradiation, giving evidence of 0.2-0.3 μm of annealing over year to decade time scales.

Keywords: Fission Track; Etching Procedure; Step Etching; Ambient temperature; Annealing; Confined Track Length.

°Corresponding author: tamer@jsg.utexas.edu

2. Introduction

Fission tracks are radiation damage trails in solid materials produced by fission decay. Natural, spontaneous fission decay of ^{238}U creates fossil or spontaneous tracks, while thermal-neutron induced fission of ^{235}U in nuclear reactors (Meitner and Frisch 1939) creates induced tracks. In apatite the damage trails from fissioning nuclei initially leave a track with a length of ~21 μm (Bhandari et al. 1971; Jonckheere 2003) and a diameter of ~10 nm (Paul and Fitzgerald 1992). Thermal annealing of the radiation damage leads to gradual repair of the crystal structure as a function of time and temperature (Fleischer et al. 1964) which results in shortening of the tracks. Fission tracks are thermochronometers (Wagner 1981); each fission track carries information on the temperatures it has experienced since its formation.

Fission tracks become observable under optical microscopes after being enlarged by a suitable etching protocol. Etching of a polished mineral surface reveals tracks intersecting that surface, and the etchant penetrates into confined fission tracks below the surface through cracks and cleavages (TINCLE) or other surface-intersecting host tracks (TINT) (Bhandari et al. 1971). In any particular sample, the confined track lengths vary due to non-identical etching duration of the individual tracks (Green et al. 1986), crystallographic anisotropy (Green and Durrani 1977; Watt et al. 1984; Donelick et al. 1999), the stochastic nature of nuclear splitting and particle interactions (e.g. Zeigler et al., 2008), and varying degrees of annealing during the thermal history of the host rock. Over geological time scales, in most apatites fossil tracks are erased above 120˚C (Naeser 1981), partially annealed above 60˚C (Gleadow and Duddy, 1981; Wagner et al. 1989) and subject to slow annealing at <60˚C (Gleadow and Duddy 1981; Donelick et al. 1990; Spiegel et al. 2007). The track length distributions reflect the temperature history of the apatite since it last cooled into the ~60-120 °C window, the so-called Partial Annealing Zone (Gleadow et al. 1986). Apatite fission-track modeling uses the information from individual track length measurements from a sample to reconstruct the time and temperature conditions that the sample has undergone (Green et al. 1989; Gallagher 2012; Ketcham 2005). No other geo-thermochronometer provides a comparable level of detail. However, to be a reliable tool for reconstructing thermal histories, length measurements need to be robust and reproducible within and among laboratory groups (Ketcham et al. 2009, 2015, 2018).

2.1 Confined track-length revelation

Fission-track analysis requires the use of optical microscopes, and therefore any feature we observe must be etched. This makes the etching procedure a crucially important part of any fission track study. An etching procedure has three elements, etchant concentration, duration, and temperature, which to some degree can be traded off against each other to achieve roughly equivalent results. Since the advent of fission-track dating, track etching has been understood primarily in terms of a fast etching velocity along the track (v_T) and a slower etching velocity in the bulk mineral (v_B). Standard calculations of etching efficiency (i.e. the fraction of tracks crossing a surface that are revealed by etching) assume these two rates are constant (Fleischer et al. 1964). Under this model, a fully etched confined track was defined as spanning the full extent of the line segment with v_T and a small to negligible amount of the v_B region (Laslett et al. 1984), while tracks that do not extend to the ends of the v_T region are under-etched and tracks extending significantly into the v_B region are over-etched.

How do we know when we etch too little, too much, or just enough? Early analysts defined optimal etching conditions using consecutive step etch experiments (e.g., Watt and Durrani 1985; Laslett et al. 1984; Green et al. 1986; Carlson et al. 1999), where confined track length measurements were performed between short-duration etching steps. As summarized by Jonckheere et al. (2017, their Fig. 1) the pattern observed in most of these experiments is that mean confined lengths would rise quickly and then reach a near-plateau value, after which they would rise slowly or not at all with further etching. Laslett et al. (1984) proposed that the onset of this plateau or linear slow lengthening defines where most revealed tracks are fully etched or slightly over-etched.

However, Jonckheere et al. (2017) show that a constant-v_T line segment model is oversimplified. By conducting step-etching experiments using an image-capture system that allowed them to follow the evolution of individual tracks rather than measuring a random sample after each step, they demonstrated that track etching velocity changes along the length of the track, decreasing continuously as the ends are approached. Importantly, they also showed that the region of enhanced etching in spontaneous and unannealed induced tracks in Durango apatite extends considerably beyond their conventionally accepted lengths (Jonckheere et al. 2017).

2.2 Reproducibility of confined track length and etch figure measurements.

A series of studies has revealed an unexpectedly and persistently poor level of reproducibility in confined track length measurements for both spontaneous and induced tracks at a range of annealing levels (Miller et al. 1990; Barbarand et al. 2003b; Ketcham et al. 2009, 2015, 2018), as well as for etch figure diameters (D_{par}) (Sobel and Seward, 2010; Ketcham et al., 2015, 2018). Variation between measurements from different laboratories consistently exceeds statistical expectation, even when they use the same etching procedures. These disparities could result from a mixture of several factors, such as variation in etching time or temperature; inexact etchant concentration; different types of microscopes, lens configurations, and light sources; analyst biases in finding and selecting tracks as suitable for measurement; and length measurement procedures. Because the measurement process is very analyst-dependent, tracking down which of these factors is dominant, and thus the best candidate for community-wide improvement, is not straightforward. In this study, we narrow down this set of problems to the final stages of the analysis process, in which tracks are etched, identified, evaluated for measurement, and measured, by having two analysts measure the same individual tracks and etch figures in a range of samples etched with two protocols.

2.3 Low-temperature annealing at laboratory time scales

It has been reported that induced fission tracks experience shortening at ambient temperatures on time scales from minutes to months in various apatites (Donelick et al. 1990; Belton 2006; Tamer and Ketcham 2018). If significant annealing of freshly formed induced tracks does occur during the months to decades after irradiation, then this may be another factor causing variation in measurements of standard materials. This phenomenon may also cloud the meaning of the “unannealed” length used to normalize all track lengths for annealing modeling (Laslett and Galbraith 1996; Ketcham 2019).

This short-time-scale annealing may reflect the same process responsible for spontaneous tracks that have only experienced low temperatures over geological time scales being typically ~1-2 μm shorter than induced tracks (Ketcham 2019), or it may reflect a different process. Expanding the experimental database for this phenomenon should help shed light on its cause.

2.4 Aim of this study

In this study, we conduct a series of experiments to address the problems of length measurement reproducibility and low-temperature annealing. Taking a selection of apatites that have experienced ambient-temperature annealing of induced tracks for as many as 36 years, we attempt to isolate the causes of variation by having two experienced analysts measure two aliquots of each, etched with different protocols. We isolate user-specific sources of variation by taking advantage of image-capture software to ensure that each analyst is observing the same tracks in the same way. The image-capture system also enabled us to conduct a follow-on step-etch analysis to help discern the nature of etching rate variation near track tips.

This work explores four main points: 1- How much agreement do two analysts have on a single track-length or etch-figure measurement? 2- What is the difference between two major etching protocols? 3- What is "under-etching", and can we recognize it based solely on the appearance of etched tracks? 4- Do induced tracks anneal at low temperatures on time scales from years to decades?

3. Experimental Details

Figure 1 summarizes the organization of the experiment. Eight apatite samples with induced tracks were selected, including three Durango, two Mud Tank and two Renfrew apatites, and one sample from a population-method-dated granite from King Island, Bass Strait, Australia (Gleadow and Lovering, 1978), which were irradiated from ~2 to ~43 years prior to etching. Sample details are summarized in Table 1. Two epoxy grain mounts were prepared for each sample (Fig. 1A, first row). Procedures and devices used in sample preparation followed those described in Gleadow et al. (2015). After polishing, ^{252}Cf irradiation (Donelick and Miller 1991) was carried out to increase the number of measurable confined tracks. To ensure identical etching conditions, two super-mounts were prepared by separately affixing the two grain mounts for each apatite sample to two petrographic slides (Fig. 1A, second row). One super-mount was etched using 5.0M HNO_3 at 20°C for 20s (Gleadow et al. 1986), and the other with 5.5M HNO_3 at 21°C for 20s (Carlson et al. 1999) (Fig. 1A, third row); for simplicity, hereafter these are referred to as 5.0M and 5.5M, respectively. Etching followed procedures described by Donelick et al. (2005).

Grain images were acquired at 1000x magnification on a Zeiss M1m Axio-Imager microscope operating under TrackWorks control software at the University of Melbourne (Gleadow et al. 2015). Transmitted and reflected light image stacks, with depth step sizes of 0.3 μm for the 5M mount and 0.2 μm for the 5.5M mount, each covering a depth of ~6-7 μm of selected c-axis-horizontal grains, were captured automatically (Gleadow et al. 2009). Measurements on captured digital image sets at an ~8000x effective magnification were performed by two analysts, An1 and An2, each with ~10 years of experience in fission-track analysis. Each analyst had a copy of all of the image sets and processed them independently in their home laboratories using FastTracks V3 image processing software. Parameters determined manually on the image sets included c-axis directions for each grain, >100 3D confined track lengths (azimuth (2D) angle to c-axis, dip and true (3D) angle to c-axis were calculated automatically) and >40 Dpar values (from both induced and Cf-irradiation tracks) per sample. Location and measurement data were saved in XML-format files. The FastTracks software displays the data as overlays on the image stacks, and allows the visibility of recorded features to be turned off and on individually.

To assess the reproducibility of the measurements, two sets of data were acquired. In Set 1, lengths from the 5.0M mount were selected, imaged, and measured by An1 and lengths from the 5.5M mount were selected, imaged and measured by An2. The respective XML files were then given to the non-selecting analyst, who viewed the measurements with the numerical data layers turned off, leaving only markers of where the measurements were taken (Fig. 1B, second row). The second analyst then re-measured the selected features. Any length selected by one analyst, but assessed as inappropriate or non-measurable by the other, was excluded from the final results. To test for operator differences in track selection, in Set 2 the mounts were switched, with An1 locating and measuring tracks in different areas of the 5.5M mount, and An2 doing the same for the 5.0M mount. For Set 2, the non-choosing analyst did not repeat the measurements. Length distributions obtained from all measurements were used to evaluate and compare the length selection criteria (or identification pattern) of each analyst.

To help evaluate the effects on length measurements of changing etching conditions and etching rates near the track tips, the 5M and 5.5M mounts were then etched for an additional 5 seconds at their respective acid concentration and temperature. All tracks measured in Set 1 for selected samples were then re-imaged and re-measured. After this additional etching step, no "new" track was added to the population; only tracks measured after the first etching step were re-measured. Those tracks that could not be measured after further etching, due to their ends overlapping with neighboring features and those that could no longer be identified as confined fission-tracks, were excluded.

4. Results

In total, 8391 confined fission-track length measurements were performed, along with 1480 etch figure diameter measurements. For Set 1, 3100 tracks were found by the analysts that first examined each mount; of these, 422 were rejected by the other analyst, an average rejection rate of 14%. Statistics for Set 1 include 2678 confined track lengths (Tables 2 and 3) and 740 etch figures (Table 4), all measured by both analysts. In Set 2, 921 confined track lengths were selected and measured by An1 and 895 by An2, independently from each other's measurements (Table 5). Details of 788 confined length measurements from one sample each of the three major apatite species (Durango, Renfrew and Mud Tank) etched for an additional 5s and re-imaged are in Table 6.

4.1 Measurement comparison at the individual track level

Track length results for measurement Set 1 are provided in Table 2, and differences in measurements between analysts calculated on a track-by-track basis are reported in Table 3. Figure 2 shows comparisons of length (A,B), azimuth (C,D) and dip angle values (E,F) of the same individual tracks measured by the two analysts with the two etching protocols. Track lengths from 5.5M etching (Figure 2A) are longer and slightly more consistent between analysts than observed with the 5.0M protocol (Figure 2B). Notably, results for both analysts feature multiple track lengths below 14 μm for the 5.0M etchant in most samples, whereas with the 5.5M protocol there are only a few tracks measured under 14 μm. Results for both etching protocols feature a high level of agreement between the analysts for track azimuth (Fig. 2C, D), although there is a slightly higher level of dispersion in the 5.0M data. Dip angles are also slightly more dispersed in the 5.0M data (Fig. 2E, F), although part of the discrepancy stems from An1 finding more high-dip tracks than An2 in the initial selection process. Low dip angles are limited to particular values due to the depth intervals being restricted to specific planes in the image stacks, as detailed by Li et al. (2018).

