https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

The Need for Fission Track Data Transparency

Murat T. Tamer[1,2], Ling Chung[3], Richard A. Ketcham[2], Andrew J. W. Gleadow[3]

[1]State Key Laboratory of Earthquake Dynamics, Institute of Geology, China Earthquake Administration, Beijing 100029, China
[2]Jackson School of Geosciences, The University of Texas at Austin, 78712 Austin, Texas, U.S.A.
[3]School of Geography, Earth and Atmospheric Sciences, University of Melbourne, 3010 Melbourne, Victoria, Australia

Correspondence to: Murat T. Tamer (murat@ies.ac.cn)

Abstract. We report a new image-based inter-analyst study to investigate fission-track grain selection and analysis by 13 participants from an image data set that included grains of variable quality. Results suggest that participants with less experience show a higher rate of selecting unsuitable grains, while participants from the same laboratories generally provide similar results. Less analysis experience may result in the rejection of suitable grains, or inclusion of unsuitable ones. While inappropriate omission and inclusion can both bias results, the latter is more pernicious due to the standard practice of achieving a predecided number of analyses; particularly in difficult samples, there is a danger of "squeezing the rock" by weakening selection criteria. Juxtaposing selected regions of interest (ROIs) on the same grains indicates that zoned grains and grains with inclusions and defects yield varying track density estimates, indicating that ROI placement can be an influential factor. We propose developing image data repositories for global data transparency, a global guidance for fission-track analysis, digital teaching modules, and open science. We also point out the need for new approaches for zeta calibration that include consideration of grain quality, methods of uranium determination, and etching protocols.

1 Intoduction

Fission-track dating and thermal history modeling are widely used for near-surface research in earth sciences, across a large spectrum of subjects such as landscape evolution (Reiners and Shuster 2009; Lemot et al., 2023; Gallen et al., 2023), climate change (Barnes et al., 2012; Qiu and Liu 2018; Yu et al., 2022), glacier-driven exhumation (Balestrieri et al., 1991; Fitzgerald and Goodge 2022; Karaoğlan et al., 2023), natural resource exploration (Dumitru et al., 1991; Deng et al., 2015; Qiu et al., 2023; Gülyüz et al., 2024) and biodiversity (Kohn et al., 1992; Torres et al., 2013; Bernet et al., 2023). Five essential 'ingredients' are required for fission-track time-temperature modeling: the track densities calculated from (1) track counts over (2) a selected region of interest; (3) preferably more than a few tens of confined track lengths per sample; (4) average etch pit diameter (Dpar) measurements per grain, or chemical information to infer kinetics; and (5) an estimate of the 238U concentration (Tagami and O'Sullivan 2005). While laser ablation mass spectrometry has become an alternative (Hasebe et al., 2014) to the widely used external detector method (EDM) (Gleadow and Lovering 1977) for uranium content determination, the first four inputs are still largely analyst-driven, although recent developments in image analysis and AI have

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

contributed significant advances in auto-counting and auto-measurement (Gleadow et al., 2009, 2019; Nachtergaele and De Grave 2021; Li et al., 2022; Ren et al., 2023; Boone et al., 2023). Previous apatite fission-track inter-laboratory and inter-analyst experiments showed significant variation in measurements for the same samples and even standards (Naeser et al., 1981; Miller et al., 1985, 1990, 1993; Barbarand et al., 2003a; Ketcham et al., 2009; Sobel and Seward, 2010; Ketcham et al., 2015; Ketcham et al., 2018; Tamer et al., 2019). These variations have been attributed to a broad range of factors including instrumentation, analytical preferences, etching protocol, and analyst selection criteria.

The process of selecting suitable grains, and regions of interest within those grains, has been relatively unexplored in these studies but may exert a first-order influence on the data quality and extractable thermal history information. From a given set of grains, grain selection influences results in several ways. Grains where oily fluids have penetrated into tracks may hinder the recognition of some surface tracks, causing confined tracks to appear shorter and thus more annealed (Ketcham et al., 2015), leading to underestimation of ages and overestimation of temperatures. Grains with excessive defects, such as polishing artefacts or etched dislocations, may cause misidentification of some spurious features as actual tracks and cause overestimation of ages. Track density can vary by up to 35% if the grain is oriented without the c-axis in the viewing plane (Aslanian et al., 2022), making both the resulting age and etch figure dimensions (e.g. Dpar) inaccurate, thereby affecting estimates of kinetics and initial track length. A perceived need to meet targets for the number of grains analyzed may cause an analyst to select borderline-acceptable grains or tracks that may not have been selected otherwise (Tamer and Ketcham, 2023).

Whereas the area counted for fission-track density determinations has typically been defined by boxes in an eyepiece reticule, modern image-based systems allow the user to draw an arbitrarily shaped region of interest. In both cases, this process must be executed with care. Regions of interest need to be placed so that the grain surface they encompass is not biased with respect to the ability to host detectable tracks. Regions of interest within one fission-fragment range of the grain edge will not sample tracks from a full 4π geometry (Donelick et al., 2005), and including sizable defects and cracks in the region of interest may result in uncountable areas; both effects will bias ages lower. Regions of interest that include zones with different U content complicate the accurate determination of U across the track-generating region (Vermeesch, 2017), and suffer edge effects from sampling a 4π region that is variable. This bias can result both from methods using laser ablation, where typically a smaller area is sampled for the U-determination than for the spontaneous track count, and the external detector method, where perfect matching between spontaneous and induced track regions of interest can be difficult to achieve, especially where the track density is low.

The suggested number of grains for age measurements for igneous-type samples is typically ~20, or more if there is any indication of kinetic variation (Donelick et al., 2005), while for detrital samples it is ~120 or more (Vermeesch 2004). If grains

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

are few or of low quality, an analyst may consciously or unconsciously "squeeze the rock" by adding some borderline-quality grains to provide the expected data quantity. Aiming for specific quantities of data may eventually cause a loss of quality.

We carried out a new inter-analyst experiment designed to investigate variability in grain selection and region of interest definition criteria. Building upon a previous two-analyst study (Tamer et al., 2019), we also tested the identification of confined track lengths. Participants were asked to perform apatite FT analysis on a selection of grains drawn from an identical image set featuring variable grain quality using software that records all details of the analysis as overlays in a .xml file, thereby allowing for subsequent review. Analysts were also asked to fill out a questionnaire about their approach.

2 Materials and Methods

2.1 Image Data Repository

We created an image data repository consisting of 41 grain and 3 graticule images from the in-house fission-track data repositories at the University of Melbourne (UM) and the University of Texas at Austin (UT). Images from UM were captured by Ling Chung using a Zeiss Axio Imager M1m microscope with an IDS µEye camera and white balance correction, while images from UT were taken by Sean Sanguinito and Murat Tamer, using a Zeiss Axio Imager M2m microscope with Olympus SYS UC30 camera and no white balance. Grains from UT were etched with 5.5M HNO3 at 21 ℃ for 20s (Carlson et al., 1999), while the grains from UM were etched with 5M HNO3 at 20 ℃ for 20s (Gleadow et al. 1986). 36 grains were selected for track density measurements and 5 for confined track length measurements. To test the self-reproducibility of the analytical results we repeated one grain image as two different grain areas (Grains 07 and 16). The grain descriptions are given in supplementary file Table S1.

2.2 Announcement and Participant Instructions

The announcement of the study was made at the 17th International Conference on Thermochronology, 2021, Santa Fe (Tamer et al, 2021) and in relevant email lists. The participants were asked to perform track density and confined track length and Dpar measurements using their preferred approach, including any analytical software, manual measurement, or AI-based analysis. The participants were not instructed to reach a given number of grains or confined track length analyses but were instructed to skip or accept grains for analysis according to their own judgement.

