Response to Reviewer #2's comments on manuscript egusphere-2023-1293

"An expanded workflow for detrital rutile provenance studies: An application from the Neotethys Orogen in Anatolia"

Megan A. Mueller, Alexis Licht, Andreas Möller, Cailey B. Condit, Julie C. Fosdick, Faruk Ocakoğlu, and Clay Campbell

November 15, 2023

We appreciate the thoughtful reviews of the 3 referees. First, we summarize the broad themes from the three reviews before responding in detail to Reviewer #2 below. The reviews critiqued (1) the novelty of the study, (2) the number of U-Pb analyses discarded during data reduction, (3) the potential bias of discarding data and the validity of interpreting discordant U-Pb analyses, and (4) the apparent lack of novelty and complexity in geochemical data interpretations. Regarding these points, the goal of the manuscript was to provide a transparent workflow for detrital rutile geochronology using new data from sedimentary basins in Anatolia. We acknowledge that it was confusing or misleading to present this work as a revised workflow. We intended to emphasize that we found a paucity of methods papers that provide a straightforward approach. Although we did not claim to present new workflows with the geochemical data, this was perhaps unintentionally implied with the title. We have presented as much information as we could squeeze out of our particular geochemical dataset. The revised manuscript will scale back statements on new workflows and instead refocus the title and introduction on suture zone settings and Anatolian geology while maintaining the broad overview of detrital rutile provenance and the details of our methods.

We also acknowledge that it is surprising to see the number of data discarded during data reduction. However, we contend that this is a common practice with detrital U-Pb geochronology in common Pb-bearing minerals. We are not the first to discard a significant number of analyses; we have found this practice in many published detrital rutile datasets, although it is not discussed much in the literature. We further expand on this point in our detailed reply and will add this context to the revised manuscript. This manuscript gives precedence for papers to be transparent in data reporting—including the number of grains analyzed and criteria for rejection—as well as examination of the full dataset in Tera-Wasserburg diagrams. Our manuscript provides an opportunity to show the consequence and potential value of the full dataset, which is relevant to others working with this type of data. While the use of common Pb-bearing minerals is common in some labs and geographic settings, the application of these tools is still far behind detrital zircon geochronology, for which there is a well-established global framework. We have encountered many detrital geochronology users who want to add detrital rutile to their toolkits but are still uncertain in how to collect and interpret these data. Here, we present a complicated detrital rutile U-Pb dataset that can serve as an example for how to treat and interpret complex, yet potentially meaningful, discordant data.

Reviewer #2 provided a critical review of our manuscript with comments that centered around (1) the U-Pb data reduction and handling of the reference materials, (2) rejection of U-Pb analyses, and (3) the validity of provenance interpretations on discordant data. Reviewer #2 included numerous in-text comments that are summarized in their review text. We thank Reviewer #2 for the thorough feedback. Reviewer #2's comments will strengthen the clarity of the revised manuscript, yet we generally disagree with Reviewer #2's characterization of points (2) and (3) and provide a rebuttal.

Reviewer #2 provided comments in the review and line-by-line comments in the text. We address the comments below by theme. Reviewer #2's comments are included below in black text. Our response is in purple text and the specific changes we will implement are highlighted in ***bold, italic purple text***.

1. **Comments regarding U-Pb analytical protocol, data reduction, and data handling of reference materials**
   - the data handling of the reference materials (unreported scatter in the measured and corrected isotopic ratios of the reference materials, inconsistency between uncertainties before and after the correction, lack of representation of the data as Concordia diagrams)
   - Line 145: The Th/U ratio of NIST SRM 612 is expected to be close to 1 (e.g. see Guillong et al 2003 JAAS "U/Th. The comparison of Th and U has the advantage that isotopic abundance, concentrations and first ionisation potentials are similar and a ratio determination should therefore give approx. 1."). This means that the joint tuning of the MS and laser system was not optimal.
   - Line 175: For R10, the uncert. of the uncorrected 6/38 age ("2SE prop abs" from Iolite in table "Uncorrected data output", supplem S2) ranges from 2.2 to 4.1%, for secondary ref mat LI04-08 the uncert. ranges from 2.3 to 4.3%. In contrast, the uncert. of the 6/38 208-corrected dates is 0.4-1.2% and 0.5-1.1%, respectively (208 corrected Table, supplem S2). The uncertainty of the corrected dates appears to be substantially smaller than the uncorrected dates. See also further comments in the U-Pb method section.
   - Line 250: MAJOR COMMENT.
     - The uncorrected R10 data (suppplem table S2, final 7/6 and 6/38 ratios and 2SE prop uncert) are scattered along Concordia and some datapoints are largely discordant. The authors should clarify if any correction has been applied to the primary ref material before normalization of the samples and secondary ref mat. ratios; if not, why.
     - The INDIVIDUAL undertainty of the final UNCORRECTED R10 ratios (table S2, Iolite output: 2SE propagated uncert) is on average 2.3% for the 7/6 ratio and 3.6% for the 38/6 ratio. In contrast, the reported uncertainty of the CORRECTED R10 data is an avg of 0.7% for the 38/6 ratio or 6/38 age (208 correction, 207 correction, 207 correction with 208 as initial age, 207 then 208 correction, table S2). The average individual uncert for the 7/6 age after the correction is 2.1% (all correction approaches, with ca 50% of the individual uncertainties smaller than the original uncertainty and the rest larger up to 0.5%.) This raises questions about the uncertainty propagation protocol associated with the correction(s).
   - Line 250: MAJOR COMMENT the U-Pb data of the reference materials (and the samples) must be plotted on a Concordia diagram. Additionally, presenting the results as weighted averages values only, without reporting the corresponding MSWD, masks any potential scatter in the data (Fig. A2 and Table "Std reproducibility" in supplem file S2). Additionally, it is unclear how the R10 uncorrected wtd average age reported in Table "Standard reproducibility" (supplem S2) has been calculated., as it differs from the wtd avg calculated (with IsoplotR) using the data reported in Table "uncorrected data output". E.g. the final uncorrected 7/6 age wtd avg (calculated using the 2SE propagated uncert.) is 1088.3 +/- 2.8 Ma, n=215/215, MSWD =0.41, vs the 1091.7 +/-0.5 Ma age reported in in Table "Standard reproducibility"; the final uncorrected 6/38 age wtd avg is 1092.21 + /- 2.32 Ma, n=215/215, MSWD 0.35. This should be clarified also for the secondary reference materials, and averages re-calculated form the presented data when they do not match.
   - Line 250: MAJOR COMMENT. The plot in figure A2 masks significant scatter in the corrected data. E.g. the "208 corrected ages" for R10 range from ca 1060 to 1120 Ma (largely not overlapping at the 2s level), with an average of ca. 1091 Ma however with a MSWD of 5.1, n=210/214 ; this would be apparent if the data had been plotted on a Concordia diagram or weighted avg or linearized probabilty plot. The "207c with 208c as initial age" corrected avg ages

for R10 are: 7/6 corrected age = 1138.6+/- 1.6, n=214, MSWD 8.7; 6/38 corrected age = 1093.7 +/-0.5, 212/214, MSWD 7.6. Such a scatter must be reported and discussed.

- Line 255: MAJOR COMMENT The need for rejection of more than 50% of the analyses due anomalous spiky patterns possibly points to an underlying issue with the analytical protocol, data acquisition or the samples themselves, or their preparation. Were similar spiky patterns observed for the reference materials? Had masses 204-202 been monitored, the occurrence of inclusions or unevenly distributed common Pb would have been identified by simply observing the corresponding signal intensities/pattern in the time resolved data.
- Line 255: On which basis such a 20% 7/6 uncertainty threshold was chosen? Considering that the rutile populations of this study are dominated by 100-300 Ma rutiles (Fig.5), a threshold should have been set on the 6/38 ratio uncertainty instead. (Cf Govin et al 2018 Geology Table DR7; LA U-Pb detrital rutile data similarly collected with a single collector SF-ICPMS)
- Table A2: The choice of 3 passes (instead of 1) for U-Pb analysis is quite unusual. It causes that each set of 3 passes (one "pass" being a sweep across the mass range of interest, 206-238 in this case) is averaged by the Element software into one of the 100 runs, and although a total of 300 sweeps will be measured, the data output will consist of 100 datapoints, not 300. This has implications in terms of counting statistics nd final uncertainty of the ratios. Why this choice, as opposed to measuring the trace elements by means of 120 runs x 1 pass?

The analytical protocol is sound. The same or very similar protocols and data acquisition settings have been in multiple previous studies carried out in the same lab (e.g., Rösel et al., 2019). (1) Anomalous signal patterns were not observed in the reference materials. The rejection of data appears to be a characteristic of this detrital sample set, similar to several other published detrital datasets (e.g., Govin et al. 2018 discarded nearly 40% of data from a much smaller dataset, n=147). See our reply to point #2 below. (2) Regarding the analytical set-up, the tuning was optimized for high sensitivity and low oxide production, while accepting a less than optimal element fractionation. This was intended to allow analyses of low U and Pb rutile. Since the elemental fractionation is corrected during the data reduction process, this does not invalidate or adversely affect the ability to use data, whereas low sensitivity would exclude even more low U and Pb rutile from being able to be used. ***This sentence will be updated in the revised manuscript (line 145).*** (3) The choice of multiple passes is not unusual and has been standard practice for U-Pb dating in multiple labs since before 2007 (e.g., Frei & Gerdes 2009). We are aware of the effects. We hold that the chosen method provides adequate data even for rutile with very low U and Pb concentrations, and fast enough mass scans to identify inclusions. The obvious difference between the U-Th-Pb and trace elements method is that each scan through the masses for the trace element method is much longer than for Pb-Th and U, which does not require settling times of the magnet. A single pass with adequate counting times was therefore chosen for the trace element method and yields a similar number of data points.