Table 3 reveals no evidence of a systematic difference in length or dip measurements between analysts; mean differences are small and equally likely to be positive or negative. The standard deviations of the differences are virtually all lower in measurements from the 5.5M-etched mount, indicating a higher level of agreement. The sole exceptions are the 3D length and dip of sample 12, which are probably correlated.

Because 3D lengths are influenced by dip, we compare 3D length difference with dip difference in Figure 3. The dispersion for the 5.0M data in Figure 3A is less compact (R^2=0.021) but also less correlated than in Figure 3B (R^2=0.085). The slight correlation in the latter indicates that a non-negligible component of length dispersion is due to dip dispersion, while the weaker correlation in the former suggests that other sources of dispersion are more influential.

Results from each analyst measuring the same set of etch figures in each sample are provided in Table 4 and Figure 4. For most samples, An1 and An2 generally agree at low Dpar values, but as Dpar increases An1 tends to get longer Dpar values. The extent of deviation is greater for the 5.0M etch than the 5.5M, although Sample 8 shows considerable variability in both etching protocols.

4.2 Track selection / rejection / angular bias

Following the unexpected difference between confined track lengths for the two etching protocols shown in Figure 2A and 2B, we tried to ascertain whether an operator bias could be partially responsible. Due to our use of an image capture system, we were able to focus on the differences in measurements at the individual track level, a scale not possible in previous studies. As documented by the difference in numbers of tracks found and measured by the first analyst to analyze a mount (N_{init}) and the number accepted by the second analyst (N_{acc}), on average ~14% of the tracks selected by one analyst were not accepted by the other for measurement (Table 2). Disagreements were primarily due to perception of an indistinct or obscured track end, or intersection of a track end with the polished and etched surface, and were probably influenced by choice of observation mode.

Figure 5 shows examples of reflected and transmitted light images of confined tracks selected and measured by An1 and rejected by An2, while Figure 6 shows rejections by An1 of tracks found and measured by An2. Disagreements primarily concern the appearance of track ends that are near the borderline of resolvability. The results probably reflect the respective length observation and measuring strategies of the two analysts. An1 uses both reflected and transmitted light to locate and measure tracks, whereas An2 uses transmitted light only. Nonetheless, the selection criteria were sufficiently similar that the great majority (86%) of tracks were judged to be acceptable by both analysts.

The effects of track rejection on overall length determinations were varied. Except for two of the four Mud Tank apatite analyses, the rejected lengths were on average shorter than the accepted ones, but often not by very much. With the 5.5M etch, the mean length of rejected tracks was within two standard errors of the accepted tracks for all samples except the King Island apatite. Conversely, the 5.0M etch of samples 7, 8 and 12 featured mean rejected lengths that were more than two standard errors shorter than the accepted lengths, and samples 9 and 10 are at the edge of the 2-SE limit. However, even in the most egregious cases, removal of the rejected tracks changed the mean track length by at most 0.12 μm.

The relative frequency of confined track orientations relative to the c-axis found by each analyst is shown in Figure 7. There is no obvious difference at lower angles, but above ~85°, where track ends become difficult to distinguish (Fig. 6C,D) and under-etching is harder to evaluate, An2 selected them at almost twice the rate as An1. It is not clear, however, whether this is due to different etching (e.g., Jonckheere et al. 2019), different operator tendencies in using transmitted versus reflected light when scanning the grain mount, or other analyst biases pertaining to these challenging tracks.

Our Set 2 measurements were intended to test whether the discrepancy between 5.0M and 5.5M may have been due to an analyst selection bias. Figure 8A shows the length distributions for Sample 1 in measurement Set 1, in which An1 selected the 5.0M tracks and An2 the 5.5M tracks. In measurement Set 2, An1 selected and measured new tracks in the 5.5M mount and An2 in the 5.0M mount (Fig. 1, Table 5). The length distributions, shown in Figure 8B, show the same overall pattern as in Figure 8A: regardless of the analyst making the selection, measurements with the 5.0M protocol feature a higher standard deviation and lower mean track length than with 5.5M. The results clearly show that the differences observed relate to the etching protocol, and not the analyst.

4.3 Track lengthening after additional etching

In our next effort to understand the discrepancy in track lengths between the 5.0M and 5.5M protocols, we etched each mount for another 5s to study etching rates near the track tips, and test whether the short tracks observed in Figure 2A were under-etched. Only samples 3, 4, and 10, which showed the greatest number of sub-14-μm tracks in the 5.0M etch, were re-measured. The effect of the longer etch times on track appearance are shown in Figure 9. Tracks become longer and thicker with increasing etching time in both cases. At 25 seconds the track ends become relatively more distinct and less ambiguous in both apatites with both light sources, which in turn led to less dispersed measurements, as shown in Table 6.

Figure 10 compares the individual track lengths at 20s and 25s from both etching protocols in samples 3, 4 and 10. In all three cases the 5.0M re-etch shows a significantly greater increase in length compared to 5.5M, especially among shorter tracks. The same trend can be seen in 20s and 25s histogram plots of sample 3 (Figure 11). At the same time, Figure 10 also makes it evident that shorter tracks were likely to be lengthened slightly more even with the 5.5M etch. Summary statistics for these experiments are plotted in Figure 12. Standard deviation falls substantially for the 5.0M and slightly for 5.5M after re-etching, although the 5.0M data remain more dispersed than the 20s 5.5M data. Mean track length increases for all samples, but more for the 5.0M etch (~0.7 μm on average) for each sample than the corresponding 5.5M etch (~0.4 μm). The greatest increase in length and decrease in standard deviation occurs for the Renfrew apatite, Sample 3, after the 5.0M re-etching, suggesting this sample was the most under-etched after 20s. Notably this apatite has the lowest bulk etching rate of all the samples (see Dpar in Table 4) in this study due to its near end-member fluorapatite composition. The extra five seconds of etching resulted in all 5.0M mean track lengths exceeding the 20s 5.5M data

4.4 Ambient temperature annealing on year-to-decade time scales

Ambient temperature annealing over year to decade time scales is reflected in the mean induced track lengths in Figure 13. Data from both analysts and both 20s etching protocols show a mean track length decrease with increasing storage time at ambient temperature in all three apatites studied. Overall, the 5.5M data show clearer evidence of a significant change, and there is apparent variation in the behavior of the different apatites. Durango apatite, which features the largest time difference, shows significant annealing between 3.5 and 27 years, but not between 27 and 32, while Renfrew and Mud Tank apatites show some evidence of annealing between 27 and 32 or 36 years, respectively. These samples were stored in a temperature-controlled building where ambient temperature variations are mostly in the range 20±3 °C.

5. Discussion

We next consider our results in the context of our original four questions.

5.1 Agreement between analysts

DOI: https://doi.org/10.2138/am-2019-7046

One surprising result of our study was the extent to which one analyst rejected tracks selected by the other (~14%). These varying decisions are partly due to different emphases placed on transmitted versus reflected light observation, slight differences in personal criteria based on their previous experience in fission track measurement, and their respective “visual memory” based on tracks previously measured. There is also some degree of variable decision-making over time, as first documented by Barbarand et al. (2003); follow-up inspection of a rejected track frequently (roughly 50% of the time) resulted in the initial analyst recognizing and agreeing with the reason for rejection. Combined, these results indicate that utilizing image capture for fission-track analysis, and including some kind of verification step (i.e. re-evaluation by the same or a different analyst) can increase data quality.

Rejected tracks were on average slightly shorter than non-rejected ones. The magnitude of disagreement was larger for the 5.0M etch, and the 5.5M-etched sample with the largest disagreement was King Island, which has the smallest etch figures. Our interpretation is thus that 5.0M and King Island tracks were more likely to be rejected due to indistinct, poorly etched ends, which may reflect under-etching, while the other rejections included a larger share of semi-obscured track ends, which would have a less systematic effect on lengths.

The degree of difference between the two analysts is small compared to the total number of tracks accepted by both (~86%). When both analysts agreed that a confined track was valid, agreement between their measurements was excellent (Figure 2, Table 3). The slightly larger extent of disagreement in the 5.0M compared to the 5.5M confined length data (Fig. 2A, B) likely resulted from track ends being on average somewhat less clearly revealed. Azimuth angle data were extremely repeatable (Fig. 2C, D). Dip angles were less so (Fig. 2E, F), probably due to variation in determining the best image plane in which the track tips are in focus. Interestingly, however, the disagreements are roughly symmetrical, showing no strong evidence of a bias in dip measurement due to different techniques. The effect of dip disagreements on 3D length is noticeable in the 5.5M data but not in the 5.0M (Fig. 3), even though the 5.0M data included more steeply dipping tracks for which a given change in dip would have a larger effect on length. We interpret this to mean that unclear track tips exert a larger effect on measurements than dip uncertainties.

The light source used when scanning for confined tracks may affect the number of measured tracks at certain angles to the c axis (Fig. 6C, D; Fig. 7). Given that the track densities of these samples were quite high and they were not subjected to high temperature annealing, analysts identified the target number of confined tracks (~200) with ease. The actual angular bias may be higher or lower with a higher "required" number of tracks or less track availability. Possible causes for bias could include a tendency to bypass tracks with ends that are possible but challenging to distinguish, or conversely a desire to attempt to include tracks from all angles.

Our Dpar data show evidence of both slight analyst bias and dispersion (Fig. 4, Table 4). Measurements by An1 tended to be slightly longer than An2, averaging about 0.1 μm, with both etches. The 5.5M etch decreased the difference between analysts for three apatite varieties. Interestingly, however, the apatite with the largest etch figures, Mud Tank, also featured the largest dispersion, which was not alleviated with the stronger etch. The reason behind this is not clear, although one possibility is zoning in OH content, which affects etch figure size and would be difficult to detect with electron microprobe analysis.

We also note that two apatites, Renfrew and King Island, have etch figures that average 1.5 μm or below with the 5.5M protocol, placing them outside the range observed by Carlson et al. (1999) and underlying the annealing models based on those data (Ketcham et al. 1999). In the Carlson et al. (1999) data set, Renfrew apatite had the second-smallest Dpar value reported (1.65±0.03 μm). The lower values reported here, averaging 1.49±0.02 μm, may partially reflect the different measurement systems used, though our mean values for Durango apatite, 1.78±0.02 μm versus 1.83±0.02 μm in Carlson et al. (1999) are much more comparable, although differing in the same sense. Taking these reference points, the yet lower Dpar value we observe for King Island, 1.40±0.02 μm, is well-supported evidence of natural Dpar values falling outside the calibrated range. The dispersion, bias and range of Dpar values reported here provides further evidence of the need for caution in using them for thermal history analysis (Ketcham et al. 2018).

5.2 Difference between two major etching protocols.

Our data show multiple indications that the 5.0M, 20°C, 20s etching protocol resulted in some degree of under-etching. Both analysts show broader track-length distributions with the 5.0M etch (Figure 2A, B and Figure 3), and slightly worse agreement on individual track lengths, dip angles (Figure 2E, F) and Dpar measurements (Figure 4). Similarly, the additional 5s etching step provided a greater increase in lengths, especially on initially short tracks (Figure 10), narrower histograms (Figure 11), and greater reduction in standard deviation (Figure 12) at 5.0M than with 5.5M. Altogether, these results imply that after 20s some tracks still had a region of somewhat enhanced track etching rate beyond their tips. In fact, the thickness of the point clouds for the 5.5M etch in Figure 10, reflecting length increases ranging from ~0.1-1 μm, implies that some 5.5M tracks were also under-etched, though with less frequency and severity.

It is possible that there was some problem with etching conditions for one or both of our mounts, such as the etchant concentration or temperature, although we followed our usual procedures with particular care and consider this extremely unlikely. We note that our mean track lengths are well within the range reported by Ketcham et al. (2015), even by very experienced and active analysts.