This experiment was made possible by Fission Track Studio, a dual software suite developed by the Melbourne Thermochronological Research Group (MTRG) that is capable of automatic grain stack-image acquisition (TrackWorks) and image review and measurement (FastTracks). The FastTracks program offers manual and automated analytical tools for obtaining all essential parameters for FT dating as well as a cross-section tool for precise dip angle determination for length measurements. All analytical results were recorded in an .xml file that can be reloaded for a follow-up analysis and review.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

The University of Melbourne provided a temporary FastTracks license and a detailed user manual for those who wanted to participate in this study. The participants had the option to reveal their names and affiliations or to be anonymous. A participant's submission was accepted only if the analysis was performed by a single analyst.

2.3 Reviewer Criteria

In the absence of absolute standards, we used the grain selection criteria of L. Chung and the confined track length measurement judgments of M. Tamer as reference points for the review of the participant results. However, no fission-track analyst can claim complete certainty in their judgments about track features and we do not suggest that these reference results represent 'true' values. Rather they are simply used as reference values that are probably typical of reasonably experienced analysts. They were used as the starting point for a detailed grain-by-grain and track-by-track discussion with the participants to arrive at a consensus view and to ascertain which factors are most likely to lead to discrepancies between analysts. Such a detailed analysis has not previously been undertaken to our knowledge and would be all but impossible without the image-based approach used here. To test the objectivity of the reviewers, the participants were shown those of their selections and measurements that the reviewers considered questionable or in some cases, erroneous, after which they could acknowledge or dispute the review.

Grains were judged to be suitable, unsuitable, or borderline. A grain having any of the three following properties was considered unsuitable: (1) the polished surface was not parallel to the c-axis (Fig. 1 a, b), (2) fluids were present in tracks (Fig. 1 c, d), (3) an excessive number of inclusions/defects/uncertain features was intermingled with actual tracks (Fig. 1 e, f). Additionally, heterogeneous U distribution within the grain, judging from the distribution of spontaneous tracks, can be a complicating factor, especially if LAICPMS spot analysis is used for U determination, but also from misalignment of the spontaneous and induced track regions of interest using the EDM. In samples with low abundance and/or low-quality of grains, some borderline-quality grains may be included in the resulting data sets (Fig. 1 g, h). Selection of the region of interest may become a challenge for inclusion/defect-rich (Fig 1 i, j) and zoned grains (Fig 1 k, l). A confined track length is measurable as long as both ends are not exposed (Fig 1 m, n) at the surface, visible and unobscured by surrounding features (Fig 1 o, p). Moreover, it is important that confined tracks are not filled, fully or partially, with oily fluids, such as may result from fingerprints, which significantly affects their optical contrast within the host mineral (Fig 1 q, r). Any measurement that does not meet these criteria is considered invalid.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

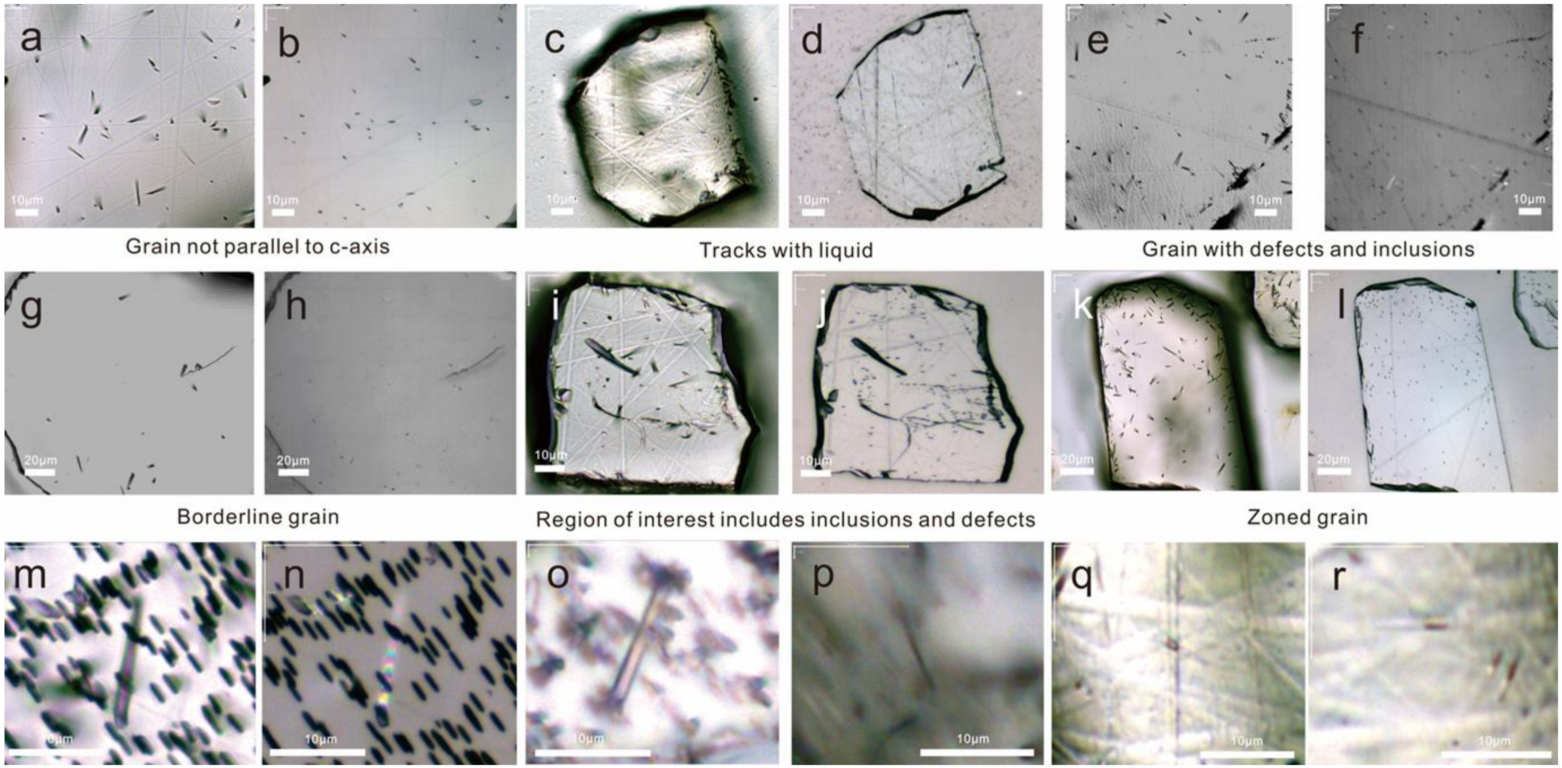


Figure 1: Examples of grains and tracks that could cause erroneous grain and region of interest selection or invalid confined track length measurement for fission-track analyses. A grain not parallel to the c-axis appears to be acceptable in transmitted light (a), but varying etch pit orientations clearly show that the grain is not parallel to the c-axis in reflected light (b). Transmitted and reflected light images showing how a dirty grain surface with liquids in tracks can impair track recognition (c,d). Transmitted and reflected light images of a grain showing an excess of non-track features such as defect and inclusion that are difficult to distinguish from actual tracks (e, f). Transmitted and reflected light images of a grain with a low and uneven track density, suggesting the possible presence of zoning that would need to be accounted for (g, h). Transmitted and reflected light images of a grain showing large defects and inclusions occupying space obscuring fission tracks, resulting in underestimation of the fission-track density if these features are included in the region of interest (i,j). Transmitted and reflected light images of a zoned grain, for which different placements of the region of interest and location of ablation point(s) may yield divergent ages (k,l). Transmitted and reflected light images of a would-be confined track, where the reflected light image shows one tip may be exposed to the surface (m,n). Confined tracks may also be rendered invalid for measurement by obscuring features (o,p), or partial fluid fillings (q,r).