Reviewer #2 raises an important point about the reference materials. We address this point by first discussing the primary standard R10, then by clarifying the analytical protocol and discussing the secondary standard results. There is no excessive scatter or discordance in the calibration standard R10. The uncorrected results are displayed in Wetherill space in Figure R1 (using propagated uncertainties). For all analyses (n=210), the concordia age is 1091.7 ± 1.7 Ma with a MSWD of 0.66 and MSWD for concordance and equivalence of 1.2. A common Pb correction was not applied to R10 as Luvizotto et al. (2009) writes, *"LA-ICP-MS data show that the non-radiogenic Pb concentrations (measured as 208 Pb) are very low (average of 0.08 ppm), meaning that common Pb can be neglected for [R10]."* ***The revised manuscript will include concordia diagrams of the reference materials and will clarify that the results are uncorrected for common Pb.***

The U-Pb data was collected in 4 analytical sessions. The analytical protocol was modified from session to session to optimize for the analysis of low U and Pb unknowns and high U and Pb reference materials. In the first two analytical sessions, 21RtF and 21RtG, Pb and Th isotopes were measured with the secondary electron multiplier (SEM) operating in counting detection mode. The secondary standards Wodgina and Kragerø did not perform well during those analytical sessions: there is an increase in ^{206}Pb and $^{206}Pb/^{238}U$ ratio and sharp decrease in $^{207}Pb/^{206}Pb$ at the beginning of the analysis (Figure R2). During those sessions, Wodgina reached ca. 4 million ^{206}Pb counts per second, which exceeds the limit of linear behavior in counting detection mode. Kragerø has the second highest counts with a maximum of around 3 million ^{206}Pb counts per second, and the other standards remained within the linear behavior range. We conclude that the poor performance of Wodgina and Kragerø on those days was due to this counting mode issue and could explain the extreme increase in $^{206}Pb/^{238}U$ ratio at the start of the analysis (Figure R2). This issue was corrected in the following two sessions, 21RtA and 21RtB, in which the SEM operated in "both" detection mode for Pb and Th isotopes, and Wodgina performed well (Figure R2). Figure R3 displays the Wodgina results from all analytical sessions in Wetherill space: data from the first two sessions are discordant and were rejected, data from the final two sessions are acceptable with a concordia age of 2819.0 ± 3.3 Ma (2s) (less than -1% offset from TIMS age). Note that only one analysis of the unknown came close to the high count rates observed in the Archean high-U rutile Wodgina, so we argue that the issue did not affect the unknowns. We compare the Wodgina results to that of secondary standard 9826J because 9826J was not affected by the SEM detection mode issue (ca. < 750,000 CPS ^{206}Pb). We compare the performance of 9826J across analytical sessions (Figure R4). In sessions 21RtF and 21RtG, 9826J has a concordia age of 378.1 ± 1.9 Ma (2s) with an MSWD for concordance and equivalence of 2.2, indicating excess scatter. This age is slightly younger than the TIMS age of 381.9 ± 1.1 Ma (Kylander-Clark, 2008), but only offset by about -1%, the long-term reproducibility range of LA-ICP-MS data. From session 21RtA, 9826J has a concordia age of 385.9 ± 3.8 Ma (2s), is within the uncertainty of the TIMS age, and has an MSWD for concordance plus equivalence of 0.88, indicating slight underdispersion. The concordia ages are satisfactory for all analytical sessions, but there is excess scatter in the first two analytical sessions which could be natural.

Next, we investigated whether excess scatter in the secondary standards was a result of the drift correction curve. In Iolite, we compared the effects of weighted linear fit, SplineSmooth5 and SplineSmooth9 curves for drift correction. The choice of spline had little effect on the final age of the reference materials but impacted the MSWD. Scatter in the reference materials is not explained by the drift correction curve. We prefer the weighted linear fit as this model reproduces the secondary standard ages and brings the MSWD closest to 1 for each standard. ***The revised manuscript will include the results that were reduced using the weighted linear fit drift correction. All figures and tables in the main text and supplement will be updated. We do not anticipate that this change will produce significant changes to the results or interpretations.***

We maintain that the poor performance of Wodgina and Kragerø in the first analytical sessions was a limitation of the SEM detection mode and does not indicate that the entire analytical sessions should be discarded: (1) the R10 and 9826J results were acceptable for analytical sessions 21RtF and 21RtG, (2) the signal patterns and results were acceptable for Wodgina in analytical sessions 21RtA and 21RtB, (3) in the unknowns, the Pb and Th isotopes signal intensities were within the linear range of the counting mode as evidenced by the data on the other reference materials. ***In the revised manuscript, the Wodgina and Kragerø analyses from sessions 21RtF and 21RtG will be excluded. The text and supporting information will be updated to explain the changes in analytical protocol across sessions. Concordia diagrams of reference materials will be added. Here we have shown concordia diagrams of several of the standards, and diagrams of all standards will be included in the revised manuscript. We will discuss in the text and/or supplement the day-by-day performance of the standards.***

Regarding the $^{207}Pb/^{206}Pb$ ratio threshold, the 20% rejection criteria is following previous studies carried out in the same lab (e.g., Lippert, 2014). In the study suggested by Reviewer #2, Govin et al. (2018), the rejection criterion is based on the uncertainty on the corrected age. The threshold changes based on the corrected age; for example, analyses with uncertainties > 10 % are discarded for ^{207}Pb-

corrected ages > 100 Ma. This age-dependent threshold is similar to approaches used in detrital zircon and can induce bias (Malusà et al., 2013). This rejection criterion depends on how the uncertainty is calculated, including whether uncertainties on the Stacey and Kramers (1975) values are accounted for, for example. For our dataset, Govin et al.'s age uncertainty filter excludes about 20% of the analyses that are included by the $^{207}Pb/^{206}Pb$ filter. However, Govin et al. (2018) do not exclude any data based on discordance, which, for our dataset, would include more analyses than the protocol we used. ***For the revised manuscript, we will re-examine whether to discard analyses based on the $^{207}Pb/^{206}Pb$ ratio, $^{206}Pb/^{238}U$ ratio and/or the uncertainty on the corrected age. If warranted, we will include a brief discussion in the revised manuscript.***

We calculated individual uncertainty in the following way, after Odlum et al. (2019) (see lines 177-178). The uncertainty in percent is calculated using the initial $^{206}Pb/^{238}U$ ratio and internal uncertainty. The uncertainty on the final age is calculated from the percent uncertainty on the initial ratio. For example, if the initial $^{206}Pb/^{238}U$ ratio has 2% uncertainty at 2 sigma and the corrected age is 200 Ma, then the corrected age uncertainty is ±4 Ma (2s). We used the internal uncertainty rather than propagated uncertainty, which is likely why the reviewer calculated larger uncertainties. ***The revised manuscript will use the propagated uncertainty. The data tables, figures and text will be updated. The revised manuscript will include Tera-Wasserburg diagrams of the reference materials using the 2s propagated uncertainty. We will display the concordia age and MSWD. We will eliminate Figure A2 and Dataset S2 tab 'Standard Reproducibility' in favor of the new Tera-Wasserburg plots.*** Regarding the uncertainty on the corrected and uncorrected ages of the reference materials, the discrepancy described by Reviewer #2 is likely a result of comparing the internal and propagated uncertainties. ***This should be resolved in the revised manuscript with the use of the propagated uncertainty.***