It should also be noted that the temperature used for our 5.0M protocol, 20°C, while employed for a number of classic studies (e.g., Gleadow et al. 1986) and data sets (Green et al. 1986; Barbarand et al. 2003b), seems to have drifted out of current practice. In the inter-laboratory study reported by Ketcham et al. (2015), nine different lab groups provided data in which they used 5.0M HNO_3, but at temperatures from 21°C to 24°C; none reported using 20°C. While temperature was carefully controlled in the present study, it is probable that some or all of these reported values reflect ambient etchant temperatures in the laboratory at the time, rather than controlled conditions.

Increasing etching temperature and/or concentration would be expected to increase etching rate and thus reduce the occurrence of under-etching. However, the data reported by Ketcham et al. (2015) do not show clear evidence of a systematic difference in mean track lengths with etching temperature at a given etchant strength, or between 5.0M and 5.5M results, across laboratories. This implies that factors aside from etchant protocol, such as analyst training, are also important for length measurements and their reproducibility. However, by effectively controlling for analyst-specific variation, this study demonstrates that subtle changes in etching conditions can have large effects on mean track-length measurements.

5.3 Can under-etching be evaluated reliably?

If some tracks etched with the 5.0M protocol were under-etched by more than 1 μm, were they improperly accepted for measurement? Although in some cases (illustrated in Fig. 5, 6, and 9) the ability to locate track ends reliably was debatable, leading to disagreement between the analysts, re-inspection of the image data for the overwhelming majority of sub-14-μm tracks accepted by both analysts did not change either analyst's mind about their appropriateness for measurement. According to each analyst's training and experience, they were valid confined tracks. The Set 2 data, in which the opposite analyst identified the tracks for each etch (Fig. 8), corroborates this conclusion.

We propose that the difficulty in evaluating under-etching is caused by the gradual, rather than sudden, diminishing of etching velocity along the track as its ends are approached (Fleischer et al., 1969; Jonckheere et al., 2017). If we call this velocity $v_T(x)$, with x denoting the distance along the track from its center to one end, then as the end is approached $v_T(x)$ becomes closer to, but still higher than, the bulk etching velocity (v_B) (Figure 14). As this occurs, the disparity between bulk and along-track etching may diminish sufficiently that the track tip can start to widen, making it easier to see and thus measure. This scheme maintains the original definition of "under-etched" as leaving some region where $v_T(x)>v_B$ unetched, although it may be unclear how different their values need to be to define the end of a given track.

Such an under-etching effect may be one of the mechanisms underlying the poor reproducibility of confined length data, either because the under-etching is undetected, or because analysts are more likely to disagree about track suitability. Tracks etched even for only 10 and 15s (Tamer 2012) can demonstrate reasonable/noticeable track ends, despite being under-etched. Taking the observation of track ends into account, it is not clear how to define a quantitative criterion to decide if a track is under-etched. Track thickness, the proximity to the crystal surface, and the number of intersecting tracks all play a role in the complex track-etching problem, which needs to be addressed in future experiments.

The issue of under-etching is not a trivial one. Track lengths are used to infer thermal histories (Ketcham et al. 2007; Gallagher 2012) and a shorter length implies more time spent at higher temperature. If a sufficient number of under-etched tracks are included in the measurement, or for that matter in the original annealing databases underlying the annealing models, the thermal history may be distorted. Furthermore, because there is no statistical test for over-dispersion of lengths as there is for ages, such a problem is likely to go undetected.

It may thus be worth re-evaluating whether the method by which an optimal etching protocol has been defined using step etching (Laslett et al. 1984, etc.) is overly conservative in trying to prevent over-etching, rather than minimize under-etching. By halting the etch immediately upon reaching the plateau mean track-length value, we may be effectively "setting up camp on the edge of a cliff," where a small change in etching conditions, or apatite composition or solubility, can have a large effect on the degree of etching and track appearance. In contrast, over-etching at the low v_B rate will only have a small effect on confined track lengths. However, further data are needed on whether and how this effect extends to spontaneous tracks (e.g., Jonckheere et al., 2017), or tracks with a greater degree of annealing.

It should also be remembered that new confined tracks are intersected continuously as surface-intersecting tracks widen with increasing etching at depth, and in fact the confined track revelation rate increases with etching time (Jonckheere et al., 2007). As a result, the presence of some under-etched tracks is unavoidable. There were under-etched tracks with both etching protocols in this study, as shown by the re-evaluation exercise (Fig. 5, 6). What will evolve with progressive etching is the relative number of under-etched tracks versus fully or over-etched ones.

5.4 Low-temperature annealing at year to decade time scales

The difference in mean track length versus time since irradiation provides intriguing evidence of track annealing at ambient temperatures over year to decade time scales. The fact that the differences are larger in the 5.5M data is probably traceable to the reduction of under-etching seen in the 5.0M data, increasing the clarity in the underlying signal. Our experimental design rules out these differences being due to variation in etching, measurement system, or analyst. However, we cannot rule out that there may be subtle differences in the different samples used of each apatite variety, irradiated at these separate times, that are responsible for the observed length variation. The fact that the trend is similar in all three, however, lends confidence to the interpretation that they represent a true annealing effect. Distinguishing whether the differences in apparent annealing rates are due to variation within the apatites, or variable annealing behavior between them, will require more data.

In Durango apatite, the mean length decrease between 2.6 and 27 years, on the order of 0.2-0.3 μm, substantially exceeds the predicted amount; the Ketcham et al. (1999) curvilinear annealing model predicts length reduction of ~0.03 μm, an order of magnitude less than we observe here. This may be due to the annealing model being based entirely on high-temperature annealing experiments (150°C and above), reducing its ability to capture annealing in this region of time-temperature space. It may be possible to combine these data with the Carlson et al. (1999) data set to improve the annealing model, but this would require a reconsideration of how much time can pass after irradiation to define an "initial" track length.

Other studies have also suggested that some possibly non-thermal reconstruction of the crystal lattice in apatite may take place at low temperature over short durations from minutes to days (Donelick et al. 1990; Belton, 2006). Further insight into early-stage fission track annealing may be obtained by searching for possible effects at high temperatures at very short time scales (e.g. Murakami et al. 2006), or at lower ambient temperature conditions (>0°C) at long time scales (Gleadow and Duddy 1981; Spiegel et al 2007).

6. Implications

Our results shed considerable light on the factors underlying the suboptimal reproducibility of fission-track length measurements documented in previous studies. When two operators observe the same well-etched confined fission track, their results are highly consistent (Table 3). As tracks ends become less distinct due to a lower degree of etching, measurement consistency diminishes somewhat, and eventually disagreements arise as to whether a track is fully etched or not. Many of these disagreements can be resolved, and measurement consistency improved, by utilizing an image capture system and including a re-evaluation step in analysis protocols.

An unexpected finding is that, even prior to this point of disputed appropriateness for measurement, confined tracks can be substantially under-etched by over 2 µm compared to near-full-extent of the region of enhanced etchability defining the latent track. By carefully controlling for analytical effects due to instrumentation and analyst bias, we have demonstrated significant and potentially influential differences between two of the principal etching protocols that have been employed for apatite fission-track analysis. The protocol using 5.0M HNO_3 at 20°C for 20s appears to be more susceptible to a slight and sometimes significant degree of under-etching. Moreover, this under-etching is not easily detectable under standard optical microscope observation.

Under-etching, combined with analyst-varying criteria for what constitutes a well-etched track end, may account for a significant component of the poor reproducibility in track length measurements. It is clear that the definition of under-etching needs to evolve, accounting for the variability in etch rates as track tips are approached. The community has not fully explored how apatite fission-track etching starts, progresses and stops in a given etching protocol, and how it is affected by varying apatite solubility. Track etching velocities and bulk etching velocities must be established to figure out when tracks are fully etched in a step etching protocol to understand the etching characteristics in a more complete way.

If sufficiently pervasive, under-etching will affect thermal history inverse modeling, biasing results toward higher temperatures to replicate the apparent increased degree of annealing. It is likely that under-etching has a direct and deleterious effect on the reproducibility of such modeling (Ketcham et al., 2018).

We have also documented evidence for ambient temperature annealing of induced tracks at year to decade time scales in multiple apatites over times up to 36 years, an observation not predicted by annealing models based exclusively on high-temperature experiments. This finding supports previous studies (Donelick et al. 1990; Belton 2006) in implying that our conception of the initial confined track length should evolve to account for time since irradiation (e.g., Laslett and Galbraith, 1996). Combining ambient temperature annealing data with higher-temperature annealing data sets (Carlson et al. 1999; Barbarand et al. 2003b) may also improve apatite fission track thermal history modeling.

7. Acknowledgments

This research was supported by the Geology Foundation of the Jackson School of Geosciences and the School of Earth Sciences at University of Melbourne. We thank R. Jonckheere and E. Sobel for thorough and helpful reviews, which improved the quality of this work.

8. References

Barbarand, J., Carter, A., Wood, I., and Hurford, T. (2003a) Compositional and structural control of fission-track annealing in apatite. Chemical Geology,198(1), 107-137.

Barbarand, J., Hurford, T., and Carter, A. (2003b) Variation in apatite fission-track length measurement: Implications for thermal history modelling. Chemical Geology,198(1), 77-106.

Belton, D.X. (2006) The low-temperature thermochronology of cratonic regions. University of Melbourne. PhD thesis, 306 pages.

Bhandari, N., Bhat, S.G., Lal, D., Rajagopalan, G., Tamhane, A.S., and Venkatavaradan, V.S. (1971) Fission fragment tracks in apatite: recordable track lengths. Earth and Planetary Science Letters, 13, 191–199.

Carlson, W.D., Donelick, R.A., and Ketcham, R.A. (1999) Variability of apatite fission-track annealing kinetics: I. Experimental results. American Mineralogist, 84, 1213-1223.

Donelick, R.A., Roden, M.K., Mooers, J.D., Carpenter, B.S., and Miller, D.S. (1990) Etchable length reduction of induced fission tracks in apatite at room temperature (≈ 23°C): Crystallographic orientation effects and "initial" mean lengths. International Journal of Radiation Applications and Instrumentation. Part, 17, 261–265.

Donelick, R.A., and Miller, D.S. (1991) Enhanced tint fission track densities in low spontaneous track density apatites using 252Cf-derived fission fragment tracks: A model and experimental observations. International Journal of Radiation Applications and Instrumentation, 18, 301-307.

Donelick, R.A., Ketcham, R.A., Carlson, W.D. (1999) Variability of apatite fission-track annealing kinetics: II. Crystallographic orientation effects. American Mineralogist, 84, 1224-1234.

Donelick, R.A., O'Sullivan, P.B, Ketcham, R.A. (2005) Apatite Fission-Track Analysis. Reviews in Mineralogy & Geochemistry, 58, 49-94.

Fleischer, R.L., Price, P.B., Walker, R.M., and Hubbard, E.L. (1964) Track registration in various solid-state nuclear track detectors. Physical Review, 133, 1443-1449.

Fleischer, R.L., Price, P.B., Woods, R.T. (1969) Nuclear-particle-track identification in inorganic solids, Physical Review, 188(2), 563-568.

Gallagher, K. (2012) Transdimensional inverse thermal history modeling for quantitative thermochronology. Journal of Geophysical Research: Solid Earth, 177, ppB02408.

Gleadow, A.J.W., and Lovering, J.F (1978) Fission track geochronology of King Island, Bass Strait, Australia: Relationship to Continental Rifting. Earth and Planetary Science Letters, 37, 429-437.

Gleadow, A.J.W., and Duddy, I.R. (1981) A natural long-term track annealing experiment for apatite. Nuclear Tracks, 5, 169-174.

Gleadow, A.J.W., Duddy, I.R., Green, P.F., and Lovering, J.F. (1986) Confined fission track lengths in apatite: a diagnostic tool for thermal history analysis. Contributions to Mineralogy and Petrology, 94, 405-415.

Gleadow, A.J.W., Gleadow, S.J., Belton, D.X., Kohn, B.P., Krochmal, M.S., and Brown, R.W. (2009) Coincidence mapping - a key strategy for the automatic counting of fission tracks in natural minerals. Geological Society, London, Special Publications, 324, 25-36.