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

3 Results and Discussion

Thirteen analysts returned completed questionnaires, though one of them could not provide the .xml file due to technical problems. Although L. Chung and M. Tamer evaluated each other's analyses, we consider them as reviewers in this study. Two additional analysts from the same laboratory with different years of experience submitted answers in one combined .xml file. Because they were unable to disentangle their results or reconduct them independently, their results are not included in the analysis below. While some participants wished to remain anonymous, others chose to be transparent with their identities; Table S2 provides the list of participants. Participants' overall and recent experience and their current fission-track setup, a summary of the questionnaire, selection rates for grains, and validity rates for confined track lengths are provided in Table 1. Excluding one of the repeated grains for checking self-reproducibility (Grain 16), of the 35 grain image sets, L. Chung estimated that 22 are suitable and 13 are unsuitable for fission track analysis. In grain-by-grain checking, we counted how many of the suitable grains were selected and how many of the unsuitable grains were rejected. The reasulting suitable and unsuitable grain selection rates are reported as percentages. M. Tamer examined every measured confined track length to check validity as determined above. The percentage of valid measurements is reported as the confined track length measurement validity rate. We did not evaluate how many valid tracks were excluded, as there was no way to determine whether such tracks were intentionally omitted or simply missed.

3.1 Graticule Calibration

Graticule images taken by microscopes at UM and UT were included in the data set for calibration. Although calibration is an essential step before performing an analysis, only five participants reported measuring them. Some omissions may have been due to not fully understanding the terms of the experiment. To make the comparison of results easier, we used the default graticule calibration for all analysts.

3.2 Self-Reproducibility

Grains 7 and 16 are duplicated images of the same grain in our data set. While some participants skipped Grain 16 after noticing the repetition, some performed density measurements on both grains. Although these remeasurements demonstrated high self-reproducibility (Fig 2), minimum and maximum densities vary by ±30%. The difference can be traced to the varying region of interest selection, light source preference, and track counting routines. Although the zeta method (Hurford and Green, 1983) is intended to normalize some differences among analysts, the degree of variation shown here is more severe than that implied by typical variation in zeta factors (~20%).

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

Analyst	Total years of experience	Activity in the past two years	In-house fission-track setup	Auto c-axis assignment tool	Graticule measurement	Total number of grains selected as suitable	Suitable grain selection rate (%)	Unsuitable grain selection rate (%)	Number of measured confined track lengths	Confined track length measurement validity rate (%)
1	7	Yes	AS	Yes	No	22	100	8	16	69
2	14	Yes	AS	No	Yes	22	100	0	14	Reviewer
3	7	Yes	AS	No	Yes	29	95	62	31	32
4	6	Yes	AS	Yes	No	23	91	23	8	88
5	5	Yes	AS	Yes	No	11	50	0	19	63
6	4	Yes	C	Yes	No	26	95	39	5	80
7	6	Yes	C	No	No	26	100	31	10	60
8	4	Yes	AS	No	No	35	100	100	16	56
9	5	Yes	AS	No	No	34	100	93	33	55
10	17	Yes	AS	Yes	Yes	22	Reviewer	Reviewer	18	94
11	30	No	D	No	Yes	22	95	8	16	81
12	2	Yes	AS	No	No	31	95	77	4	25
13	5	Yes	AS	No	No	28	95	54	49	30
14	9	Yes	AS	No	Yes	24	100	15	18	78
15	2	Yes	AS	No	No	22	86	23	22	59
16*	40	Yes	C	N/A	N/A	22	100	0	N/A	N/A

Table 1: Summary of the questionnaire, rates of grain selections, and confined track length measurement validities.

AS: Autoscan, C: Custom; D: Dumitru System. *: The analyst did not participate in the experiment but evaluated the grains as suitable and unsuitable.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

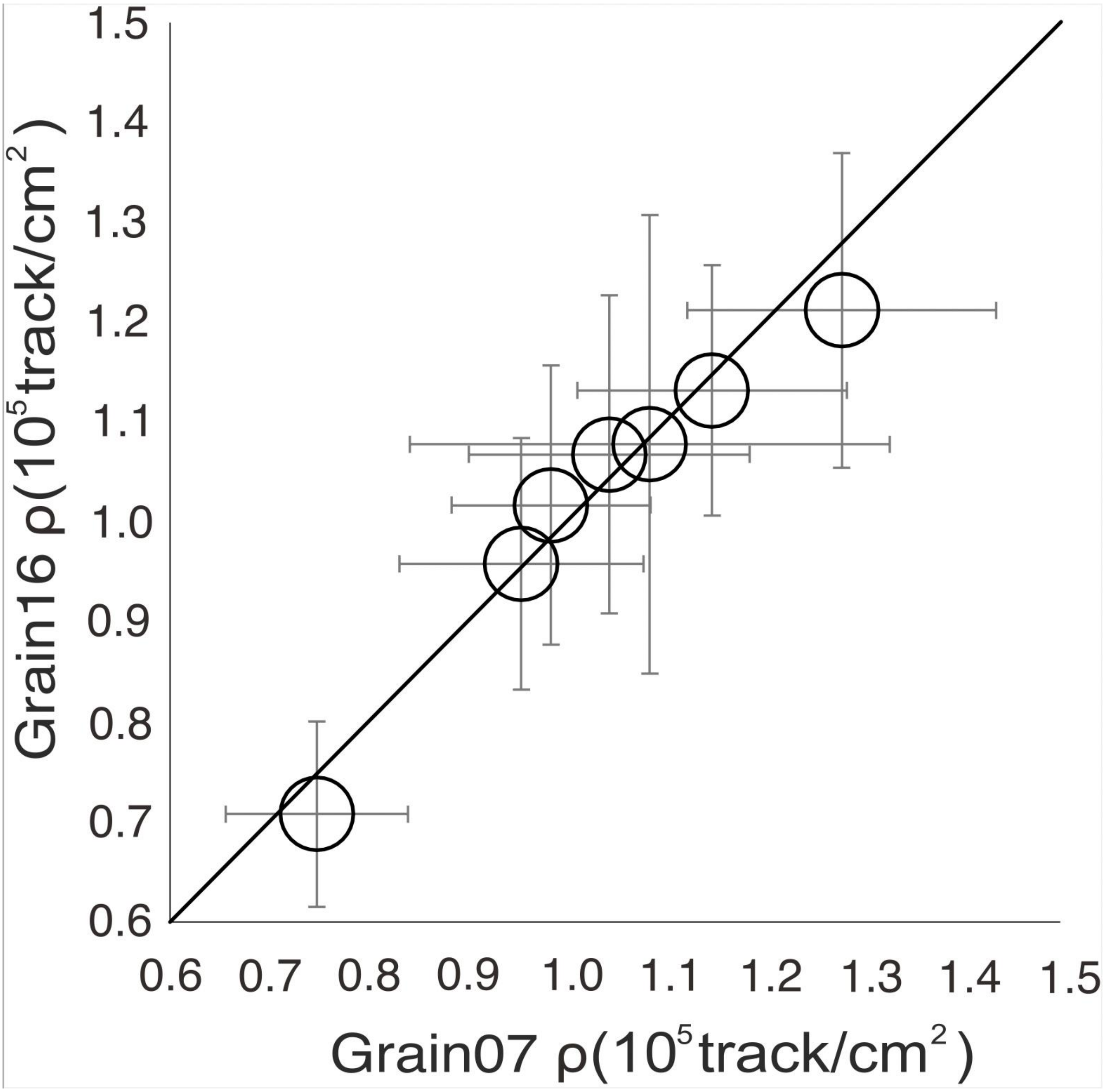


Figure 2: Self-reproducibility of track density (ρ) determinations on replicated grain images for seven analysts.