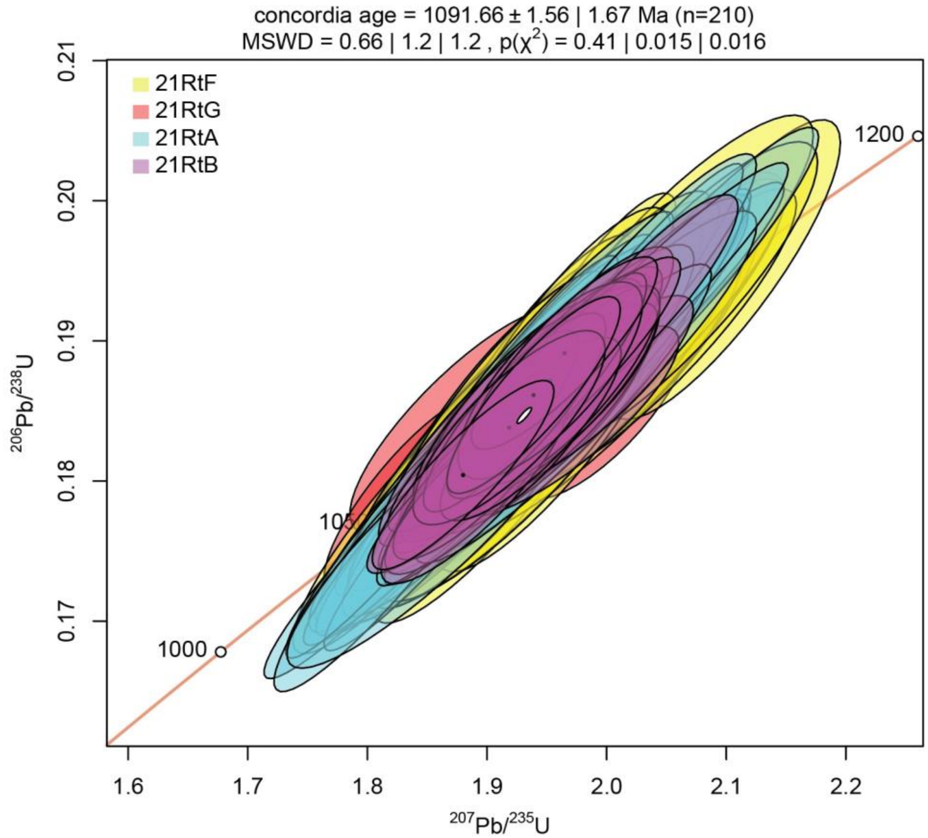
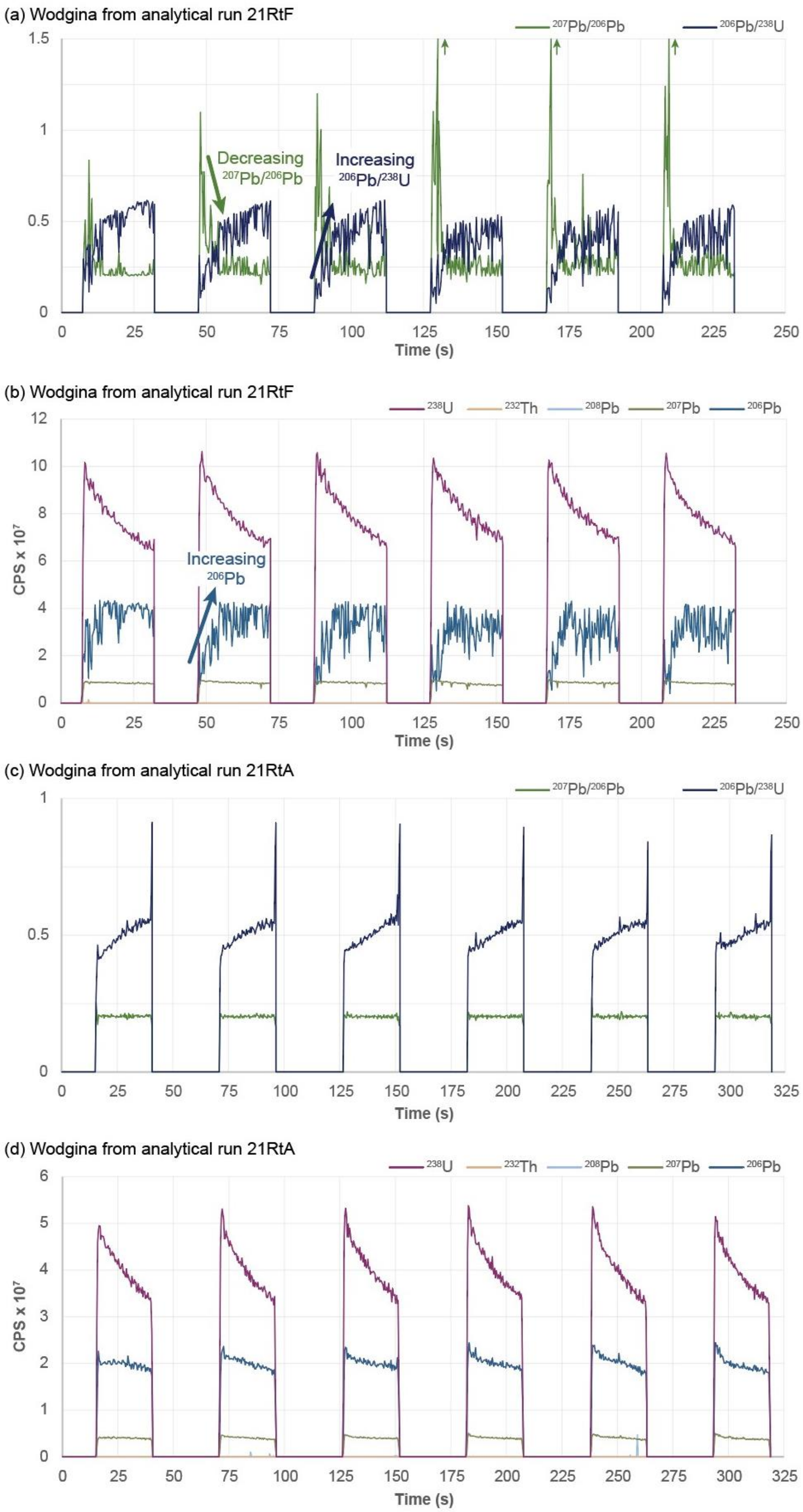


Figure R1. All R10 results are displayed in Wetherill space. The ellipses are colored by the analytical session. For all analytical sessions, the concordia age is 1091.7 ± 1.6 Ma with an MSWD of 0.66 and MSWD for concordance and equivalence of 1.2. Uncertainties and ellipses are propagated 2s absolute. The calculated concordia age is shown as a white shaded ellipse. A common Pb correction was not applied. Figure made using IsoplotR (Vermeesch, 2018).

Figure R2 (next page). Comparison of secondary standard Wodgina from analytical sessions 1 (21RtF) and 3 (21RtA), representing the switch from operating the secondary electron multiplier (SEM) detection mode in counting to both. The first 6 Wodgina analyses from each session are displayed. Note the increase in ^{206}Pb and $^{206}Pb/^{238}U$ ratio and sharp decrease in $^{207}Pb/^{206}Pb$ downhole during session 21RtF (counting mode).


(a) Wodgina from analytical run 21RtF
207Pb/206Pb
206Pb/238U
Decreasing 207Pb/206Pb
Increasing 206Pb/238U
Time (s)
(b) Wodgina from analytical run 21RtF
238U
232Th
208Pb
207Pb
206Pb
CPS x 10^7
Increasing 206Pb
Time (s)
(c) Wodgina from analytical run 21RtA
207Pb/206Pb
206Pb/238U
Time (s)
(d) Wodgina from analytical run 21RtA
238U
232Th
208Pb
207Pb
206Pb
CPS x 10^7
Time (s)

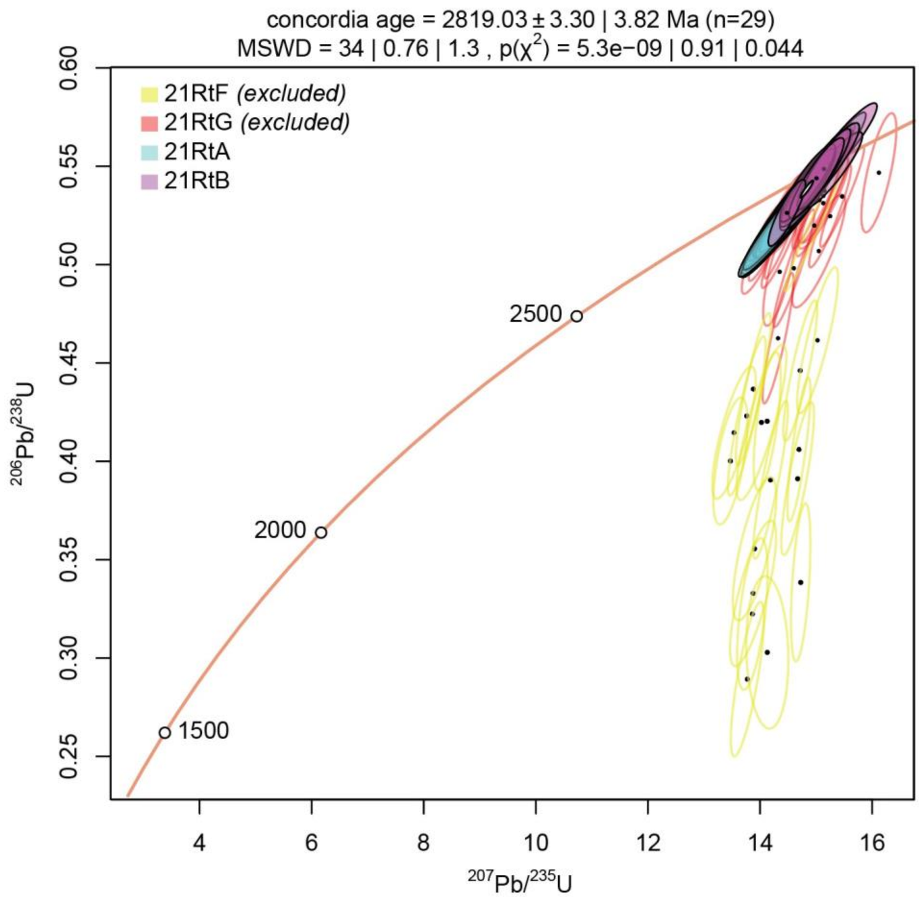
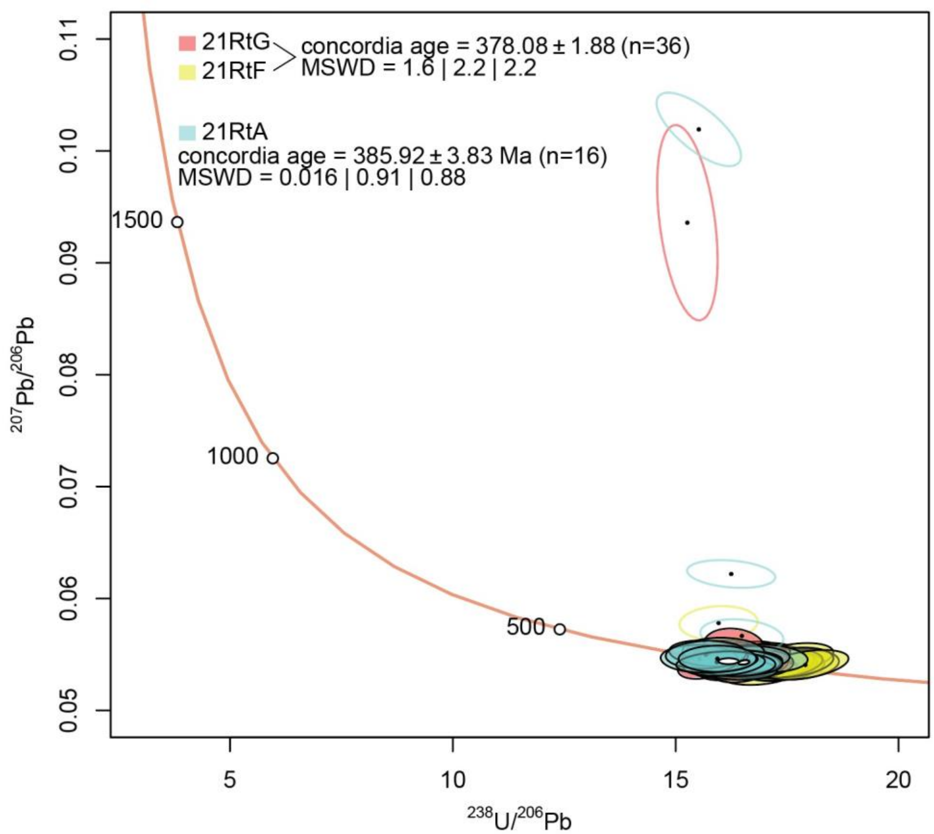