Gleadow, A., Harrison, M., Kohn, B., Lugo-Zazueta, R., and Phillips, D. (2015) The Fish Canyon Tuff: A new look at an old low-temperature thermochronology standard. Earth and Planetary Science Letters, 424, 95-108.

Green, P.F., and Durrani, S.A. (1977) Annealing studies of tracks in crystals. Nuclear Track

Detection, 1, 33-39.

Green, P.F., Duddy, I.R., Gleadow, A.J.W., Tingate, P.R., and Laslett, G.M. (1986) Thermal annealing of fission tracks in apatite. 1. A qualitative description. Chemical Geology: Isotope Geoscience Section, 59, 237-253.

Green, P.F., Duddy, I.R., Laslett, G.M., Hegarty, K.A., Gleadow, A.J.W., and Lovering, J.F. (1989) Thermal annealing of fission tracks in apatite 4. Quantitative modelling techniques and extension to geological timescales. Chemical Geology: Isotope Geoscience Section, 79, 155-182.

Jonckheere, R. (2003) On methodical problems in estimating geological temperature and time from measurements of fission tracks in apatite. In Radiation Measurements, 36, 43-55.

Jonckheere, R., Enkelmann, E., Min, M., Trautmann, C., Ratschbacher, L. (2007) Confined fission tracks in ion-irradiated and step-etched prismatic sections of Durango apatite. Chemical Geology, 242, 202-217.

Jonckheere, R., Tamer, M.T., Wauschkuhn, B., Wauschkuhn, F., and Ratschbacher, L. (2017) Single-track length measurements of step-etched fission tracks in Durango apatite: "vorsprung durch Technik." American Mineralogist, 102, 987-996.

Jonckheere, R., Wauschkuhn, B., and Ratschbacher, L. (2019) On growth and form of etched fission tracks in apatite: A kinetic approach. American Mineralogist, 104, 569-579.

Ketcham, R.A., Donelick, R.A., and Carlson, W.D. (1999) Variability of apatite fission-track annealing kinetics: III. Extrapolation to geological time scales. American Mineralogist, 84, 1235-1255.

Ketcham, R.A. (2005) Forward and Inverse Modeling of Low-Temperature Thermochronometry Data. Reviews in Mineralogy and Geochemistry, 58, 275-314.

Ketcham, R.A., Carter, A., Donelick, R.A., Barbarand, J., and Hurford, A.J. (2007) Improved measurement of fission-track annealing in apatite using c-axis projection. American Mineralogist, 92, 799-810.

Ketcham, R.A., Donelick, R.A., Balestrieri, M.L., and Zattin, M. (2009) Reproducibility of apatite fission-track length data and thermal history reconstruction. Earth and Planetary Science Letters, 284, 504-515.

Ketcham, R.A. (2015) Technical Note: Calculation of stoichiometry from EMP data for apatite and other phases with mixing on monovalent sites. American Mineralogist, 100,1620-1623.

Ketcham, R.A., Carter, A., and Hurford, A.J. (2015) Inter-laboratory comparison of fission track confined length and etch figure measurements in apatite. American Mineralogist, 100, 1452-1468.

Ketcham, R.A., van der Beek, P., Barbarand, J., Bernet, M., and Gautheron, C. (2018) Reproducibility of Thermal History Reconstruction From Apatite Fission-Track and (U-Th)/He Data. Geochemistry, Geophysics, Geosystems, 19, 2411-2436.

Ketcham, R.A. (2019) Fission track annealing: from geologic observations to thermal modeling. In P.G. Fitzgerald and M. Malusa, Eds., Fission track thermochronology and its application to geology (P. Fitzgerald and M. Malusa, eds.) pp. 49–75. Springer

Textbooks in Earth Sciences, Geography and Environment, Springer.

Laslett, G.M., Gleadow, A.J.W., and Duddy, I.R. (1984) The relationship between fission track length and track density in apatite. Nuclear Tracks and Radiation Measurements (1982), 9, 29-38.

Laslett, G.M., and Galbraith, R.F. (1996) Statistical properties of semi-tracks in fission track analysis. Radiation Measurements, 26, 565-576.

Meitner, L., and Frisch, O.R. (1939) Disintegration of uranium by neutrons: A new type of nuclear reaction. Nature, 143, 239-240.

Miller, D.S., Eby, N., McCorkell, R., Rosenberg, P.E., and Suzuki, M. (1990) Results of interlaboratory comparison of fission track ages for the 1988 fission track workshop. International Journal of Radiation Applications and Instrumentation. Part, 17, 237–245.

Murakami, M., Yamada, R., and Tagami, T. (2006) Short-term annealing characteristics of spontaneous fission tracks in zircon: A qualitative description. Chemical Geology, 277, 214-222.

Naeser, C.W., Zimmerman, R.A., Cebula, G.T. (1981) Fission-track dating of apatite and zircon: an interlaboratory comparison. Nuclear Tracks, 5, 65-72.

Paul, T.A., and Fitzgerald, P.G. (1992) Transmission electron microscopic investigation of fission tracks in fluorapatite. American Mineralogist, 77, 336-344.

Qingyang L., Gleadow, A.J.W., Seiler, C., Kohn, B., Vermeesch, P., Carter, A., Hurford, A. (2018) Observations on three-dimensional measurement of confined fission track lengths in apatite using digital imagery. American Mineralogist, 103, 430-440.

Sobel, E. and Seward, D. (2010)Influence of etching conditions on apatite fission-track etch pit diameter. Chemical Geology, 271, 59-69.

Spiegel, C., Kohn, B., Raza, A., Rainer, T., and Gleadow, A. (2007) The effect of long-term low-temperature exposure on apatite fission track stability: A natural annealing experiment in the deep ocean. Geochimica et Cosmochimica Acta, 71, 4512-4537.

Tamer, M.T. (2012) The Lengths of Fossil and Induced Fission Tracks in Durango Apatite. Technische Universitaet Bergakademie Freiberg. Masters Thesis, 37p.

Tamer, M.T., and Ketcham, R.A. (2018) Is low-temperature fission-track annealing in apatite a thermally controlled process? 16th International Conference On Thermochronology, Germany/Quedlinburg. Conference proceedings.

Wagner, G.A. (1981) Fission-track ages and their geological interpretation. Nuclear Tracks, 5, 15-25.

Wagner, G.A., Gleadow, A.J.W., and Fitzgerald, P.G. (1989) The significance of the partial annealing zone in apatite fission-track analysis: Projected track length measurements and uplift chronology of the transantarctic mountains. Chemical Geology: Isotope Geoscience Section, 79, 295-305.

Watt, S., and Durrani, S.A. (1985) Thermal stability of fission tracks in apatite and sphene: Using confined-track-length measurements. Nuclear Tracks and Radiation Measurements (1982), 10, 349–357.

Watt, S., Green, P.F., and Durrani, S.A. (1984) Studies of annealing anisotropy of fission tracks in mineral apatite using Track-in-Track (tint) length measurements. Nuclear Tracks and Radiation Measurements (1982), 8, 371–375.

Ziegler, J.F., Biersach, P.J. and Ziegler, M.D. (2008), SRIM The Stopping and Range of Ions in Matter, v05ed., 390pp. SRIM Co. Chester, Maryland.

Table 1: Apatites used in this study, neutron irradiation dates and ages of induced tracks.

Apatite	Composition (apfu)[1]			Irradiation Location	Sample Number	Irradiation Date	Track Age (years)[4]	Track Age (ln(s))
	F	Cl	OH					
Durango	1.81[2]	0.13[2]	0.07[2]	Melbourne	1	2/20/1985	32.11	20.74
				Melbourne	7	3/8/1990	27.07	20.56
				Austin	10	7/30/2014	2.65	18.24
Mud Tank	1.30[3]	0.03[3]	0.67[3]	Melbourne	4	3/25/1981	36.02	20.85
				Melbourne	8	3/8/1990	27.07	20.56
Renfrew	1.95[2]	0.01[2]	0.04[2]	Melbourne	3	2/20/1985	32.11	20.74
				Melbourne	9	3/8/1990	27.07	20.56
King Island	1.99[3]	0.01[3]	0.00[3]	Melbourne	12	11/26/1973	43.36	21.04

[1] Stoichiometry calculations from Ketcham (2015).
[2] Composition data from Carlson et al. (1999).
[3] Composition data from University of Melbourne, using Cameca SX50 electron microprobe (settings: accelerating voltage = 15 kV; beam current = 35 nA; beam spot size = 10 μm).
[4] All samples initially etched on 3/25/2017.

Table 2: Set 1 confined track length measurements.

Apatite	Sample	Initial Measurements[1]			Rejected Measurements[2]				Accepted Measurements[3]				
									N	Analyst 1		Analyst 2	
		N	MTL (μm)	StDev (μm)	N	Rejection Rate (%)	MTL (μm)	StDev (μm)		MTL (μm)	StDev (μm)	MTL (μm)	StDev (μm)
		5.0M etch[4]											
Durango	1	212	15.52(07)	1.04	28	13.0	15.37(28)	1.48	188	15.54(07)	0.97	15.54(07)	0.91
	7	170	15.40(06)	0.83	19	11.2	14.57(20)	0.87	151	15.50(06)	0.76	15.53(07)	0.81
	10	223	15.55(06)	0.96	37	16.6	15.29(19)	1.16	186	15.60(07)	0.91	15.66(07)	0.92
Mud Tank	4	214	16.10(07)	0.97	26	12.1	16.34(26)	1.32	188	16.07(07)	0.91	16.03(07)	0.90
	8	210	16.01(07)	1.06	22	10.5	14.97(28)	1.33	188	16.13(07)	0.96	16.08(07)	0.89
Renfrew	3	202	15.33(09)	1.32	28	13.9	15.02(33)	1.76	174	15.38(09)	1.24	15.32(07)	1.17
	9	205	15.45(07)	1.03	36	17.6	15.16(19)	1.13	169	15.52(08)	1.00	15.53(07)	0.94
King Island	12	174	15.07(09)	1.19	13	7.5	14.18(45)	1.64	161	15.14(06)	1.14	15.19(07)	1.07
		5.5M etch[5]											
Durango	1	188	15.83(05)	0.73	29	15.4	15.78(15)	0.79	159	15.90(06)	0.71	15.84(06)	0.72
	7	188	15.89(06)	0.79	29	15.4	15.75(17)	0.91	159	15.87(06)	0.80	15.92(06)	0.77
	10	200	16.05(05)	0.70	31	15.5	15.96(14)	0.80	169	16.15(05)	0.71	16.06(05)	0.68
Mud Tank	4	190	16.25(06)	0.78	31	16.3	16.43(15)	0.85	159	16.18(06)	0.82	16.21(06)	0.76
	8	201	16.33(05)	0.70	29	14.4	16.27(15)	0.79	172	16.34(05)	0.71	16.34(05)	0.68
Renfrew	3	196	15.62(05)	0.73	26	13.3	15.49(15)	0.75	170	15.63(06)	0.74	15.64(06)	0.73
	9	202	15.77(05)	0.69	38	18.8	15.75(10)	0.61	164	15.75(06)	0.72	15.78(06)	0.71
King Island	12	134	15.52(07)	0.77	13	9.7	15.13(18)	0.65	121	15.56(07)	0.81	15.57(07)	0.77

[1] Initial measurements of 5.0M etch by An1; initial measurements of 5.5M etch by An2.
[2] Subset of initial tracks rejected by the other analyst.
[3] Subset of initial tracks accepted by both analysts.
[4] Tracks etched using 5.0M HNO_3, located, imaged, and measured initially by An1; measurements repeated on the same images by An2.
[5] Tracks etched using 5.5M HNO_3, located, imaged, and measured initially by An2; measurements repeated on the same images by An1.

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2019-7046

Table 3: Difference between An1 and An2 Set 1 confined track measurements.