3.3 Post-review follow-up and objectivity of the review

After the initial review of grain and confined track length measurements, a follow-up meeting with each participant was conducted to discuss each judgment deemed questionable or unsuitable by the reviewers. Virtually all participants

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

acknowledged all inappropriate grain selections and confined track length measurements, except Participant 6 considered Grain 08 and Grain 31 as borderline instead of unsuitable. This high rate of acknowledgment by the participants supports the soundness of the criteria utilized by the reviewers. According to the participants, inappropriate selection and measurements stemmed from different factors. While some participants cited a lack of attention to details (e.g. poor identification of track ends), others stated that they have been choosing some unsuitable grains in their routine fission-track studies since their training. Some of the participants mentioned that they knowingly added unsuitable grains to the data sets in the past to meet the expected number of grains. Some of the participants used FastTracks' automatic tools for c-axis orientation and dpar length measurements.

3.4 Track Density and Confined Track Length Distributions

The track density and confined track length distributions of each participant show several patterns (Fig 3). The density distributions of suitable grains appear to be somewhat more relatable and consistent than for unsuitable grains, and the inclusion of unsuitable grains in all cases skewed the track density distribution to lower values. A similar effect tends to be present in the confined track length distributions, but not in all cases. The confined track length distributions show that the analysts observed and measured different numbers of confined track lengths with a varying number of measurements on valid tracks. Combining all valid tracks yields a histogram with confined track lengths ranging from 8 to 17 µm. An analyst adopting different selection criteria may select confined tracks that were exposed to etchant for different amounts of time (Tamer and Ketcham 2023). Participants 1 and 10 and Participants 8 and 9 are from the same two laboratories and show similarities in their respective track density results. This may be related to the shared training and/or analytical routine in counting, though for length measurements participants 8 and 9 had more divergent results, possibly due to different personal selection criteria. Dpar measurements on suitable grains tend to be more similar with some outliers, while the Dpars of unsuitable grains provide higher dispersion (Fig S1). Post-experiment interviews with participants suggested that the dispersion of Dpars on suitable grains may have stemmed from different levels of zoom applied.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

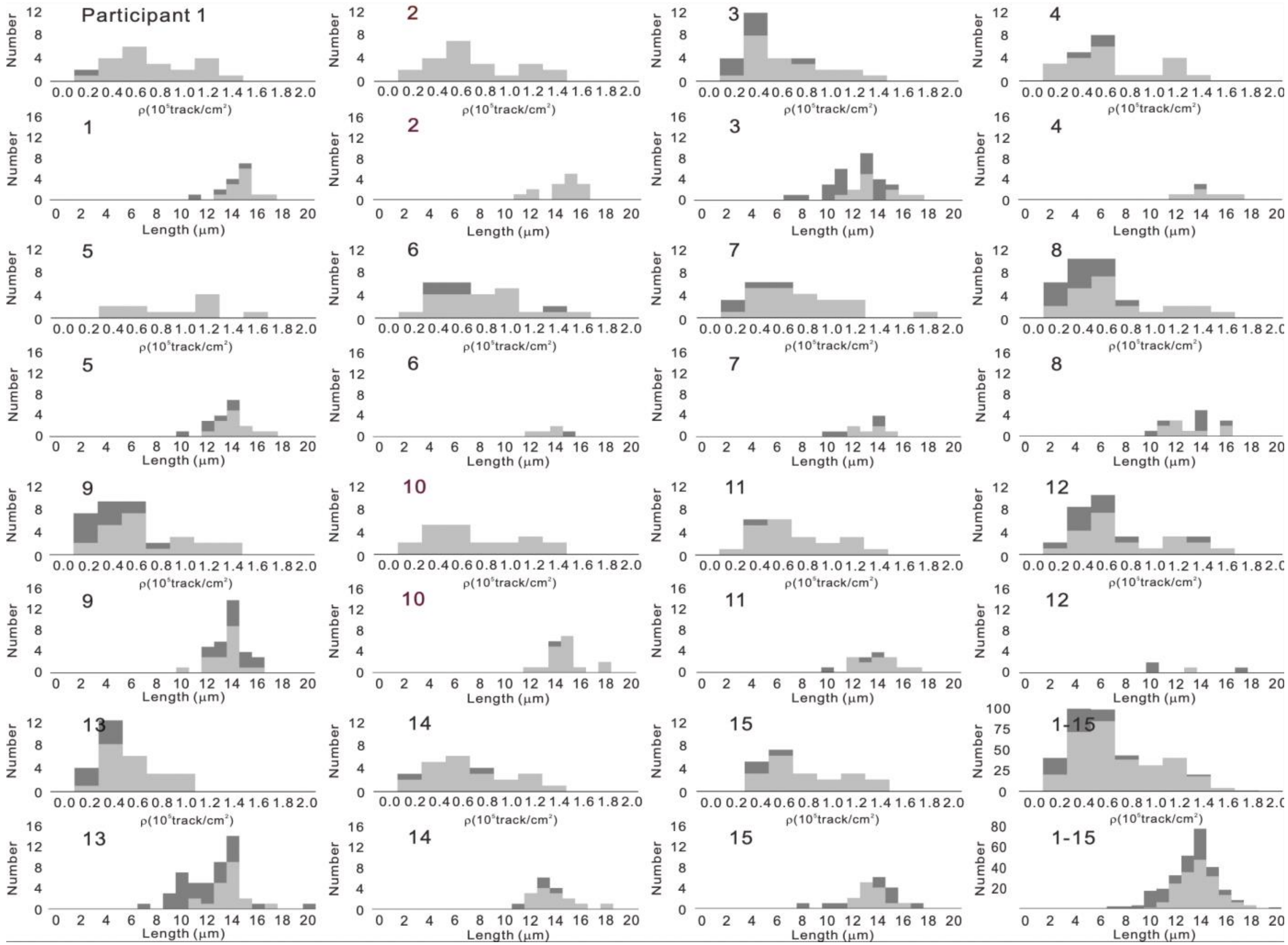


Figure 3: Track density (ρ) and confined track length distributions of each participant and the reviewers (2 and 10, in red). Participant numbers are indicated at the top left. The cumulative result for all participants is shown at the bottom right (1-15). Dark grey shows measurements of grains and confined tracks assessed to be unsuitable, and light grey displays the measurements of suitable selections.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

3.5 Impact of Experience

The track density and confined track length distributions of each participant show several patterns (Fig 3). The density distributions of suitable grains appear to be somewhat more relatable and consistent than for unsuitable grains, and the inclusion of unsuitable grains in all cases skewed the track density distribution to lower values. A similar effect tends to be present in the confined track length distributions, but not in all cases. The confined track length distributions show that the analysts observed and measured different numbers of confined track lengths with a varying number of measurements on valid tracks. Combining all valid tracks yields a histogram with confined track lengths ranging from 8 to 17 μm. An analyst adopting different selection criteria may select confined tracks that were exposed to etchant for different amounts of time (Tamer and Ketcham 2023). Participants 1 and 10 and Participants 8 and 9 are from the same two laboratories and show similarities in their respective track density results. This may be related to the shared training and/or analytical routine in counting, though for length measurements participants 8 and 9 had more divergent results, possibly due to different personal selection criteria. Dpar measurements on suitable grains tend to be more similar with some outliers, while the Dpars of unsuitable grains provide higher dispersion (Fig S1). Post-experiment interviews with participants suggested that the dispersion of Dpars on suitable grains may have stemmed from different levels of zoom applied.

3.6 Region of Interest (ROI) Selection

Track density measurements on defect/inclusion-free grains with homogeneous track distributions may not be greatly affected by different region of interest (ROI) selections, but inclusion of large defects in the ROI can cause an underestimation of density by obscuring tracks (Fig 4d). Similarly, the selection of high and low track density areas within a zoned grain can yield widely varying density determinations (Fig 4e). A single-spot or even dual-spot laser ablation approach on such grains may result in a significant dispersion of dates depending on the analyst's ROI selection. Several participants placed ROIs in too-close proximity to the mineral border (<~10 μm, Fig 4 d,e), where track registration is below the required 4π geometry (Donelick et al., 2005).