Figure R3. All Wodgina results displayed in Wetherill space. The ellipses are colored by the analytical session. Wodgina from sessions 21RtF (yellow) and 21RtG (red) are excluded from the concordia age calculation. For analytical sessions 21RtA and 21RtB, the concordia age is 2819.0 ± 3.3 Ma with an MSWD for concordance and equivalence of 1.3. Uncertainties and ellipses are propagated 2s absolute. Ellipses that are not shaded are excluded from concordia age calculations (i.e., analyses from sessions 21RtF and 21RtG). The calculated concordia age is shown as a white shaded ellipse. A common Pb correction was not applied. Figure made using IsoplotR (Vermeesch, 2018).

Figure R4 (next page). Results of secondary standard 9826J from analytical sessions 21RtF, 21RtG and 21RtA. From sessions 21RtF and 21RtG, 9826J has a concordia age of 378.1 ± 1.9 Ma (2s) with an MSWD for concordance plus equivalence of 2.2, indicating excess scatter. From session 21RtA, 9826J has a concordia age of 385.9 ± 3.8 Ma (2s) with an MSWD for concordance plus equivalence of 0.88, indicating slight underdispersion. Uncertainties and ellipses are propagated 2s absolute. Ellipses that are not shaded are excluded from concordia age calculations. The two calculated concordia ages are shown as white shaded ellipses. A common Pb correction was not applied. Figure made using IsoplotR (Vermeesch, 2018).



2. Comments regarding rejected and discordant U-Pb data

- the sample dataset which should serve as the case study to validate the workflow (rejection of >50% of the initial analysis due to "anomalous spiky patterns", followed by further rejections due to large 7/6 uncertainty, leaving with only 30% of the initial data, with no possibility to carry out provenance interpretation at the sample level due to too small n per sample; last but not least vast majority of these remaining data being largely discordant despite the different common Pb correction approached used
- Line 195: Following the correction, the data from this study are largely discordant.
- Line 280: MAJOR COMMENT. U-Pb data should be plotted as Concordia diagrams with uncertainty represented as ellipses, particularly in a manuscript submitted to a geochronology-focused journal. Alternative ways to plot the data are welcome in addition.
- Table A2: Where are these Concordia diagrams with 2s ellipses?

Reviewers #2 and #3 raised concerns about the number of analyses discarded during data reduction. The workflow that we present includes the analyses of all detrital rutile grains in a sample. *"Rutile grains were handpicked; all rutile grains were picked from most samples, except for samples 16SKY26, 16SKY42 and 17OZK05 for which 260–320 grains were selected"* (lines 238-239). Part of the reason that provenance interpretation is not possible at the sample level is due to the low rutile yield in some samples. For example, sample 18DMN01 only had 5 rutile grains, all of which were analyzed, and 4 ages were obtained. We suspect that this was a feature of local geology (i.e., metamorphic sources were not exposed at the surface at the time the sample was deposited), however, in the future we will collect larger samples to try to increase heavy mineral yield. The second reason is due to the exclusion of analyses during data reduction, for reasons including poor signal intensity and large uncertainty in the $^{207/206}$Pb ratio. Anomalous signal intensity was likely due to a combination of very low uranium concentrations, "*elemental heterogeneity from ablating into small inclusions and/or lamellae, and inhomogeneities due to micro-cracks with different element/isotope composition"* (lines 257-258).

The exclusion of analyses during data reduction is not unique to this study, but represents a quite common problem with large-*n* detrital rutile provenance work. Many published studies discuss analyzing a larger number of grains than are presented—discarding analyses—due to low U and/or low radiogenic Pb content (Bracciali et al., 2013; Caracciolo et al., 2021). For example, Shaanan et al. (2020) discard 60% of their detrital rutile U-Pb data due to discordance. Similarly, Reviewer #1 points to the study by Caracciolo and co-authors (2021) that analyzes 712 detrital rutile grains without a U-filter, yet, after discordance filtering, only 347 grains remained (48%) (from what we can tell as the data is not available online). Additionally, Caracciolo et al. (2021) did not have enough rutile ages per sample to discuss sample-by-sample provenance interpretations, which we also experienced with our dataset. This points to a larger problem in trying to scale up detrital rutile to large-*n* provenance applications. For this reason, we wanted to confidently include as many U-Pb analyses as possible in our interpretations, which led to the exploration of U-Pb discordance.

To address this comment, ***the revised manuscript will emphasize that the rejection of analyses during data reduction is not a unique limitation of this study. We will clarify that this is common in studies that have attempted large-n detrital rutile U-Pb. The original text emphasized the role of inclusions and lamellae in potentially contributing to poor signal intensity patterns and did not clearly emphasize that, in addition, low signal intensity could be from low U and/or low radiogenic Pb contents. The revised manuscript will also include low U and low Pb as a potential cause of poor raw signal intensity. As suggested by Reviewer #3, we will also include a data treatment section of the Appendix that includes a figure with examples of signal intensity patterns in the unknowns.***

We are unclear about Reviewer #2's criticism (1) of the data as discordant after common Pb correction and (2) the apparent absence of concordia diagrams. We suspect that (1) is due to a lack of clarity in what is being shown in the figures. The uncorrected results are displayed in Tera-Wasserburg space (Figures 4, 7b, 8c, 10). The common Pb corrections force concordance, so all analyses are concordant after correction. The corrected data are shown as histograms, KDEs and CDFs. The figures with data displayed in Tera-Wasserburg space are the uncorrected ratios, whereas the data displayed as histograms, KDEs and CDFs are the common Pb-corrected data (Figures 5, 7c, 12, A3); although one KDE in Figure 5 shows the full uncorrected dataset. Regarding (2), Tera-Wasserburg diagrams are concordia diagrams, but perhaps the reviewer is referring to Wetherill diagrams. However, it is generally more useful to display data from common Pb-bearing minerals in Tera-Wasserburg space as it is easy to see the linear discordia arrays between the common and radiogenic ratios (many examples, including Chew et al., 2014). ***The revised manuscript text and figure captions will clarify whether the uncorrected or common Pb-corrected data are displayed. The revised manuscript will include the figures displaying the reference materials in Tera-Wasserburg diagrams with 2s error ellipses.*** The current Figure 4 displays the unknowns in Tera-Wasserburg space with error bars so that the figure is less crowded, ***and the revised manuscript appendix will include the same figure with 2s error ellipses.***

3. Comments regarding provenance interpretations in common Pb-bearing minerals

- attempted provenance analysis based on such a largely discordant dataset, with age modes identified in the KDEs distributions of up to 60% discordant data and unsupported claim that the KDEs distributions of variably discordant data are similar hence meaningful, and can be used to constrain provenance

- Line 85: The conclusions of this study are not supported by the data; following the application of the common Pb correction methods the sample data remain largely discordant, rising questions about the effectiveness of the correction approaches and or assumptions. Importantly, no common Pb correction appears to have been applied to the primary reference material R10
- Line 130: In this study a very large portion of the "common Pb corrected" data are rejected, which makes the authors' approach as biased as the U-based filtering used by others.
- Line 140: Published studies have shown detrital rutile samples where a proportion (of the uncorrected data) is concordant, in addition to (typically one) array of discordant dates intercepting the cluster of concordant dates (Bracciali et al 2013, 2015, 2016).
- Line 255: rejection of 70% of the sample data is remarkably high, inevitably biasing the dataset and hindering any provenance interpretation (even if the remaining data were close to concordant, which are not)
- Line 285: MAJOR COMMENT. Final sample sizes are small following rejection of 70% of the initial data. The inability to assess the approach presented in this at the individual sample level is a major limitation.
- Line 290: MAIN COMMENT. The authors aim to use a set of natural samples to test their approach. Not only i) the original dataset is largely biased by rejection of ca 70% of the collected data (due to "anomalous -spiky- signals") and 7/6 uncert >20%, lines 253-256, but ii) following (207 or 208) correction, the data are largely discordant. Identifying "age modes" in KDEs generated from largely discordant data is pointless and such practice should be avoided as the modes or peaks of discordant data are geologically meaningless. The final aim of any detrital provenance study is to identify the timing of real geological events which can be tracked back to the rock sources. Age peaks or modes derived from distributions of largely discordant U-Pb data are geologically meaningless. Because of i) and ii) the results of this study cannot support any robust provenance interpretation.
- Line 300: The 100-40% concordant group includes in the first place largely discordant dates which are geologically meaningless. The similarity between the 0-40% conc and the 100-40% conc distribution does not justify inclusion of discordant dates.
- Line 300: I am afraid I have to disagree here. The comparison of the KDEs distributions of Fig 5, bottom panel indicates significant differences. A representation of the same data at a smaller scale (e.g. < 500 Ma) would enhance such differences (presence or absence of peaks, position of youngest peak). Such an overinterpretation is misleading and should not be encouraged as an acceptable practice.