Apatite	Sample	N	2D MTL (μm)[1]	StDev	3D MTL (μm)[2]	StDev	Dip (°)	StDev	3D c-ax (°)[3]	StDev
	5.0M etch									
Durango	1	188	-0.01	0.43	-0.02	0.45	0.01	2.89	-0.15	1.08
	7	151	0.00	0.43	-0.02	0.44	-0.09	4.37	-0.03	0.97
	10	186	-0.08	0.50	-0.06	0.51	0.73	2.64	-0.03	1.10
Mud Tank	4	188	0.06	0.41	0.04	0.43	-0.14	2.81	0.04	1.17
	8	188	0.02	0.38	0.05	0.38	0.93	3.03	0.07	1.12
Renfrew	3	174	0.03	0.37	0.05	0.37	0.68	3.27	0.02	1.09
	9	169	-0.01	0.35	-0.02	0.34	0.04	3.31	0.07	0.96
King Island	12	161	-0.07	0.38	-0.05	0.34	0.28	3.41	-0.09	1.27
	5.5M etch									
Durango	1	159	-0.06	0.21	-0.05	0.25	0.22	2.70	0.06	0.88
	7	159	0.05	0.31	0.05	0.32	-0.08	2.14	-0.10	0.67
	10	169	-0.09	0.27	-0.09	0.28	0.21	2.38	-0.02	0.57
Mud Tank	4	159	0.07	0.26	0.04	0.30	-0.61	2.70	-0.15	1.00
	8	172	-0.01	0.22	0.00	0.23	0.10	2.54	-0.10	0.69
Renfrew	3	170	0.04	0.23	0.01	0.25	-0.82	2.86	-0.19	0.98
	9	164	0.02	0.27	0.02	0.28	-0.01	2.46	0.01	0.64
King Island	12	151	0.00	0.36	-0.01	0.38	0.17	3.65	-0.02	0.97

[1] Mean difference in mean confined track length without accounting for dip.
[2] Mean difference in 3D mean confined track length.
[3] Mean difference in 3D angle to c-axis

Table 4: Etch figure measurements.

Apatite	Sample	N	Analyst 1		Analyst 2	
			Dpar (μm)	StDev (μm)	Dpar (μm)	StDev (μm)
		5.0M etch[1]				
Durango	1	38	1.82(03)	0.16	1.77(02)	0.11
	7	44	1.75(02)	0.14	1.79(02)	0.11
	10	45	1.75(02)	0.12	1.62(02)	0.12
Mud Tank	4	44	2.18(04)	0.28	1.99(04)	0.26
	8	47	2.33(03)	0.24	2.18(04)	0.30
Renfrew	3	48	1.89(03)	0.19	1.57(02)	0.16
	9	47	1.67(02)	0.16	1.48(02)	0.14
King Island	12	45	1.39(02)	0.12	1.30(02)	0.13
		5.5M etch[2]				
Durango	1	50	1.78(02)	0.13	1.75(02)	0.11
	7	50	1.80(02)	0.16	1.78(02)	0.11
	10	40	1.77(02)	0.14	1.77(02)	0.15
Mud Tank	4	50	2.20(04)	0.30	2.05(03)	0.08
	8	50	2.42(03)	0.22	2.27(04)	0.25
Renfrew	3	42	1.55(03)	0.18	1.45(02)	0.15
	9	50	1.48(02)	0.14	1.48(02)	0.17
King Island	12	50	1.42(02)	0.17	1.37(02)	0.17

[1] Samples etched using 5.0M HNO_3, and etch figures located, imaged, and measured initially by An1; measurements repeated on the same images by An2.
[2] Samples etched using 5.5M HNO_3, and etch figures located, imaged, and measured initially by An2; measurements repeated on the same images by An1.

Table 5: Set 2 confined track length measurements, switching the analyst locating tracks from Set 1.

Apatite Species	Sample	Analyst 1 (5.5M)			Analyst 2 (5.0M)		
		N	MTL (μm)	StDev (μm)	N	MTL (μm)	StDev (μm)
Durango	1	131	15.84(07)	0.76	97	15.47(11)	1.12
	7	108	15.90(07)	0.77	99	15.14(09)	0.92
	10	134	15.95(06)	0.68	223	15.48(07)	1.06
Mud Tank	4	110	16.32(08)	0.87	94	16.03(08)	0.74
	8	109	16.30(08)	0.87	97	16.00(08)	0.77
Renfrew	3	111	15.66(08)	0.80	94	15.00(11)	1.04
	9	111	15.79(08)	0.84	94	15.17(10)	0.97
King Island	12	107	15.60(07)	0.76	97	15.13(12)	1.15

Table 6: Confined track length measurements after 5s additional etching.

Apatite	Sample	20s			20+5s		
		N	MTL (μm)	StDev (μm)	N	MTL (μm)	StDev (μm)
		5.0M etch					
Renfrew	3	174	15.38(09)	1.24	143	16.20(08)	0.99
Mud Tank	4	188	16.07(07)	0.91	146	16.72(07)	0.84
Durango	10	186	15.60(07)	0.91	138	16.17(07)	0.84
		5.5M etch					
Renfrew	3	170	15.64(06)	0.73	128	16.02(06)	0.71
Mud Tank	4	159	16.21(06)	0.76	117	16.61(07)	0.72
Durango	10	169	16.06(05)	0.68	116	16.50(06)	0.67

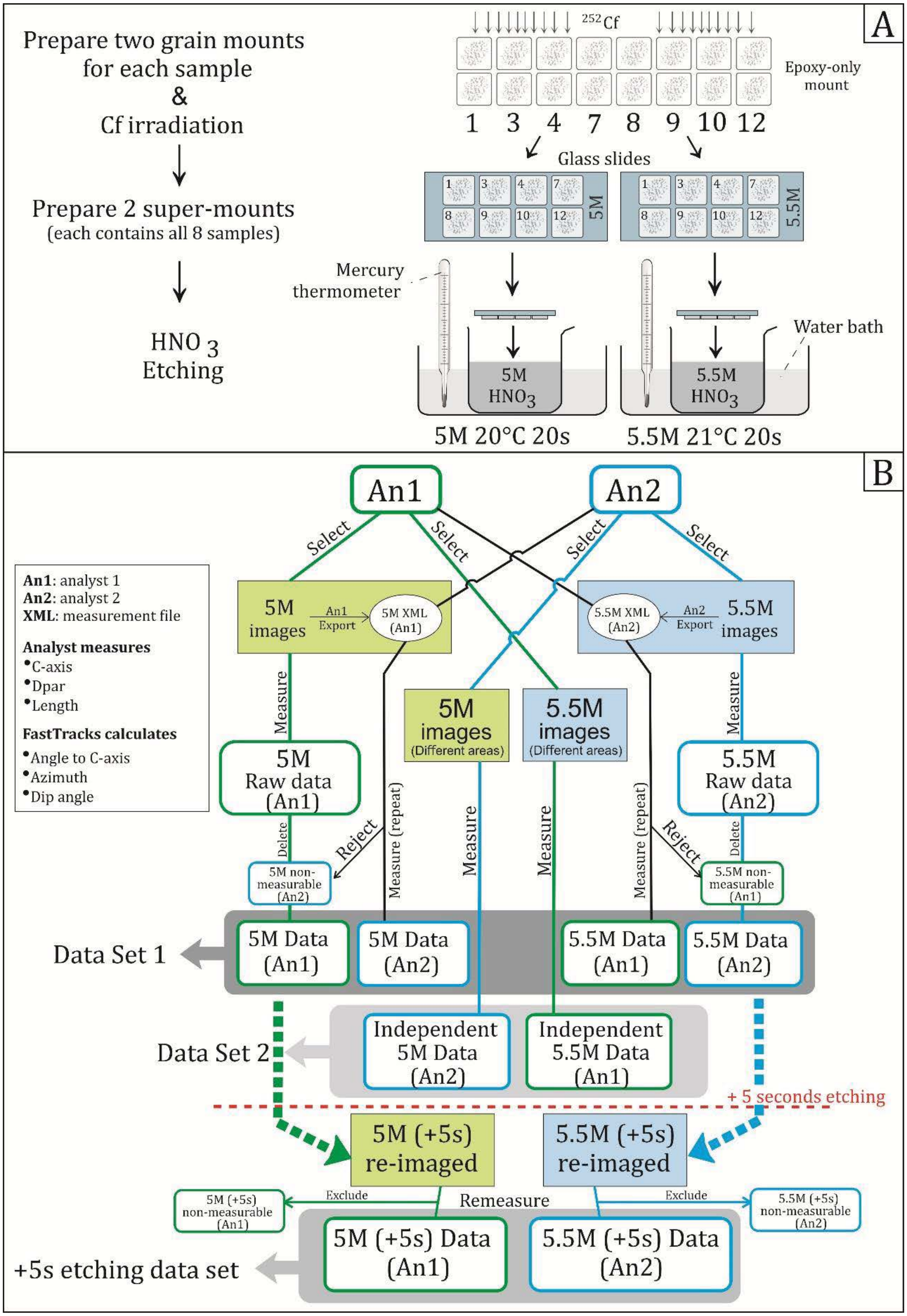

Figure 1: Experimental design of the study in etching (A) and measurement, track selection and re-etching procedures (B).

DOI: https://doi.org/10.2138/am-2019-7046

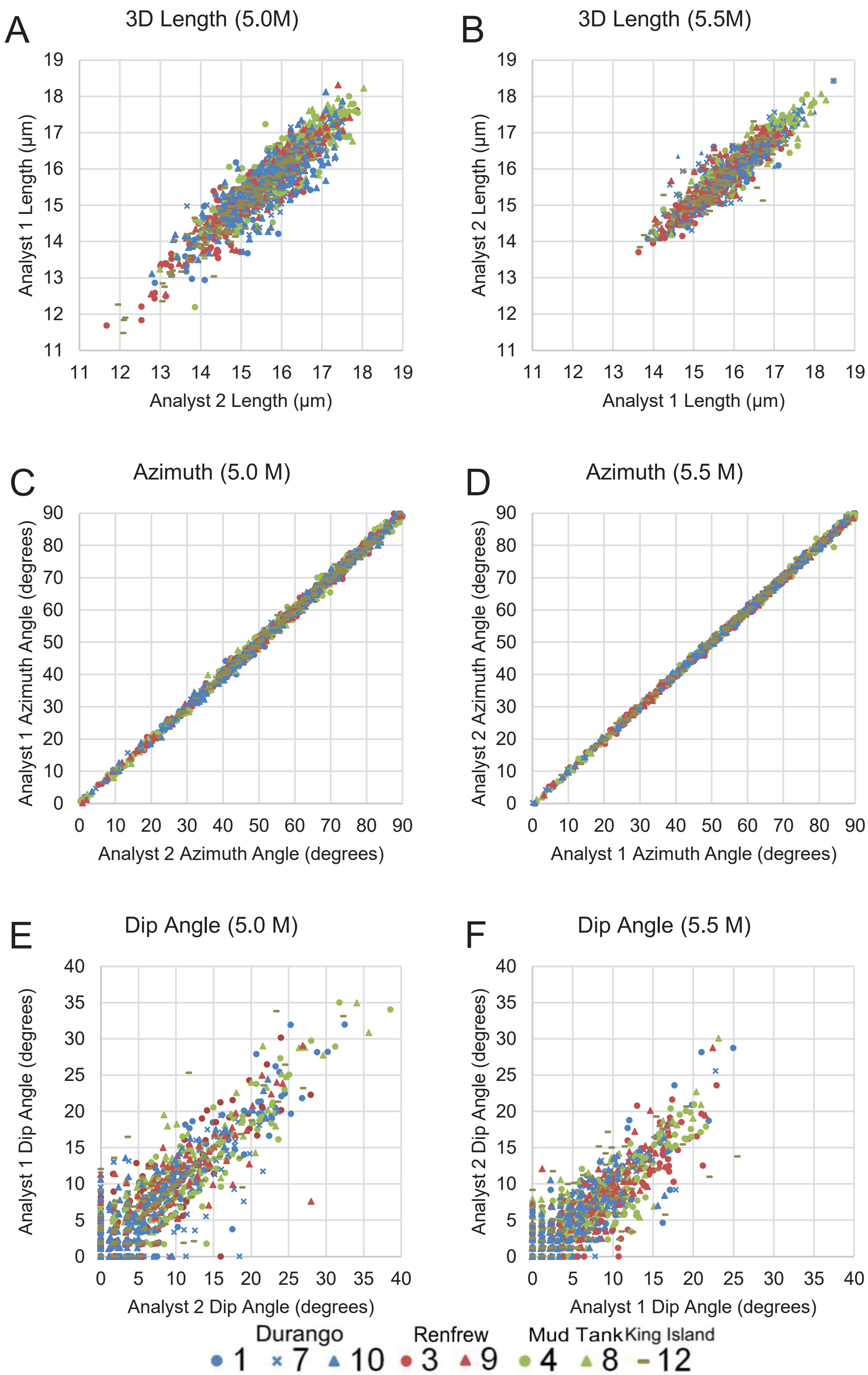

Figure 2: Comparison of the single individual confined track length measurements for two analysts with 5.0M (A, C, E) and 5.5M (B, D, F) protocols.