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

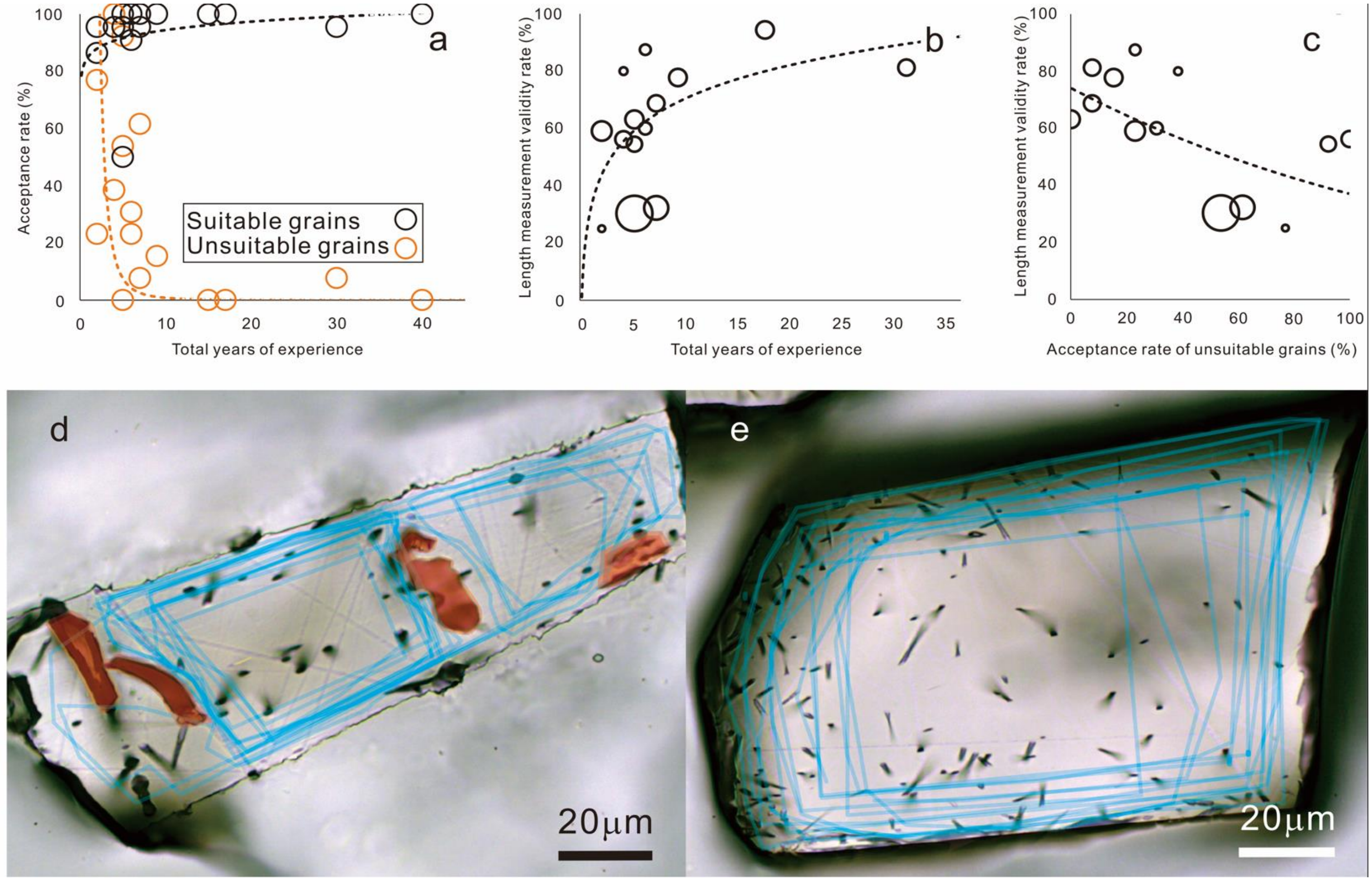
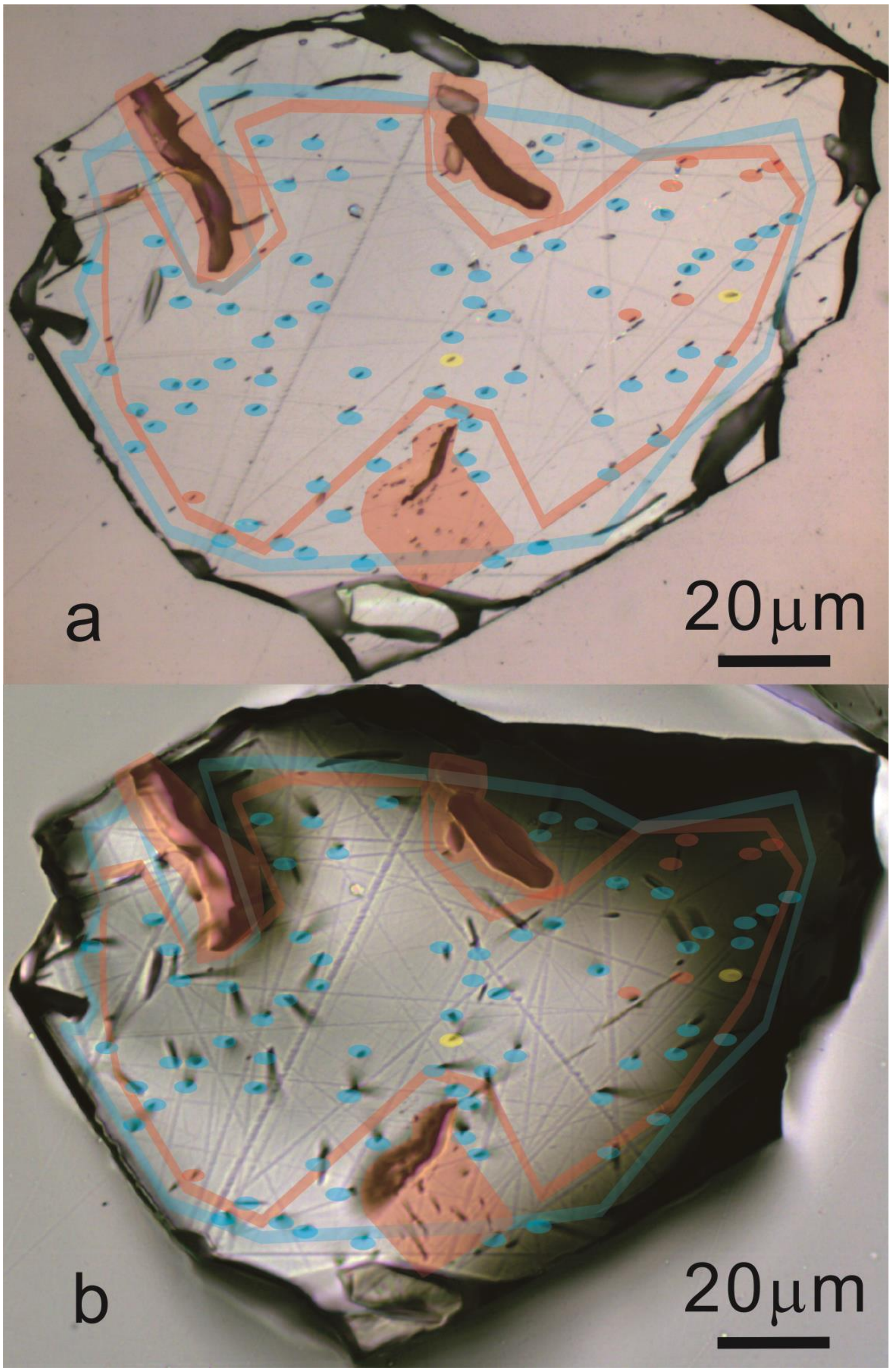


Figure 4: Correlation of years of fission-track experience against acceptance rate of suitable and unsuitable grains for density measurement (a) and confined track length measurement validity rate (b). The confined track length measurement validity rate against the acceptance rate of unsuitable grains is shown in c. The size of the circles in b and c reflect the number of confined track length measurements. Participant selections of the region of interest are juxtaposed in a defect/inclusion-rich grain (d) and a zoned grain (e).