Again, we suspect that the comment here is based on a lack of clarity about what is being shown in the figures. The uncorrected results are displayed in Tera-Wasserburg space (Figures 4, 7b, 8c, 10). The common Pb corrections force concordance, so all analyses are concordant after correction. The corrected data are shown as histograms, KDEs and CDFs. <u>The figures with data displayed in Tera-Wasserburg space are the uncorrected ratios, whereas the data displayed as histograms, KDEs and CDFs are the common Pb-corrected data (Figures 5, 7c, 12, A3)</u>; although one KDE in Figure 5 shows the full uncorrected dataset. ***This was not clear in the manuscript and will be clarified in the revised version.***

We suspect that this clarification addresses the concerns of Reviewer #2. Reviewer #2 argues that *"Identifying "age modes" in KDEs generated from largely discordant data is pointless and such practice should be avoided as the modes or peaks of discordant data are geologically meaningless. The final aim of any detrital provenance study is to identify the timing of real geological events which can be traced back to the rock sources. Age peaks or modes derived from distributions of largely discordant U-Pb data are geologically meaningless."* We disagree with this argument. In case the above clarification has not resolved the discrepancy, we provide an explanation below for why geologically meaningful interpretations can be made from initially discordant data.

We maintain that geologically meaningful interpretations can be made from initially discordant data when appropriate common Pb corrections are applied. We again emphasize that the initially discordant data are concordant after common Pb correction; our interpretations are based on concordant data. U-Pb discordance in common Pb-bearing minerals is well documented in published reference materials and unknowns (e.g., Chew et al., 2011, 2014). Practically everyone using detrital rutile data is using data that is discordant before correction. Many publications focus on common Pb corrections and how to treat discordance—so that accurate interpretations can be made from initially discordant data (e.g., Faure, 1986; Williams, 1997; Ludwig, 1998; Andersen, 2002; Chew et al., 2011, 2014; McLean et al., 2011; Thomson et al., 2012; Smye and Stockli, 2014; Vermeesch, 2020, 2021). We note that many *in-situ* studies with common Pb-bearing minerals fit discordia arrays to co-genetic grains to derive a ^{207}Pb-corrected age. In these instances, individual analyses can be nearly 100% discordant and still interpreted confidently within the population of co-genetic grains (e.g., Poulaki et al., 2023). We are unsure why Reviewer #2 dismisses discordant rutile data, especially when discordant data are present and interpreted in their own papers that they referenced.

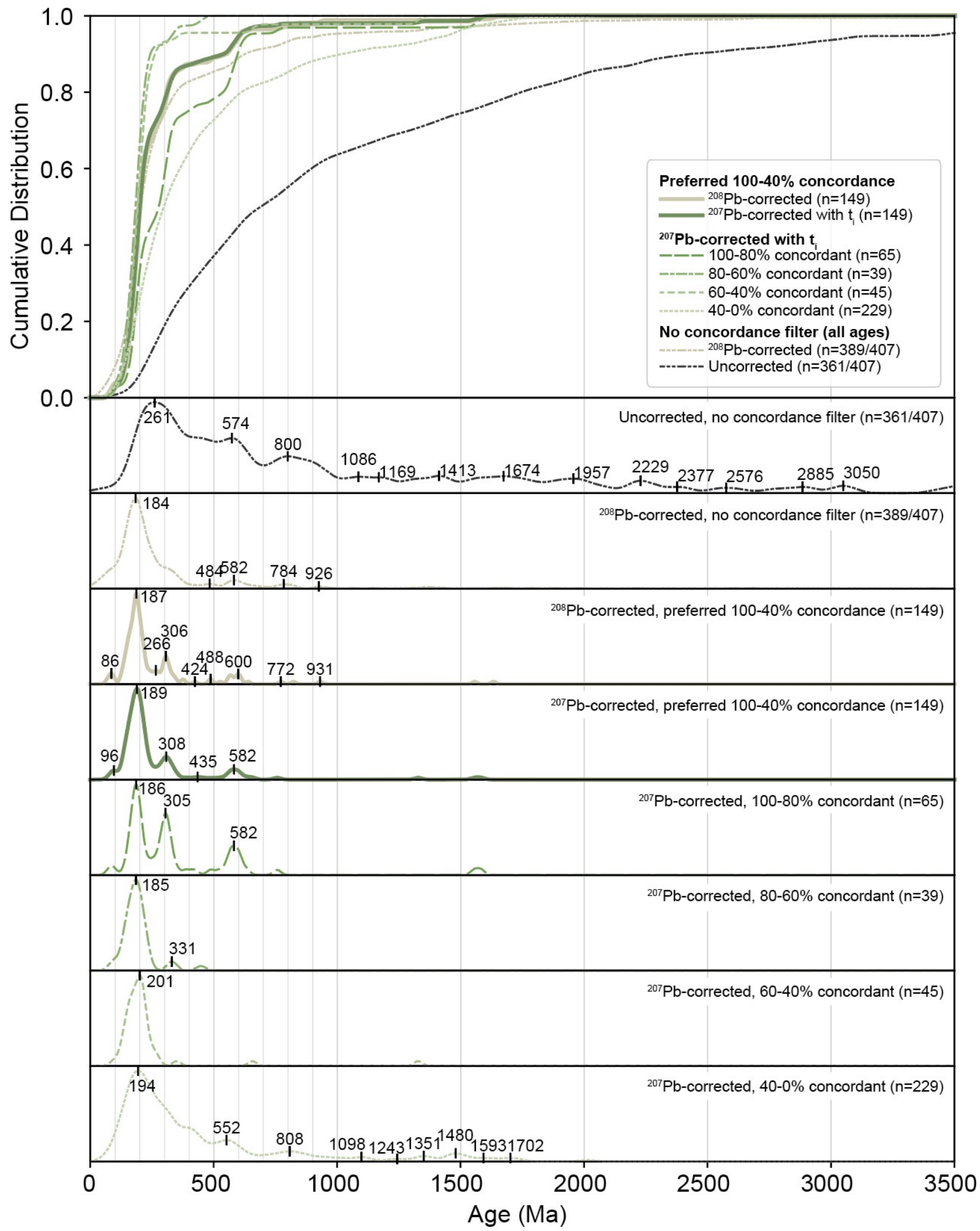
Reviewer #2 argues that *"Published studies have shown detrital rutile samples where a proportion (of the uncorrected data) is concordant, in addition to (typically one) array of discordant dates intercepting the cluster of concordant dates (Bracciali et al 2013, 2015, 2016)."* We agree that rutile U-Pb analyses can be concordant (e.g., Rösel et al., 2019, Kooijman et al. 2010, and many others), which is also the case with some of our analyses (10% of the analyses are >90% concordant), but whether a detrital population has only one age mode is dependent on the particular sedimentary system, so it is not universally true that there will be only one array of discordant analyses. We do not agree that detrital data ought to conform to one discordia array, when multiple age populations are expected in detrital samples from most tectonic settings (i.e., Govin et al., 2018 and most studies on detrital U-Pb geochronology). In Rösel et al. (2019), the detrital rutile grains are concordant and interpreted to be sourced from high grade metamorphic rocks, which could indicate a lithologic control on discordance.

Finally, with regards to our study, we used the U-Pb dates combined with protolith information and temperature from trace elements to inform provenance interpretations. In Section 7, we demonstrate that the resulting U-Pb age peaks correspond to plausible sedimentary sources that match the timing of metamorphism, protolith, and temperature of those sources. Therefore, we maintain that our dataset provides meaningful provenance information, and that the conclusions are supported by the data.