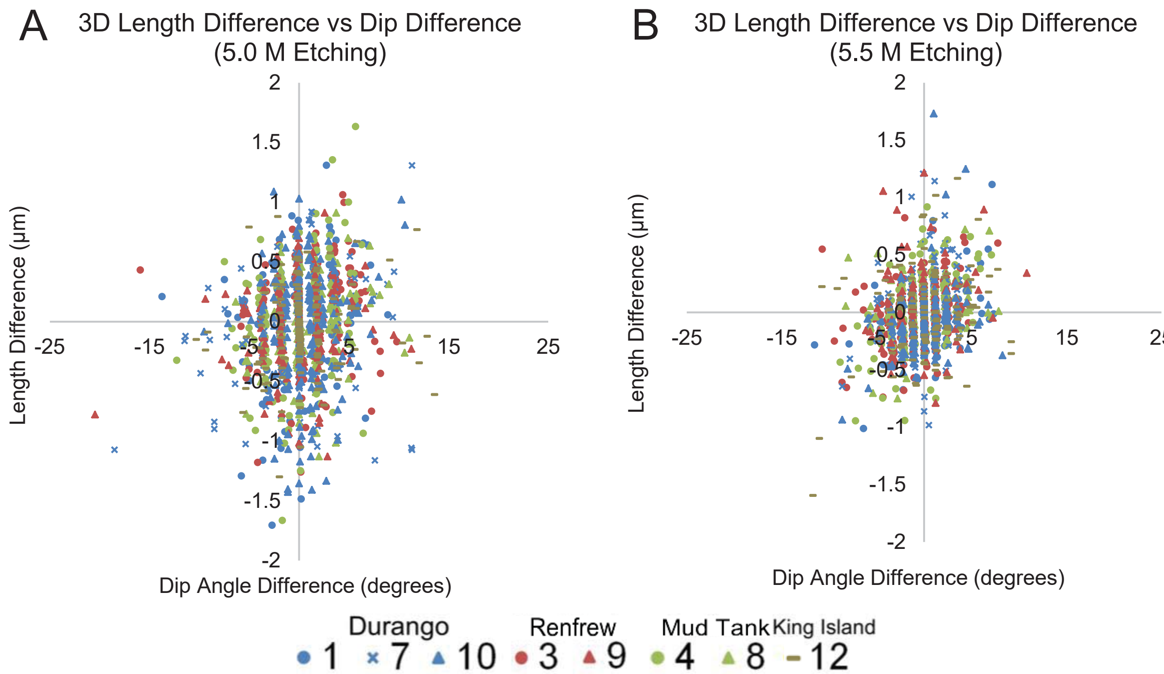

Figure 3: Relationship between the track length and dip angle difference in 5.0M (A) and 5.5 M (B).

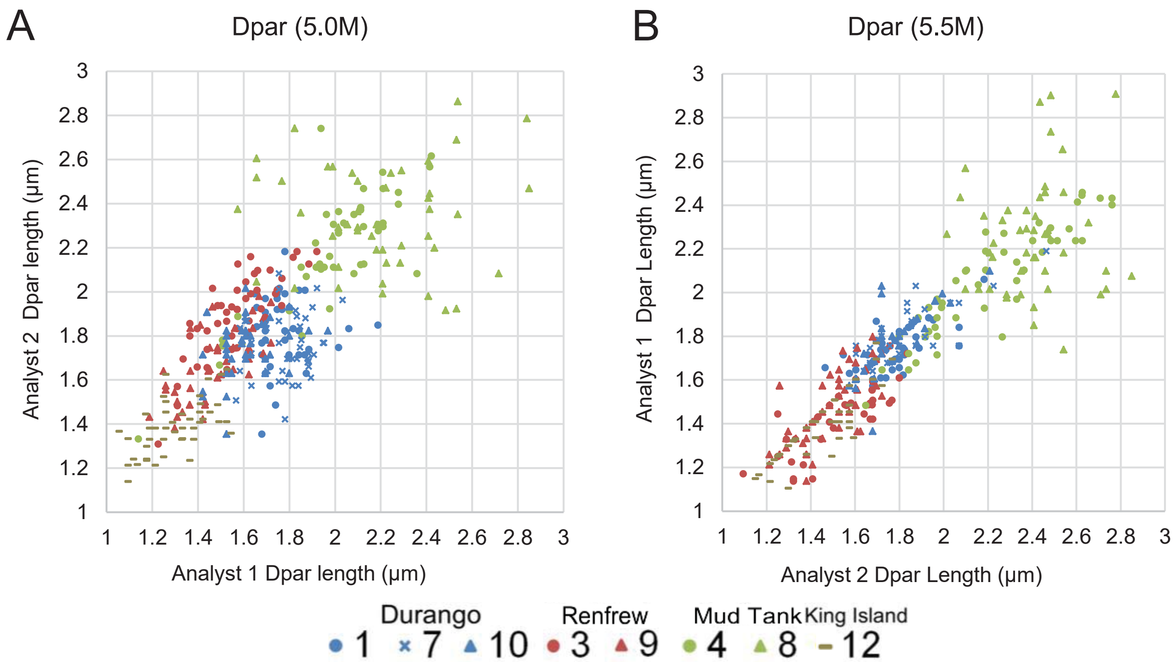

Figure 4: Comparison of individual Dpar measurements by each analyst with 5.0M (A) and 5.5 M (B) protocols.

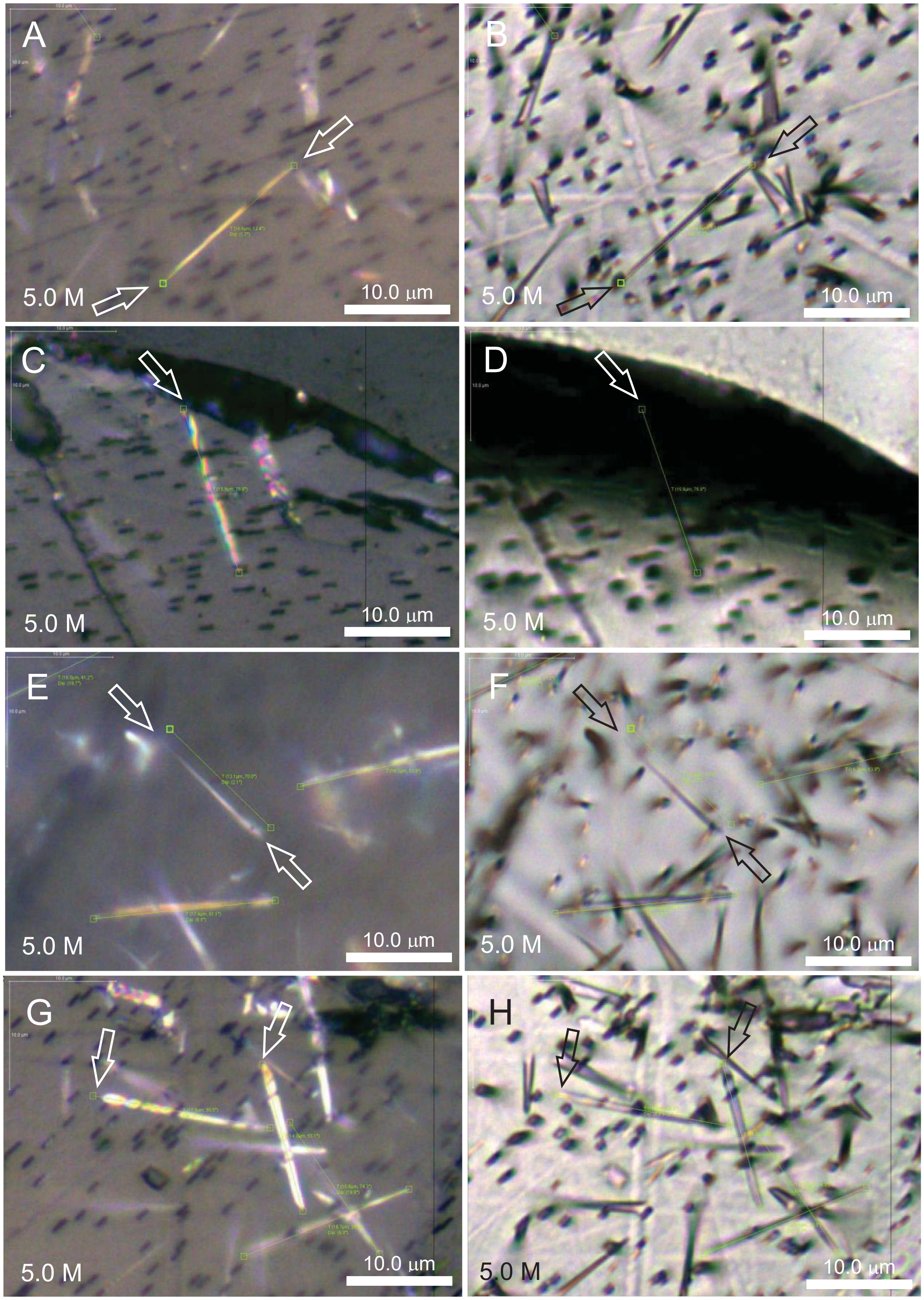

Figure 5: Reflected (left) and transmitted (right) light images of confined tracks found and measured by An1 and rejected by An2. Arrows indicate the features responsible for rejection. In A the track ends are arguably ambiguous, and in B they are obscured by other features. C and D show a track that is not unambiguously confined on both ends in reflected light, while in transmitted light

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2019-7046

half of the track is not visible. E and F show a track with indistinct ends. There are two rejections in G and H. The right track ends are distinct under reflected light but blurry under transmitted. The left track appears to intersect the surface.