3.7 Region of Interest (ROI) Selection

To demonstrate some effects of the light source and ROI specification on density measurement, at the request of one participant, we reanalyzed their track density for Grain 07, which was significantly lower than most of the group. Figure 5 shows the comparison on a track-by-track basis. Perhaps importantly, this analyst revealed a preference for counting tracks in transmitted light only. However, counting tracks solely in transmitted-light images can cause an underestimation of the track density (Aslanian et al., 2022; Tamer and Ketcham 2023). Figure 5 a and b show the transmitted and reflected light image for this particular grain, with the participant analysis in blue and re-analysis using images from both light sources in red with excluded counts in green. This reanalysis suggested that some genuine tracks had been overlooked, especially in the region on the right that is shaded in transmitted light. After adjustment of the ROI, the addition of the overlooked tracks in reflected light, and the exclusion of a track showing a defect-like pattern, the track density increases by ~35%.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024



265 **Figure 5: Transmitted and reflected light images of Grain 7 (a, b) with the analysis of one participant (blue) and a reviewer's proposed correction (red). Excluded defect areas and tracks.**

3.8. Highlights of participant comments on the data set

Participants 1 and 5 found the images taken with no white balance to be not ideal for counting, and participants 5 and 7 mentioned that the grains etched with the 5.0 M HNO3 20s 20°C (Gleadow et al. 1986) protocol appear to be under-etched.

4 Implications, Suggestions and Conclusions

Analysts may consider unsuitable grains and/or conduct invalid confined track length measurements depending on their years of experience, training, and the difficulty in finding sufficient grains to meet analytical goals. "Squeezing the rock", or indeed "the grain", to extract sufficient data in this way can lead to improper thermal history information. Comprehensive laboratory training and calibrations are essential for fission-track analysts to avoid these problems. Results of graticule and confined track length calibrations and the identity of the analyst should be stated in publications.

ROI selection may cause varying track density determinations, especially in zoned and defect-rich grains. While a single-spot laser ablation analysis is a time-efficient approach, its application on such grains may result in varying U determinations unless the laser spot covers a high proportion of the counted area (e.g., Cogné et al. 2020). Laser uranium mapping (Ansberque et al., 2021) or EDM (Gleadow and Lovering 1977) approaches require more work, but in ideal cases may better represent the selected region of interest. However, precise matching of spontaneous and induced track areas in the EDM can also be difficult in some cases. These approaches may also be more effective in identifying zoned grains when the spontaneous track density is low.

Zeta calibration (ζ) against a set of age standards is intended to normalize for uncertainties in some parameters in the age equation, such as thermal neutron fluence (ϕ) and the spontaneous fission decay constant (λf), and to account for varying counting efficiencies of different analysts (Hurford and Green 1983). This method assumes, however, that a calibration derived from measurements on near-ideal standard samples with minimal inclusions and defects (e.g, Durango, Fish Canyon Tuff apatites) also reflects analyst judgements in unknown samples, which may yield any quality of grains. Zoning, high levels of inclusions and defects, and variable ROI selection may cause significant divergences unaccounted for by zeta calibration, particularly for a less-experienced analyst. These have been studied partially (Vermeesch 2017; Cogné and Gallagher 2021) but further work is needed on this matter.

Among the available confined tracks, analysts can select tracks with different effective etch times based on their individual perceptions and criteria. Some participants highlighted that the grains etched with 5.0 M HNO3 at 20°C for 20s (e.g. Gleadow et al. 1986) appear under-etched, which agrees with a previous inter-analyst comparison experiment (Tamer et al., 2019). A

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

proposed two-step etching protocol (5.5 HNO3 21°C 20+10s) allows analysts to select any suitable track but ensures that the final confined track length data set does not contain under-etched tracks (Tamer and Ketcham 2023).
The application of AI and machine learning methods have become popular topics in various research fields in earth sciences including fission-track counting and confined track length measurements. The quality of any automated analyses will be defined by not only sophisticated algorithms but also the fission-track analysis experience of the initial "teacher" of the AI, including hardware preferences during image acquisition and the resulting image quality. Determining which grains and tracks are acceptable for measurement represents an additional challenge for AI method development; training cannot be based on good images alone but must also include features to avoid based on sometimes subtle indications.

The exposure of the "ingredients" of thermal history modeling is limited to data summary tables and sometimes raw data as supplementary files to research articles. Although fission-track data have generally fared well in inter-laboratory age comparisons in recent years, these have tended to utilize relatively straightforward samples. This study illustrates some of the potential hazards of fission-track analysis of more challenging materials, but also pathways toward improving data reliability. In particular, the opportunity is coming into view for the fission-track community to share data on a new level, allowing analysts to see, learn from, and discuss each others' image data.

Recent developments in data repositories and metadata reporting are healthy signs of an emerging open science culture and up-to-date reporting in low-T thermochronology. However, these are currently limited to collecting and presenting the data in their corresponding geo-locations (Boone et al., 2022; Boone et al., 2023) and data reporting formats and table contents (Kohn et al., 2024). Given the continuing relevance of fission-track data, we recommend building toward a global infrastructure and culture enabling and encouraging data transparency through the formation of online digital image repositories (such as geochron@home (Vermeesch et al., 2023)), which can accommodate fission-track image data. Furthermore, proper analyst training and reconsideration of laboratory routines for image acquisition are needed. It has been over 60 years since the beginning of the fission-track dating method (Price and Walker, 1962), and no clear guidelines have been formulated on "musts" and "cans" in fission-track practice. While existing fission-track laboratories develop their own preferences and routines, new laboratories often represent branching points, which can be a source of necessary and beneficial innovation but also undocumented and undesireable divergence. A global community repository housing guidelines for best-practice fission track analyses and fission track training modules are needed, as well as reference libraries of interpreted image sets. Adoption of an open science culture will ultimately benefit every fission-track laboratory and increase data quality.

Code and data availability. The image data used for this study, and the analysis results, have been uploaded to geochron@home (Vermeesch 2024). The analyst numbers in Table S2 are replaced with random letters in geochron@home to preserve anonymity.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

Author contributions. Conceptualisation: MTT, LC, RAK, AJWG. Data curation: MTT, LC. Formal analysis: MTT, LC. Software: LC, AJWG. Investigation: MTT, LC. Methodology: MTT, LC, RAK, AJWG. Resources - MTT, LC, RAK, AJWG. Writing – original draft preparation: MTT, LC, RAK, AJWG. Writing – review & editing: MTT, LC, RAK, AJWG.

Competing interests. The authors have no competing interests to declare

Acknowledgements. We appreciate all of the participants in this study for their participation and critical discussions. This research was supported by the Geology Foundation of the Jackson School of Geosciences and Fundamental Research Funds in the Institute of Geology, China Earthquake Administration (IGCEA2229), National Science Foundation of China (4231101318) and Project 3.51 of the AuScope Program within the Australian National Collaborative Research Infrastructure Strategy.

References

Ansberque, C., Chew, D. M., & Drost, K.: Apatite fission-track dating by LA-Q-ICP-MS imaging. Chemical Geology, 560, 119977, https://doi.org/10.1016/j.chemgeo.2020.119977, 2021.

Aslanian, C., Jonckheere, R., Wauschkuhn, B., & Ratschbacher, L.; Short communication: Experimental factors affecting fission-track counts in apatite. Geochronology, 4(1), 109-119, https://doi.org/10.5194/gchron-4-109-2022, 2022.