From the premise that meaningful interpretations can be made from initially discordant data, we wanted to explore, *"do different discordance filters influence the resulting age spectra and provenance interpretations or not?"* (lines 148-149), in other words, *how discordant is 'too discordant'?* Answering this question requires identifying a metric or threshold for 'too discordant.' We approached this by comparing the uncorrected, ^{207}Pb-corrected, and ^{208}Pb-corrected dates. The unfiltered, uncorrected dates yielded a large, unimodal age peak that is useless for interpretation (Figure 5). In the comparison of ^{207}Pb corrected dates, the age difference *"is less than 1% for analyses less than 60% concordant and less than 5% for analyses 60–40% concordant"* (line 295) (Figure 6). This is why we chose to include all analyses up to 60% discordant (above 40% concordant). In a separate comparison, we binned the ^{207}Pb- and ^{208}Pb-corrected dates by their percent discordance (Figure 5, see also Figure 4). Rather than zooming in on the < 500 Ma ages, here we show a modified Figure 5 (Figure R5) with labeled KDE peaks (using detritalPy; Sharman et al., 2018). There is variability in age modes and their amplitude between some of the discordance bins (which we discussed in lines 293-304). For example, the ~190 Ma peak is slightly older at 200 Ma in the 60-40% concordant bin. Between all of the concordance bins, the main age modes are generally at 95 Ma, 190 Ma, 310 Ma, and 580 Ma. We disagree with the Reviewer's characterization that *"the comparison of the KDEs distributions of Fig 5, bottom panel indicates significant differences."* The similarity between the ^{207}Pb-corrected, and ^{208}Pb-corrected dates for grains 100-40% concordant is shown in Figure 5 as KDEs and CDFs and in the statistical comparisons in supporting information (Dataset S3 in Mueller et al., 2023). We view the similarity between the ^{207}Pb-corrected and ^{208}Pb-corrected dates for grains 100-40% concordant as a positive check on the discordance filter chosen here. Therefore, we maintain that reliable provenance interpretations can be made from data above 40% concordant in our specific dataset (Section 5.1).

Regarding potential bias in data reduction and filtering, it is possible that the rutile analyses excluded due to low U, low Pb, inclusions, or high 207/206 uncertainty could impart a bias on the age distributions. It is difficult to demonstrate this one way or another as we cannot know the age distribution of excluded grains. They were excluded precisely because their age information was not usable. To reduce bias, we aimed to include as many analyses as possible by including discordant data. The effects of a strict discordance filter are shown in Figure 4. If we had applied a 20% discordance filter, as is common in detrital zircon work and less common in detrital rutile (Shaanan et al., 2020), the results would be only the 100-80% concordance group (which constitutes only 65 analyses). Although, we define concordance here as distance to concordia along the discordia (after Vermeesch, 2021) rather than the relative age difference between the ^{206}Pb/^{238}U and ^{207}Pb/^{206}Pb dates. The 100-80% concordance group has similar dominant age modes as our preferred 100-40% concordance group; however, the KDE amplitudes are different, likely related to *n*. This does not change the resulting interpretations in our dataset, but this may be important for other datasets for which a stricter concordance filter alters date distributions and/or peak amplitude is important (see also lines 293-304).

Figure R5 (next page). Distributions of U-Pb dates of all samples together displayed as kernel density estimates (KDEs) with labeled age peaks, and cumulative distribution functions (detritalPy; Sharman et al., 2018). Both uncorrected and corrected data are displayed. The ^{207}Pb-corrected data are separated into concordance bins. Note that the KDEs are not relative plots. The main age modes present throughout the date distributions are ca. 95 Ma, 190 Ma, 310 Ma, and 580 Ma.



4. Additional line by line comments, excluding typographic comments

- Line 40: Detrital rutile mineralogy and geochemistry have been routinely applied to constrain provenance for almost two decades, rutile U-Pb chron. for a decade.

We agree and included detrital rutile in our summary of detrital geochronometers in lines 62-65.

- Line 40: I suggest to rewrite the introduction focusing on detrital rutile as a provenance tracker and a complementary proxy to other single grain and/or bulk techniques. Reference to the theory of plate tectonics or the evolution of sedimentary provenance studies since the 70s is off topic.

The revised manuscript will change from a focus on new workflows to detrital rutile provenance, suture zone settings, and Anatolia geology.

- Line 45: Sedimentary provenance includes detrital geo- and thermo-chronology of a wide range of accessory minerals as established techniques, not only mineralogy and geochemistry of sediments.

We agree and included a summary of detrital geochronometers in lines 62-65.

- Line 50: In 2023, detrital geochron (especially of zircon) is by far an established provenance tool. Non relevant terminology (classic as opposed to... modern?) should be avoided. Single mineral approaches are not elevated relative to bulk methods. In fact, single and bulk methods can be complementary in constraining provenance. The choice of method(s) depends on the scientific question.

We agree and did not mean to imply that bulk methods are outdated, as we use these methods ourselves. ***We will rephrase.***

- Line 55: Zircon is an ideal mineral provenance proxy due to i) its widespread occurrence in crustal rocks (certainly not a limitation) ii) its chemical and mechanical resistance to geologic and sedimentary processes. Please avoid referrring to "the zircon problem", which is unjustified and misleading. Rutile occurs in a narrower range of crustal rocks compared to zircon, but that does not mean that a "rutile problem" should be invoked.
    - Line 60: Sedimentary provenance based on one technique (any technique) can be incomplete, or not. It depends on the scientific question(s) being addressed. An interpretation based on rutile only could also be incomplete. Please delete or rewrite.
    - Line 65: There is no such "zircon problem". Avoid using such terminology. We could likewise refer to "any mineral" problem.
    - Line 70: Please delete. Detrital rutile is an established proxy and there is no such problem to overcome.

We will clarify that zircon is an excellent provenance proxy due to the reasons stated. The rise in the application of other detrital phases has revealed quite well that zircon does a good job in reconstructing provenance, but often does not capture the entire provenance picture (e.g., Moecher and Samson, 2006; Hietpas et al. 2010; Gaschnig, 2019). We don't find it misleading to characterize this as a "zircon problem" as many studies are reporting to reconstruct provenance with zircon alone even though zircon is primarily sourced from felsic to intermediate igneous rocks, thereby underrepresenting mafic rock types in the source, for example. Pereira and Storey (2023) refer to this as "zircon's blind spots" (pgs. 3, 4). However, we did not mean to imply that rutile should replace zircon as the only provenance tool used, but rather used as a complementary dataset (line 20). We agree that bulk and single mineral approaches ought to be used together to reconstruct a full provenance history.

- Line 70: Mention the metamorphic (and igneous) rocks where rutile can form, with references, and why it can be commonly found in sedimentary rocks. Metamafic and metapelic (rutile) are terms introduced by Meinhold (2010, ESR) to broadly categorize rutile provenance.

We believe that we have already included this information in our short introductory review: *"Rutile can form in metamafic and metapelitic rocks across a range of P-T conditions, therefore, detrital rutile is especially advantageous when tracking sediment input from greenschist to eclogite facies sources (e.g., Zack and Kooijman, 2017)."* (lines 69-72).

- Line 70: MAIN COMMENT "U-Pb closure temperature" is not correct as the notion of closure refers to volume diffusion of Pb. Please cite the work by Cherniak, Contrib Min Pet 2000 and other studies showing how the open-system vs close-system behaviour of radiogenic Pb is better described by the notion of partial retention zone and on which natural parameters is dependant. The review by Smye et al 2018, Chem Geol is a good source of references.

- Line 75: This statement is not entirely correct. First cycle detrital rutile can track cooling following any thermal event capable of resetting the U-Pb clock in rutile. Please note that: i) any (high-grade) metamorphic unit must be exposed to surface before being eroded; ii) when cooling is "slow" the detrital rutile U-Pb date can substantially post-date the "age" of the metamorphic event. Please amend the introduction to include examples (e.g Flowers et al 2005 Geology, fig.2 where zircon is 110 Ma, rutile 75 Ma in basement crystalline rocks; other examples in Fig. 4 of Bracciali 2019, Geosci.)

The idea of closure temperature is an oversimplification, but is still a widely used term in the thermochronology community that will be familiar to most readers of this manuscript. The original concept of the closure temperature concept by Dodson (1973) clearly includes the existence of a partial retention zone (Dodson 1973, Fig. 1). The closure temperature defined there is a construct, whose dependencies and limitations are widely known. We argue that it is still convenient and acceptable to use this term and concept to compare the general temperature sensitivity of different chronometers. As pointed out by the reviewer, a given chronometer is sensitive (among other things) to a temperature range that is determined by the diffusion kinetics of the radiogenic isotope, in this case Pb. In this way, thermochronometers have a temperature range where child products are partially retained, the partial retention zone, which is sensitive to factors such as grain size and cooling rate. The U-Pb date is reflective of the interplay between the kinetics of diffusion and/or annealing and accumulation rates. We wrote, *"rutile U-Pb dates correspond cooling from the most recent medium to high-temperature metamorphic event that exceeded the closure temperature"* (lines 72-73); ***the revised manuscript will clarify that slow cooling rates can produce U-Pb dates significantly younger than the timing of metamorphism (e.g., Möller et al. 2000). This is relevant to our dataset as rutile U-Pb dates around 190 Ma are 10-25 Myr younger than estimates of peak metamorphism in the Karakaya Complex. We will mention this in the discussion.***

- Line 80: Alternative correction approaches are not acknowledged in this manuscript. Bracciali et al (2013 Chem Geol; 2016 ESR) corrected only the discordant data using the terrestrial Pb-evolution model of Stacey and Kramers (1975) and no Pbc correction applied to the primary reference material (containing negligibile common Pb). Please include

We discuss the use of the Stacey and Kramers (1975) terrestrial Pb evolution model to correct discordant data in the ^{208}Pb correction and ^{207}Pb correction sections (Section 2.3.1 and 2.3.2, respectively) and have not applied any corrections to the reference materials.

- Line 90: MAJOR COMMENT. This detrital rutile synopsis is only based on geochemistry of rutile (and derived thermometry). No mention to U-Pb dating of rutile (by LA and ID-TIMS), in a manuscript submitted to G-Chron. Please rewrite and inlcude relevant literature.