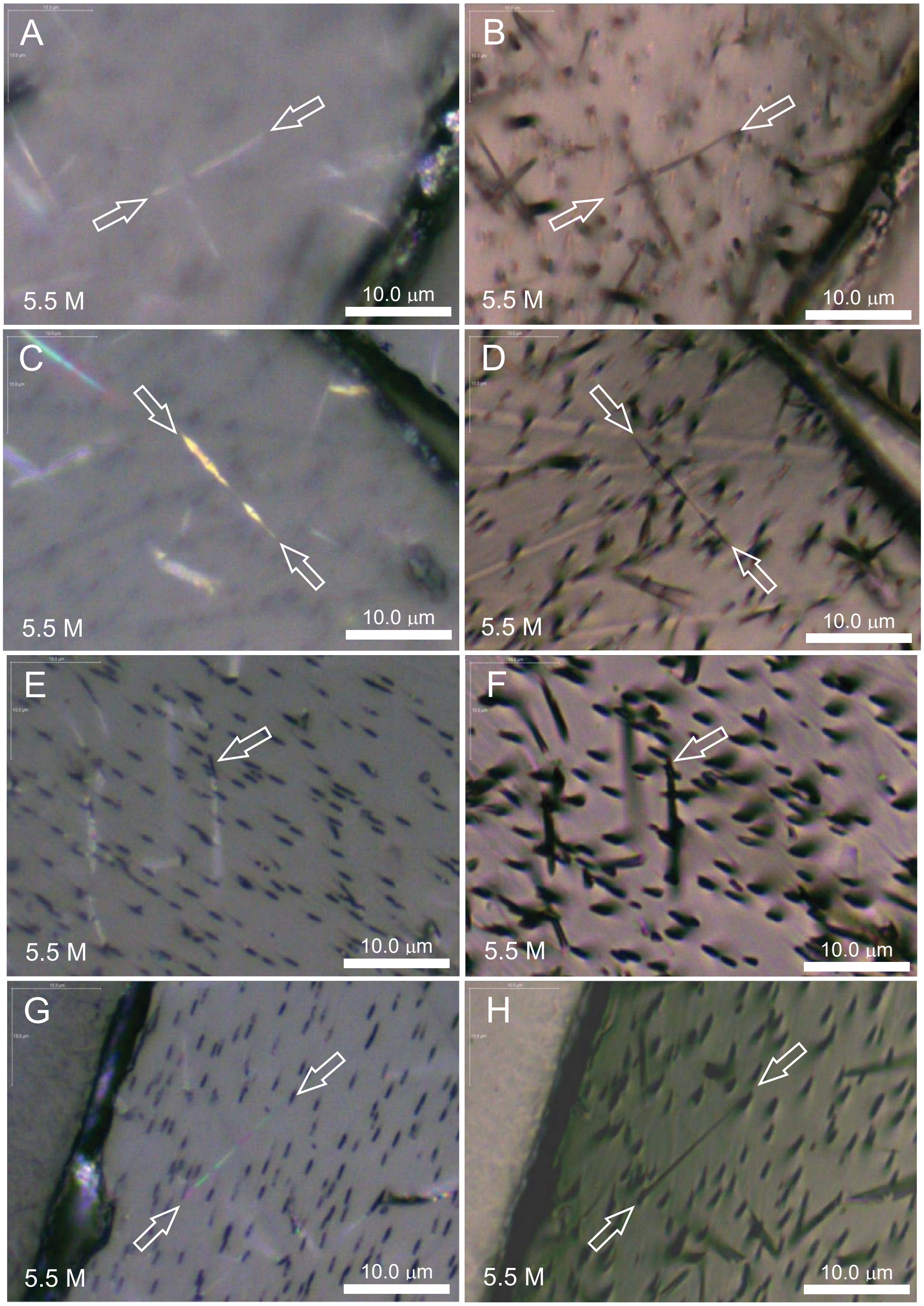

Figure 6: Reflected (left) and transmitted (right) light images of confined tracks found and measured by An2 and rejected by An 1. Arrows indicate the features responsible for rejection. Track ends in A and B are ambiguous. The track in Figure 5C and 5D was judged as under-etched

and thus cannot be measured confidently. Figure 5E and 5F show a track that appears to be a confined track under transmitted light but has a clear surface intersection under reflected light. Figure 5G and 5H show a confined track where both ends are obscured by other features under transmitted light and one end (top-right) is still difficult to determine using reflected light.

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2019-7046

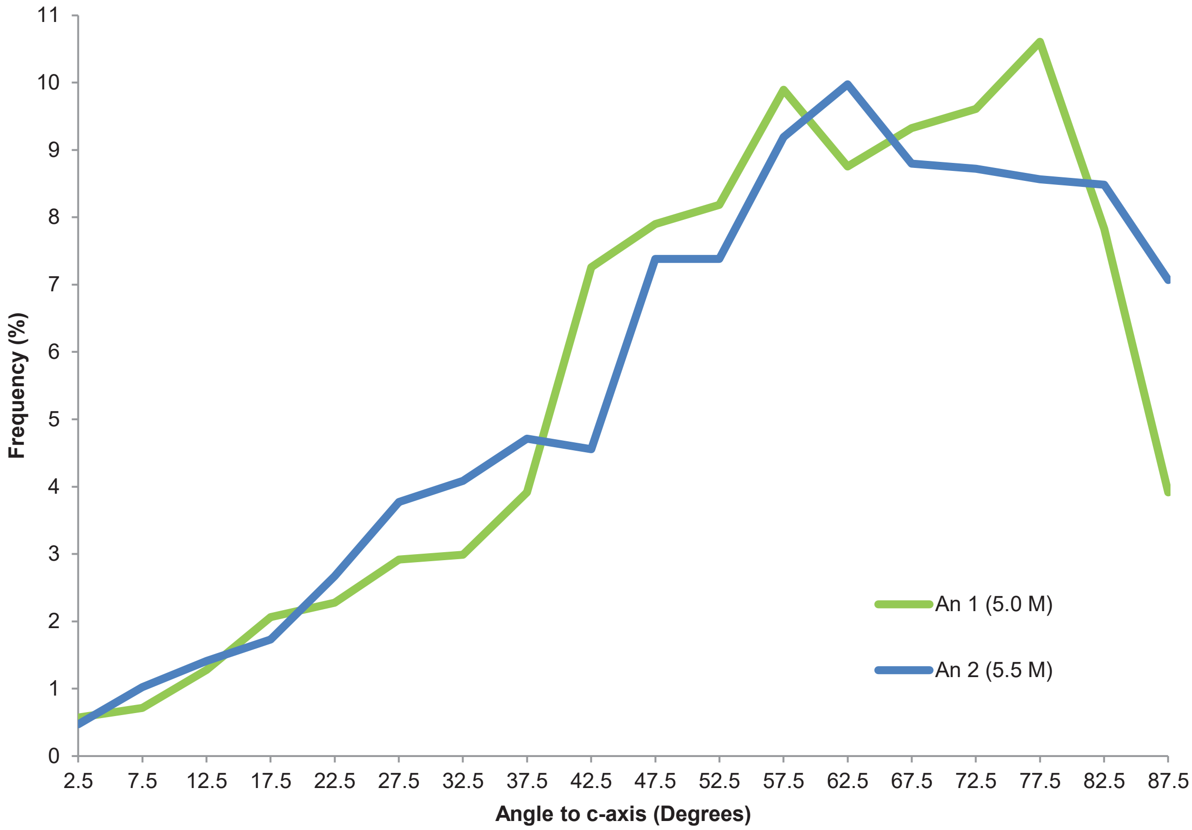

Figure 7: Comparison of the distribution of the c-axis angles of confined tracks in Set 1 measurements by the two analysts. All tracks represented in this plot were also evaluated and accepted by the non-finding analyst, and so variations reflect differences in the initial scanning and evaluation process.

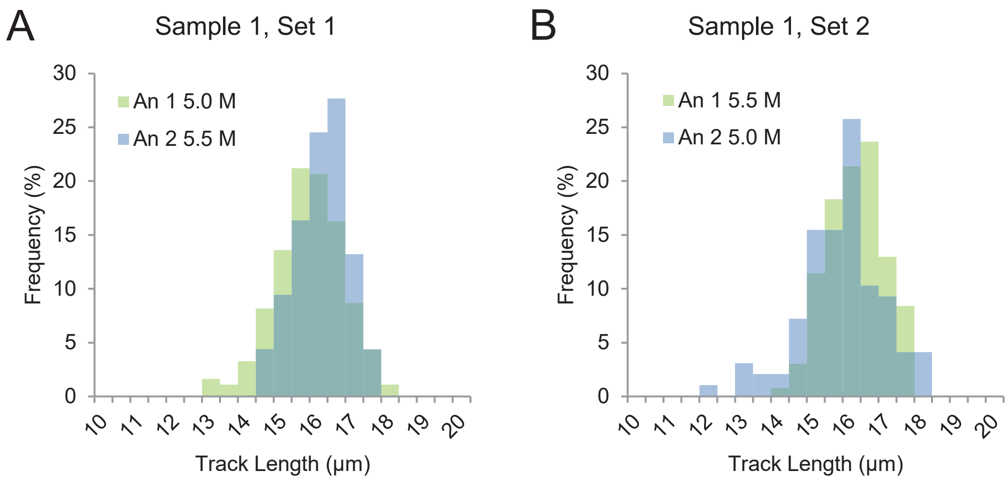

Figure 8: Histograms of confined fission-track lengths in Sample 1 located by both analysts using both etching procedures. According to both analysts, the 5.5M etch produced a slightly longer mean track length (15.90 and 15.84 µm in A and B, respectively), and lower standard deviation (0.71 and 0.76 µm), than the 5.0M etch (15.52, 0.95 and 15.47, 1.12 µm), showing that the differences are due to the etchant, and not the analyst.

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2019-7046

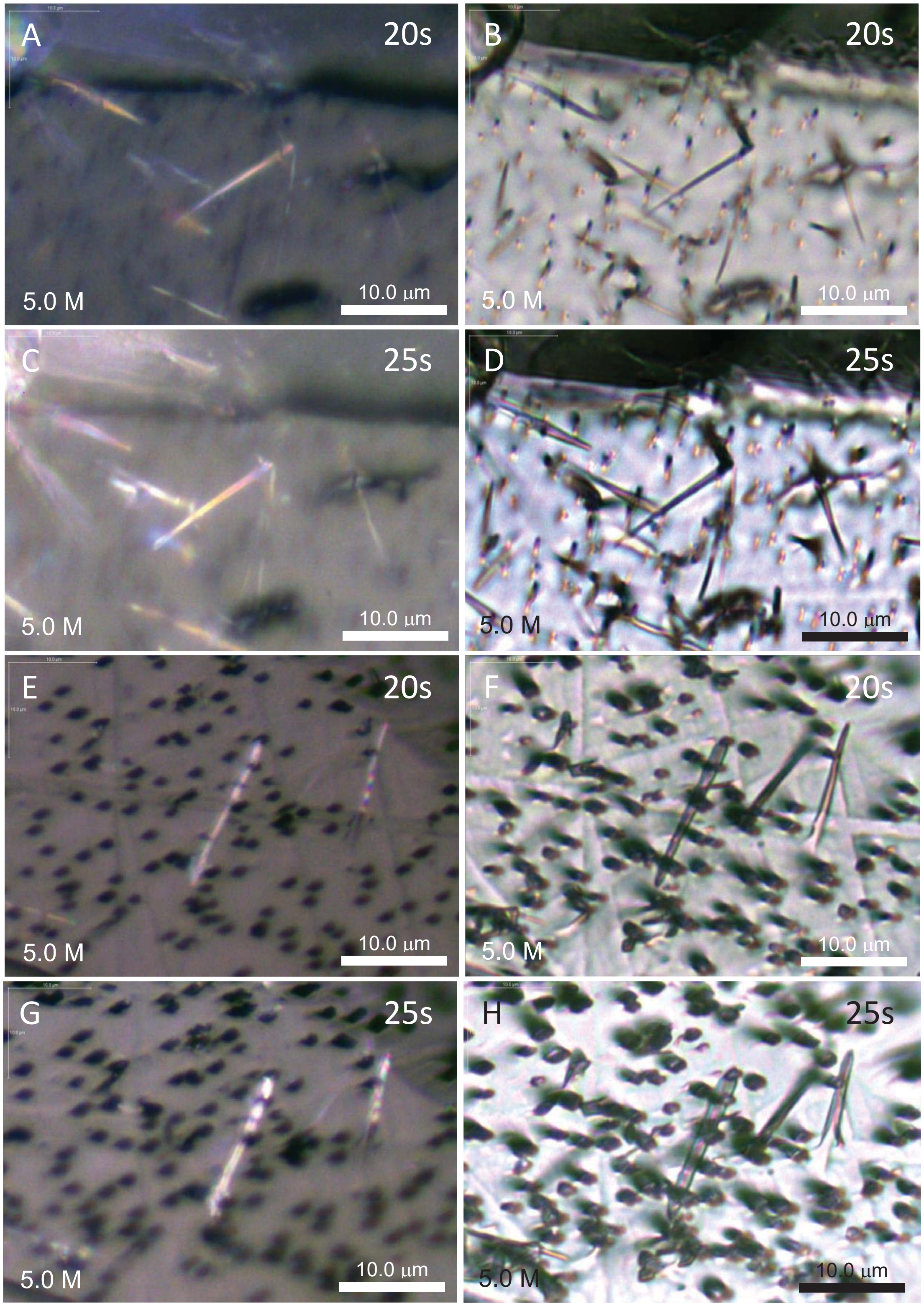

Figure 9: Images of two confined tracks in reflected (left) and transmitted (right) light at 20s (A,B) and 25s (C,D) etching in Durango apatite and 20s (E,F) and 25s (G,H) etching in Mud Tank apatite, both with the 5.0M protocol.