Barbarand, J., Hurford, T., & Carter, A. : Compositional and structural control of fission-track annealing in apatite. Chemical Geology, 198(1-2), 107-137, https://doi.org/10.1016/S0009-2541(02)00424-2, 2003.

Barnes, J.B., Ehlers, T.A., Insel, N., McQuarrie, N. & Poulsen, C.J. : Linking orography, climate, and exhumation across the central Andes. Geology, 40(12), pp.1135-1138. https://doi.org/10.1130/G33229.1, 2012.

Balestrieri, M.L., Bigazzi, G. & Ghezzo, C. : The transantarctic mountains: a natural laboratory for apatite fission-track analysis. Results from Italian Antarctic expeditions. Radiation measurements, 31(1-6), pp.621-626. https://doi.org/10.1016/S1350-4487(99)00154-7, 1999.

Bernet, M., Wootton, L., Valla, P., Lavergne, S., Boucher, F., Barun, J., Husson, L., Robert, X., & Balvay, M.: Tracing Western Alps mountain building and high mountain biodiversity since the the early-mid Miocene. 18th International Conference of Thermochronology, Riva del Garda, Italy, 2023.

Boone, S.C., Faux N., Nattala, U., Jiang, C., Church, T., Chung, L., Mcmillan, M., Jones, S., Jiang, H., Liu, D., Ehinger, K., Drummond, T., Kohn, B., & Gleadow, A.J.W. : Towards Fully Automated Digital Fission-Track Analysis Through Artificial Intelligence. 18th International Conference of Thermochronology, Riva del Garda, Italy, 2023.

Boone, S.C., Kohlmann, F., Noble, W., Theile, M., Beucher, R., Kohn, B., Glorie, S., Danišík, M., Zhou, R., McMillan, M. & Nixon, A. : A geospatial platform for the tectonic interpretation of low-temperature thermochronology Big Data. Scientific Reports, 13(1), p.8581, https://doi.org/10.1111/ggr.12419, 2023.

Boone, S. C., Dalton, H., Prent, A., Kohlmann, F., Theile, M., Gréau, Y., Florin, G., Noble, W., Hodgekiss, S. A., &Ware, B. : AusGeochem: An Open Platform for Geochemical Data Preservation, Dissemination and Synthesis. Geostandards and Geoanalytical Research 46 (2), https://doi.org/10.1111/ggr.12419, 2022.

Carlson, W.D., Donelick, R.A. & Ketcham, R.A.: Variability of apatite fission-track annealing kinetics: I. Experimental results. American Mineralogist, 84(9), pp.1213-1223. https://doi.org/10.2138/am-1999-0901, 1999.

Cogné, N. & Gallagher, K.: Some comments on the effect of uranium zonation on fission track dating by LA-ICP-MS. Chemical Geology, 573, p.120226. https://doi.org/10.1016/j.chemgeo.2021.120226, 2021.

Deng, J., Wang, C., Bagas, L., Carranza, E. J. M., & Lu, Y. : Cretaceous–Cenozoic tectonic history of the Jiaojia Fault and gold mineralization in the Jiaodong Peninsula, China: constraints from zircon U–Pb, illite K–Ar, and apatite fission track thermochronometry: Mineralium Deposita, 50, 987-1006, https://doi.org/10.1007/s00126-015-0584-1, 2015.

Donelick, R. A., O'Sullivan, P. B., & Ketcham, R. A. : Apatite fission-track analysis. Reviews in Mineralogy and Geochemistry, 58(1), 49-94. https://doi.org/10.2138/rmg.2005.58.3, 2005.

Dumitru, T.A., Hill, K.C., Coyle, D.A., Duddy, I.R., Foster, D.A., Gleadow, A.J.W., Green, P.F., Kohn, B.P., Laslett, G.M. & O'Sullivan, A.J. : Fission track thermochronology: application to continental rifting of south-eastern Australia. The APPEA Journal, 31(1), pp.131-142. https://doi.org/10.1071/AJ90011, 1991.

Fitzgerald, P.G. & Goodge, J.W. : Exhumation and tectonic history of inaccessible subglacial interior East Antarctica from thermochronology on glacial erratics. Nat Commun 13, 6217. https://doi.org/10.1038/s41467-022-33791-y, 2022.

Gallen, S.F., Seymour, N.M., Glotzbach, C. Stockli, D.F., & O'Sullivan, P. : Calabrian forearc uplift paced by slab–mantle interactions during subduction retreat. Nat. Geosci. https://doi.org/10.1038/s41561-023-01185-4, 2023.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

Gleadow, A. J. W., & Lovering, J. F. : Geometry factor for external detectors in fission track dating. Nuclear Track Detection, 1(2), 99-106. https://doi.org/10.1016/0145-224X(77)90003-5, 1977.

Gleadow, A.J.W., Duddy, I.R., Green, P.F., & Lovering, J.F. : Confined fission track lengths in apatite: a diagnostic tool for thermal history analysis. Contributions to Mineralogy and Petrology, 94, 405-415. https://doi.org/10.1007/BF00376334, 1986.

Gleadow, A.J.W., Gleadow, S.J., Belton D.X., Kohn, B.P., Krochmal M.S. & Brown, R.W. : Coincidence mapping - a key strategy for the automatic counting of fission tracks in natural minerals . Geological Society of London Special Publication 324, 25-36. https://doi.org/10.1144/SP324.2, 2009.

Gleadow, A.J.W., Kohn, B., & Seiler, C. : The Future of Fission-Track Thermochronology. In: Malusà, M., Fitzgerald, P. (eds) Fission-Track Thermochronology and its Application to Geology. Springer Textbooks in Earth Sciences, Geography and Environment. Springer, Cham. https://doi.org/10.1007/978-3-319-89421-8_4, 2019.

Green, P. F. (1981), A new look at statistics in fission-track dating: Nuclear tracks, 5(1-2), 77-86. https://doi.org/10.1016/0191-278X(81)90029-9, 1981.

Gülyüz, N., Gülyüz, E., Karaoglan, F., & Kuscu, I. : Low temperature thermochronology reveals tilting of crystalline bodies, Halilaga porphyry Cu-Au deposit, NW Anatolia: Implications for exploration of porphyry copper deposits and interpretation of low-temperature thermochronology data for regional tectonics. Ore Geology Reviews, 166, 105958. https://doi.org/10.1016/j.oregeorev.2024.105958, 2024.

Hasebe, N., Barbarand, J., Jarvis, K., Carter, A. & Hurford, A.J. : Apatite fission-track chronometry using laser ablation ICP-MS: Chemical Geology, 207(3-4), pp.135-145. https://doi.org/10.1016/j.chemgeo.2004.01.007, 2004.

Hurford, A.J. & Green, P.F., : The zeta age calibration of fission-track dating, Chemical Geology, 41, pp.285-317. https://doi.org/10.1016/S0009-2541(83)80026-6, 1983.

Karaoglan F., Karatas, B., Ozdemir, Y., Gulyuz, E., Vassiliev, O., Selbesoglu, M. O. & Gildir, S. : The geo/thermochronology of Dismal Island (Marguerite Bay, Antarctic Peninsula), Turkish Journal of Earth Sciences, 32, 8, 4. https://doi.org/10.55730/1300-0985.1887, 2023.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

Ketcham, R. A., Donelick, R. A., Balestrieri, M. L., & Zattin, M. : Reproducibility of apatite fission-track length data and thermal history reconstruction: Earth and Planetary Science Letters, 284(3), 504-515. https://doi.org/https://doi.org/10.1016/j.epsl.2009.05.015, 2009.

Ketcham, R. A., Carter, A., & Hurford, A. J., : Inter-laboratory comparison of fission track confined length and etch figure measurements in apatite, American Mineralogist, 100(7), 1452-1468. http://dx.doi.org/10.2138/am-2015-5167, 2015.