This is a good point. We discussed rutile geochemistry in the overview of Section 2.1 and common Pb corrections in Section 2.3 but did not include a discussion of U-Pb dating methods. ***The revised manuscript will include a short overview of rutile U-Pb.***

- Line 120: This is not necessarily the case. A rutile with a low U content, but "old" (e.g. Ga), having accumulated enough radiogenic Pb could be easier to date than a "young" rutile (e.g. few Ma) richer in U. Additionally, the lowest measurable radiogenic Pb signal intensities (affecting the precision of the isotope ratios) will vary depending on the LA technique of choice (quadrupole, vs SC-SF- vs MC-SF ICP-MS).

Yes, ***the revised manuscript can include statements that both low U and young rutile can be challenging to date.*** We already discussed how instrumentation can affect signal intensity in our discussion of U-threshold filtering (lines 305-320).

- Line 120: this is not correct: high U rutiles could be discordant because of of a high relative Pbc content

We show in Figure 7 that rutile above 4 ppm U, the high-U rutile, have a higher proportion of concordant analyses (n=60/68, 88%) than low U rutile. Yes, there are several high U rutile that are discordant (n=8/68, 12%) due to high Pbc (Figure 7b). We would expect that this proportion might vary across datasets. ***We will clarify in line 120 that screening low-U rutile produces a higher proportion of analyses with acceptably high U and Pb signal intensities and seems to result in a higher proportion of concordant analyses.***

- Line 125: The omission or inclusion of low-U rutile does not "make sense" depending on the geological setting: either it is deemed a valid data handling approach, or not. Filtering on U content (in essence rejecting analyses which are expected to fail) is equivalent to discarding analyses where the very low-intensity measured signals would result in poorly determined isotopic ratios with an extremely large uncertainty (e.g. tens %, 2s), since this would make the corresponding dates hardly usable for any provenance interpretation. The application of such filtering before any common Pb correction (as appropriate) must be discussed.

We are uncertain of the exact reason why the published studies used a U threshold and gave them the benefit of the doubt by assuming that they had a valid reason for doing so. The fact is that such datasets exist, and for the region of our study constitute half of the publications (2 of 4). One idea is that they were not expecting metamafic sources and therefore expected the majority of rutile to have high U. Here, we show that the U thresholds used are too high, as rutile with U < 4-5 ppm can yield dates with acceptable precision and concordance (see Section 5.2, Figure 7). And setting a specific ppm level as threshold is not good practice as a general recommendation because of the aforementioned dependence of LOD on instrumentation and analytical parameter choices. Therefore, filtering on U content is not "rejecting analyses which are expected to fail" as we demonstrate that they do not all fail. In fact, excluding rutile below 4 ppm U biases the dataset (discussed in Section 5.2). Regarding the removal of analyses with large uncertainties in the isotopic ratio: *"214 analyses were excluded for 207Pb/206Pb error above 20%"* (line 255). The analyses that were removed generally had low uranium concentrations: maximum of 8.9 ppm, average of 0.5 ppm, minimum of 0.02 ppm (Dataset S1). See also reply to point #1 above.

- Line 255: clarify which control goals were used

This is stated in the same sentence: *"Even with the modified protocol, a significant number of analyses did not meet quality control goals: 686 of 1,278 analyses were excluded due to anomalous (spiky) patterns in raw signal intensity and a further 214 analyses were excluded for 207Pb/206Pb error above 20%, leaving 378 analyses remaining (30%)."* (lines 253-256). The quality control goals are acceptable signal intensity, both high enough U and Pb for data reduction (***we will add this statement***) and smooth, non-spiky, signal with 207/206 uncertainty below 20% (2s). ***We will clarify that the 20% uncertainty is at the 2-sigma level.***

- Line 270: Trace elements in rutile can easily sum-up to 1-2%

The TiO2 internal standardization varies across the literature, with TiO2 normalization commonly 100 mass-% (e.g., Rösel et al., 2019; Plavsa et al., 2018), 99 mass-% (e.g., Ewing et al., 2013), or 98 mass-% (e.g., Hart et al., 2018). The TiO2 internal standardization value and reproducibility of standards are not consistently reported, so it is hard to assess. The choice of TiO2 normalization to values between 100 and 98 mass-% is below the reproducibility of reference materials and does not significantly influence the calculated trace element results.

Figure R6 displays the trace element results of the three reference materials for the four analytical sessions. The trace element results are within 10% for all of the main elements discussed in the paper: Cr, Nb, and Zr. For the secondary standard GSC-1G, all elements are within 10% of the published values except for Sn and Ga. The R10 rutile reference material displayed internal heterogeneity in trace element composition, which we noticed when comparing the trace element results with the laser ablation spot coordinates. This is also reflected in the range of trace element compositions reported for R10 in the GeoReM database (http://georem.mpch-mainz.gwdg.de/). The R10 results are within 10% for many elements but the offset is much higher for some elements. We note that all of the R10 results are within the range of reported values from the GeoReM database. Reproducibility of around 10% is consistent with the literature. For example, Ewing et al. (2013) report that trace element measurements generally reproduce reference materials to within 11%, but there are other reports of concentrations varying up to 30% for some elements (Plavsa et al., 2018). ***The revised manuscript will include a discussion of reproducibility in the main text and appendix.***

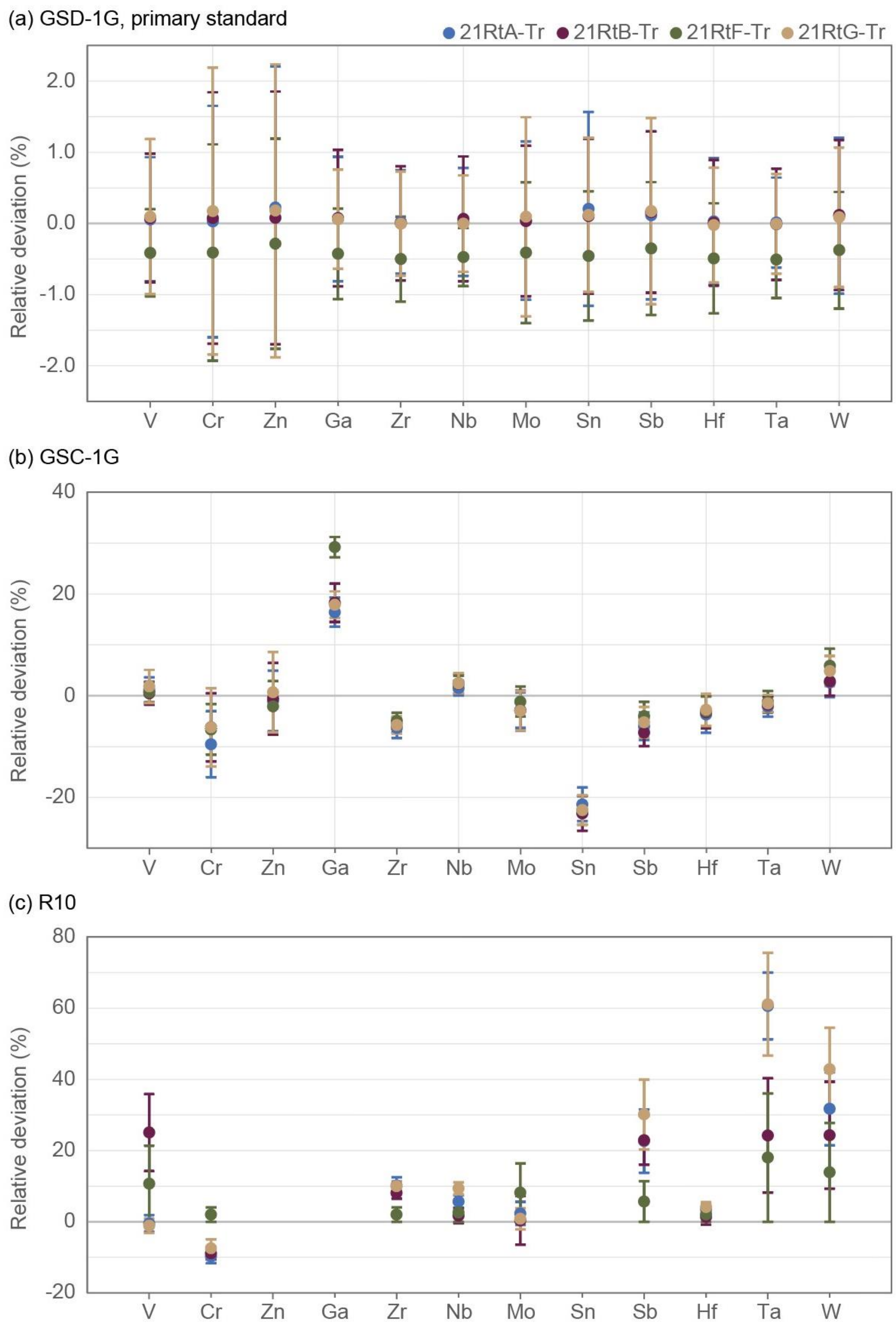


Figure R6 (previous page). Plots of percent deviation of the trace elements for reference materials GSD-1G, GSC-1G and R10. The data were calibrated using GSD-1G and Ti as an internal standard element. Error bars show the standard deviation.

- Line 315: It is expected that when a minimum signal intensity (CPS or V) has to be set to filter the data, such a threshold will vary depending on the sensitivity of the technique (Q-ICP-MS, SF-SC, SF-MC).