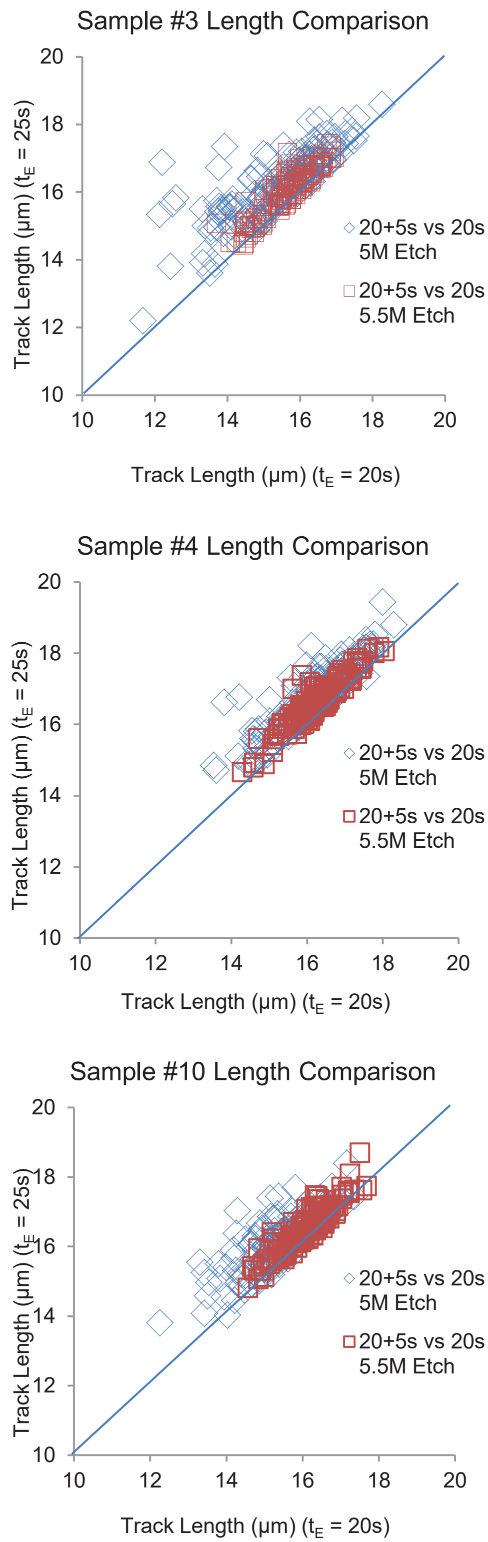

Figure 10: Comparison of 20s against 25s step-etch data on individual track lengths in three samples with both etching protocols; 5.0 M in blue, 5.5 M in red. The blue reference line shows the 1:1 relation.

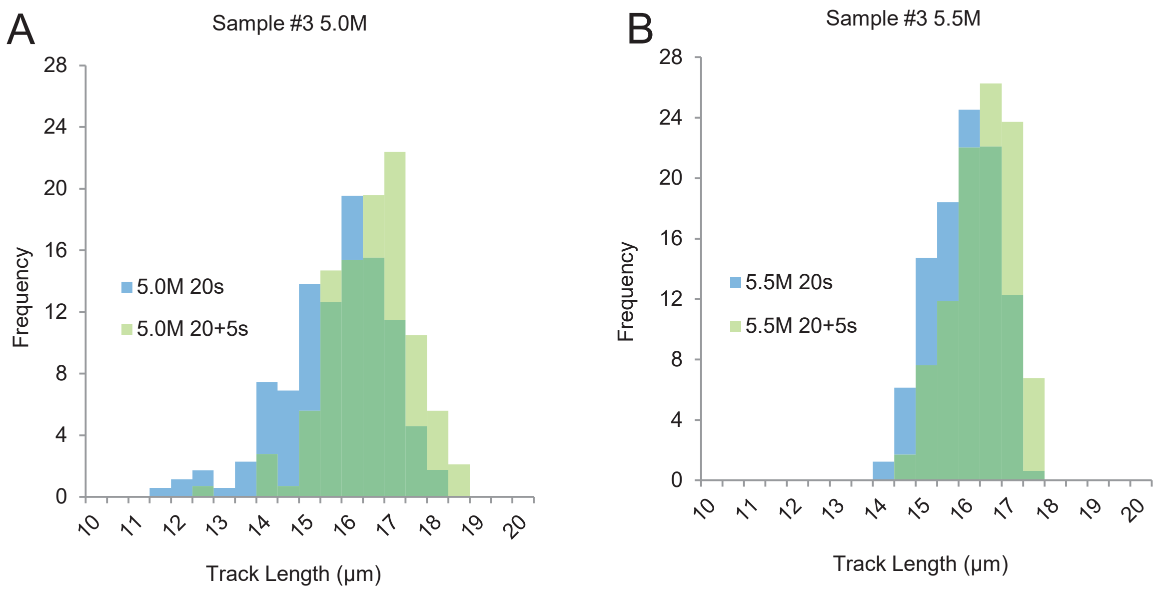

Figure 11: Confined track length distributions after 20 and 25 seconds of etching time using 5.0M (A) and 5.5M (B) etchants at 20°C.

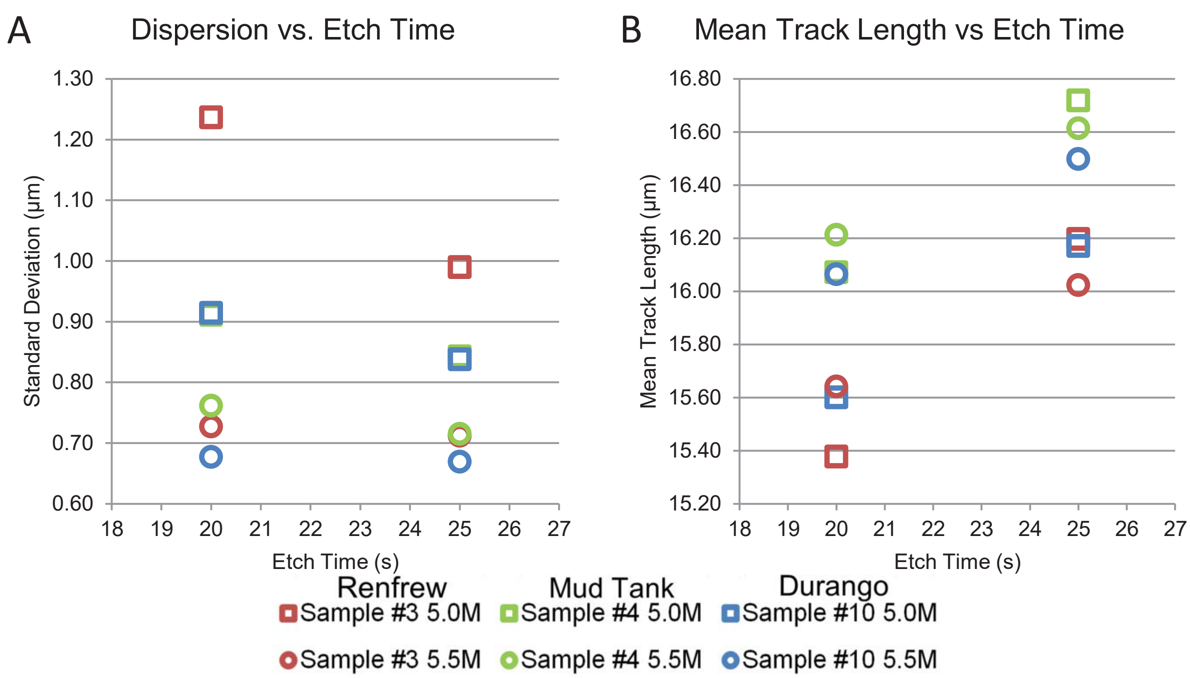

Figure 12: Standard deviation (A) and mean lengths (B) of confined fission tracks vs. etching time. Error bars are omitted for clarity.

This is a preprint, the final version is subject to change, of the American Mineralogist (MSA)
Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2019-7046

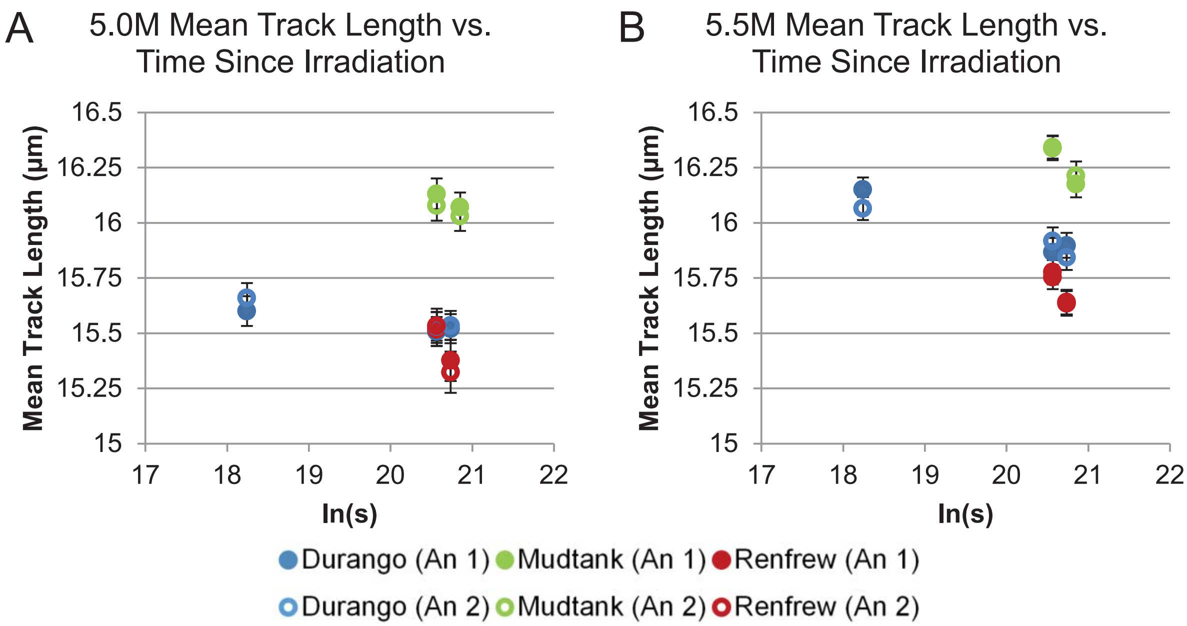

Figure 13: Changes in mean induced track lengths according to time that samples experienced ambient temperatures (~20°C) after irradiation and prior to etching using 5.0M (A) and 5.5M (B) protocols. Error bars are one standard error.

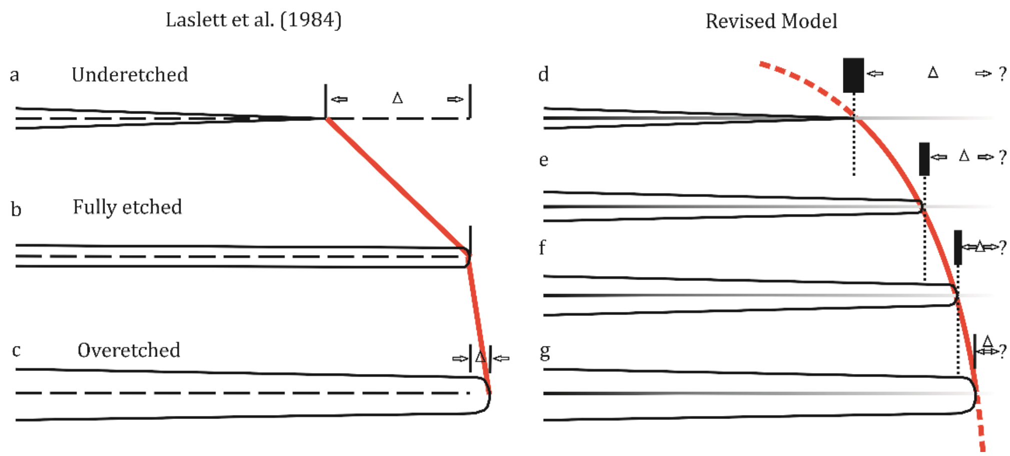

Figure 14: Track etching model from Laslett et al. (1984) and proposed revised model. The Laslett et al. (1984) model is built upon the assumption of a constant etch rate along the latent track, indicated by dashed line, while in the revised model the track etching rate changes along the track, with lighter shades of gray indicating slower etching, approaching the bulk etching rate. The black boxes symbolize the measurement precision. The delta symbols indicate distance from the etched track to the "true" track end; the question marks denote that this quantity is not straightforward to define due to the gradual reduction in etch rate.