Ketcham, R. A., van der Beek, P., Barbarand, J., Bernet, M., & Gautheron, C., : Reproducibility of thermal history reconstruction from apatite fission-track and (U- Th)/He data: Geochemistry, Geophysics, Geosystems, 19(8), 2411-2436. https://doi.org/10.1029/2018GC007555, 2018.

Kohn, B.P., Pillans, B. & McGlone, M.S., : Zircon fission track age for middle Pleistocene Rangitawa Tephra, New Zealand: stratigraphic and paleoclimatic significance. Palaeogeography, palaeoclimatology, palaeoecology, 95 (1-2), pp.73-94. https://doi.org/10.1016/0031-0182(92)90166-3, 1992.

Kohn, B. P., Ketcham, R. A., Vermeesch, P., Boone, S. C., Hasebe, N., Chew, D., Bernet, M., Chung, L., Danisik, M., Gleadow, A.J.W., & Sobel, E. R. : Interpreting and reporting fission-track chronological data, Geological Society of America Bulletin, pp. 30-. https://doi.org/10.1130/B37245.1, 2024.

Li, R., Xu, Z., Su, C., & Yang, R., : Automatic identification of semi-tracks on apatite and mica using a deep learning method, Computers & Geosciences, 162, 105081. https://doi.org/10.1016/j.cageo.2022.105081, 2022.

Lemot, F., Valla, P.G., van der Beek, P., Jagercikova, M., Niedermann, S., Carcaillet, J., Sobel, E.R., Andò, S., Garzanti, E., Robert, X. & Balvay, M. : Miocene cave sediments record topographic, erosional and drainage development in the Western European Alps. Earth and Planetary Science Letters, 621, p.118344. https://doi.org/10.1016/j.epsl.2023.118344, 2023.

Miller, D. S., Duddy, I. R., Green, P. F., Hurford, A. J., & Naeser, C. W. : Results of interlaboratory comparison of fission-track age standards: Fission-track workshop—1984, Nuclear Tracks and Radiation Measurements, 10(3), 383–391. https://doi.org/10.1016/0735-245X(85)90127-9, 1985.

Miller, D.S., Eby, N., McCorkell, R., Rosenberg, P.E. & Suzuki, M. : Results of interlaboratory comparison of fission track ages for the 1988 fission track workshop. International Journal of Radiation Applications and Instrumentation. Part D. Nuclear Tracks and Radiation Measurements, 17(3), pp.237-245. https://doi.org/10.1016/1359-0189(90)90041-U, 1990.

Miller, D.S., Crowley, K.D., Dokka, R.K., Galbraith, R.F., Kowallis, B.J., & Naeser, C.W. : Results of interlaboratory comparison of fission track ages for 1992 Fission Track Workshop, Nuclear Tracks, 21(4), 565-57. https://doi.org/10.1016/1359-0189(93)90179-H., 1993.

Nachtergaele, S., and De Grave, J. : AI-Track-tive: open-source software for automated recognition and counting of surface semi-tracks using computer vision (artificial intelligence), Geochronology, 3(1), 383-394. https://doi.org/10.5194/gchron-3-383-2021, 2021.

Naeser, C.W., Zimmermann, R.A. & Cebula, G.T. : Fission-track dating of apatite and zircon: an interlaboratory comparison, Nuclear Tracks, 5(1-2), pp.65-72. https://doi.org/10.1016/0191-278X(81)90027-5, 1981.

Price, P.B., & Walker, R.M. : Chemical etching of charged particle tracks in solids, Journal of Applied Physics, 33, 3407-3412. https://doi.org/10.1063/1.1702421, 1962.

Reiners, P.W., & Shuster, D.L. : Thermochronology and landscape evolution, Physics Today; 62 (9): 31–36. https://doi.org/10.1063/1.3226750, 2009.

Qiu, K.F., Deng, J., Sai, S.X., Yu, H.C., Tamer, M. T., Ding, Z.J., Yu, X.F., & Goldfarb, R.J. : Low-temperature thermochronology for defining the tectonic controls on heterogeneous gold endowment across the Jiaodong Peninsula, eastern China, Tectonics, 42(1), e2022TC007669. https://doi.org/10.1029/2022TC007669, 2023.

Qiu, N. & Liu, S. : Uplift and denudation in the continental area of China linked to climatic effects: Evidence from apatite and zircon fission track data, Scientific Reports, 8(1), p.9546. https://doi.org/10.1038/s41598-018-27801-7, 2018.

Ren, Z., Li, S., Xiao, P., Yang, X. & Wang, H. : Artificial intelligent identification of apatite fission tracks based on machine learning. Machine Learning: Science and Technology, 4(4), p.045039. https://doi.org/10.1088/2632-2153/ad0e17, 2023.

Sobel, E.R., & Seward, D. : Influence of etching conditions on apatite fission-track etch pit diameter: Chemical Geology 271(1-2), 59-69. https://doi.org/10.1016/j.chemgeo.2009.12.012, 2010.

Tagami, T. & O'Sullivan, P.B. : Fundamentals of fission-track thermochronology: Reviews in mineralogy and geochemistry, 58(1), pp.19-47. https://doi.org/10.2138/rmg.2005.58.2, 2005.

https://doi.org/10.5194/gchron-2024-26
Preprint. Discussion started: 9 October 2024

Tamer, M. T., Chung, L., Ketcham, R. A., & Gleadow, A. J., : Analyst and etching protocol effects on the reproducibility of apatite confined fission-track length measurement, and ambient-temperature annealing at decadal timescales: American Mineralogist: Journal of Earth and Planetary Materials, 104(10), 1421-1435. https://doi.org/10.2138/am-2019-7046, 2019.

Tamer, M. T., Chung, L., Ketcham, R. A., & Gleadow, A.J.W. : Inter-analyst comparison and reproducibility of apatite fission track analysis. 17th International Conference of Thermochronology, Santa Fe, USA, 2021. DOI: https://doi.org/10.1002/essoar.10507907.2, 2021.

Tamer, M. T., & Ketcham, R. A. : How many vs which: On track selection criteria for apatite fission track analysis: Chemical Geology, 634, 121584. https://doi.org/10.1016/j.chemgeo.2023.121584, 2023.

Torres, V., Hooghiemstra, H., Lourens, L. & Tzedakis, P.C. : Astronomical tuning of long pollen records reveals the dynamic history of montane biomes and lake levels in the tropical high Andes during the Quaternary. Quaternary Science Reviews, 63, pp.59-72. https://doi.org/10.1016/j.quascirev.2012.11.004, 2013.

Tversky, A., & Kahneman, D. : Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. Science, 185(4157), 1124-1131. https://doi.org/10.1126/science.185.4157.1124, 1974.

Vermeesch, P. : pvermees/GaHa: v.1 (v.1). Zenodo. https://doi.org/10.5281/zenodo.13777917, 2024.

Vermeesch, P. :How many grains are needed for a provenance study? Earth and Planetary Science Letters, 224(3-4), pp.441-451. https://doi.org/10.1016/j.epsl.2004.05.037, 2004.

Vermeesch, P. : Statistics for LA-ICP-MS based fission track dating. Chemical Geology, 456, pp.19-27. https://doi.org/10.1016/j.chemgeo.2017.03.002, 2017.

Vermeesch, P., Band, T., Duong, E., & Carter, A. : Crowd sourced fission track dating with geochron@home. 18th International Conference on Thermochronology, Riva del Garda, Italy. Abstract Volume, 190, 2023.

Yu, J., Zheng, D., Pang, J., Li, C., Wang, Y., Wang, Y., Hao, Y. & Zhang, P. ; Cenozoic mountain building in eastern China and its correlation with reorganization of the Asian climate regime: Geology, 50(7), pp.859-863. https://doi.org/10.1130/G49917.1, 2022.