Yes, we believe we stated this in lines 305-320, that the U concentration filter does not make sense as it is instrument and parameter dependent. To clarify again, we are not advocating for such a threshold, but it exists in the published literature on detrital rutile and its potential effects need to be discussed. ***We will check that this is clear in the revised manuscript.***

References Cited

Andersen, T.: Correction of common lead in U–Pb analyses that do not report 204Pb, Chem. Geol., 192, 59–79, https://doi.org/10.1016/S0009-2541(02)00195-X, 2002.

Bracciali, L., Parrish, R. R., Horstwood, M. S. A., Condon, D. J., and Najman, Y.: UPb LA-(MC)-ICP-MS dating of rutile: New reference materials and applications to sedimentary provenance, Chem. Geol., 347, 82–101, https://doi.org/10.1016/j.chemgeo.2013.03.013, 2013.

Caracciolo, L., Ravidà, D. C. G., Chew, D., Janßen, M., Lünsdorf, N. K., Heins, W. A., Stephan, T., and Stollhofen, H.: Reconstructing environmental signals across the Permian-Triassic boundary in the SE Germanic Basin: A Quantitative Provenance Analysis (QPA) approach, Glob. Planet. Change, 206, 103631, https://doi.org/10.1016/j.gloplacha.2021.103631, 2021.

Chew, D. M., Sylvester, P. J., and Tubrett, M. N.: U–Pb and Th–Pb dating of apatite by LA-ICPMS, Chem. Geol., 280, 200–216, https://doi.org/10.1016/j.chemgeo.2010.11.010, 2011.

Chew, D. M., Petrus, J. A., and Kamber, B. S.: U-Pb LA-ICPMS dating using accessory mineral standards with variable common Pb, Chem. Geol., 363, 185–199, https://doi.org/10.1016/j.chemgeo.2013.11.006, 2014.

Dodson, M. H.: Closure Temperature in Cooling Geochronological and Petrological Systems, Contributions to mineralogy and petrology, 40, 259–274, 1973.

Ewing, T. A., Hermann, J., and Rubatto, D.: The robustness of the Zr-in-rutile and Ti-in-zircon thermometers during high-temperature metamorphism (Ivrea-Verbano Zone, northern Italy), Contrib. Mineral. Petrol., 165, 757–779, https://doi.org/10.1007/s00410-012-0834-5, 2013.

Faure, G.: Principles of Isotope Geology, 2nd Edition., Wiley & Sons, Inc., 608 pp., 1986.

Frei, D. and Gerdes, A.: Precise and accurate in situ U–Pb dating of zircon with high sample throughput by automated LA-SF-ICP-MS, Chemical Geology, 261, 261–270, https://doi.org/10.1016/j.chemgeo.2008.07.025, 2009.

Gaschnig, R. M.: Benefits of a Multiproxy Approach to Detrital Mineral Provenance Analysis: An Example from the Merrimack River, New England, USA, Geochem. Geophys. Geosystems, 20, 1557–1573, https://doi.org/10.1029/2018GC008005, 2019.

Govin, G., Najman, Y., Copley, A., Millar, I., van der Beek, P., Huyghe, P., Grujic, D., and Davenport, J.: Timing and mechanism of the rise of the Shillong Plateau in the Himalayan foreland, Geology, 46, 279–282, https://doi.org/10.1130/G39864.1, 2018.

Hart, E., Storey, C., Harley, S. L., and Fowler, M.: A window into the lower crust: Trace element systematics and the occurrence of inclusions/intergrowths in granulite-facies rutile, Gondwana Res., 59, 76–86, https://doi.org/10.1016/j.gr.2018.02.021, 2018.

Hietpas, J., Samson, S., Moecher, D. and Schmitt, A.K., 2010. Recovering tectonic events from the sedimentary record: Detrital monazite plays in high fidelity. *Geology*, *38*(2), pp.167-170.

Kooijman, E., Mezger, K., and Berndt, J.: Constraints on the U–Pb systematics of metamorphic rutile from in situ LA-ICP-MS analysis, Earth and Planetary Science Letters, 293, 321–330, https://doi.org/10.1016/j.epsl.2010.02.047, 2010.

Kylander-Clark, A. R. C.: Slow subduction and exhumation of a thick ultrahigh -pressure terrane: Western Gneiss Region, Norway, Ph.D., University of California, Santa Barbara, United States -- California, 121 pp., 2008.

Lippert, P. G.: Detrital U-Pb geochronology provenance analyses: case studies in the Greater Green River Basin, Wyoming, and the Book Cliffs, Utah, Thesis, University of Kansas, 2014.

Ludwig, K. R.: On the Treatment of Concordant Uranium-Lead Ages, Geochim. Cosmochim. Acta, 62, 665–676, https://doi.org/10.1016/S0016-7037(98)00059-3, 1998.

Malusà, M. G., Carter, A., Limoncelli, M., Villa, I. M., and Garzanti, E.: Bias in detrital zircon geochronology and thermochronometry, Chemical Geology, 359, 90–107, https://doi.org/10.1016/j.chemgeo.2013.09.016, 2013.

McLean, N. M., Bowring, J. F., and Bowring, S. A.: An algorithm for U-Pb isotope dilution data reduction and uncertainty propagation, Geochem. Geophys. Geosystems, 12, https://doi.org/10.1029/2010GC003478, 2011.

Moecher, D.P. and Samson, S.D., 2006. Differential zircon fertility of source terranes and natural bias in the detrital zircon record: Implications for sedimentary provenance analysis. *Earth and Planetary Science Letters*, *247*(3-4), pp.252-266.

Möller, A., Mezger, K. and Schenk, V., 2000. U–Pb dating of metamorphic minerals: Pan-African metamorphism and prolonged slow cooling of high pressure granulites in Tanzania, East Africa. Precambrian Research, 104(3-4), pp.123-146.

Mueller, M., Licht, A., Möller, A., Condit, C., Fosdick, J. C., Ocakoğlu, F., and Campbell, C.: Supplemental data for: An expanded workflow for detrital rutile provenance studies: An application from the Neotethys Orogen in Anatolia, https://doi.org/10.17605/OSF.IO/A4YE5, 2023.

Odlum, M. L., Stockli, D. F., Capaldi, T. N., Thomson, K. D., Clark, J., Puigdefàbregas, C., and Fildani, A.: Tectonic and sediment provenance evolution of the South Eastern Pyrenean foreland basins during rift margin inversion and orogenic uplift, Tectonophysics, 765, 226–248, https://doi.org/10.1016/j.tecto.2019.05.008, 2019.

Plavsa, D., Reddy, S. M., Agangi, A., Clark, C., Kylander-Clark, A., and Tiddy, C. J.: Microstructural, trace element and geochronological characterization of TiO2 polymorphs and implications for mineral exploration, Chem. Geol., 476, 130–149, https://doi.org/10.1016/j.chemgeo.2017.11.011, 2018.

Poulaki, E. M., Stockli, D. F., and Shuck, B. D.: Pre-Subduction Architecture Controls Coherent Underplating During Subduction and Exhumation (Nevado-Filábride Complex, Southern Spain), Geochem. Geophys. Geosystems, 24, e2022GC010802, https://doi.org/10.1029/2022GC010802, 2023.

Rösel, D., Zack, T., and Möller, A.: Interpretation and significance of combined trace element and U–Pb isotopic data of detrital rutile: a case study from late Ordovician sedimentary rocks of Saxo-Thuringia, Germany, Int. J. Earth Sci., 108, 1–25, https://doi.org/10.1007/s00531-018-1643-5, 2019.

Shaanan, U., Avigad, D., Morag, N., Güngör, T., and Gerdes, A.: Drainage response to Arabia–Eurasia collision: Insights from provenance examination of the Cyprian Kythrea flysch (Eastern Mediterranean Basin), Basin Res., n/a, https://doi.org/10.1111/bre.12452, 2020.

Sharman, G. R., Sharman, J. P., and Sylvester, Z.: detritalPy: A Python-based toolset for visualizing and analysing detrital geo-thermochronologic data, Depositional Rec., 4, 202–215, https://doi.org/10.1002/dep2.45, 2018.

Smye, A. J. and Stockli, D. F.: Rutile U–Pb age depth profiling: A continuous record of lithospheric thermal evolution, Earth Planet. Sci. Lett., 408, 171–182, https://doi.org/10.1016/j.epsl.2014.10.013, 2014.

Stacey, J. S. and Kramers, J. D.: Approximation of terrestrial lead isotope evolution by a two-stage model, Earth Planet. Sci. Lett., 26, 207–221, https://doi.org/10.1016/0012-821X(75)90088-6, 1975.

Thomson, S. N., Gehrels, G. E., Ruiz, J., and Buchwaldt, R.: Routine low-damage apatite U-Pb dating using laser ablation–multicollector–ICPMS, Geochem. Geophys. Geosystems, 13, https://doi.org/10.1029/2011GC003928, 2012.

Vermeesch, P.: IsoplotR: A free and open toolbox for geochronology, Geoscience Frontiers, 9, 1479–1493, https://doi.org/10.1016/j.gsf.2018.04.001, 2018.

Vermeesch, P.: Unifying the U–Pb and Th–Pb methods: joint isochron regression and common Pb correction, Geochronology, 2, 119–131, https://doi.org/10.5194/gchron-2-119-2020, 2020.

Vermeesch, P.: On the treatment of discordant detrital zircon U–Pb data, Geochronology, 3, 247–257, https://doi.org/10.5194/gchron-3-247-2021, 2021.

Williams, I. S.: U-Th-Pb Geochronology by Ion Microprobe, in: Applications of Microanalytical Techniques to Understanding Mineralizing Processes, Society of Economic Geologists, 1–35, https://doi.org/10.5382/Rev.07.01, 1997.