The manuscript by Muston et al present 40Ar/39Ar furnace step-heating and 39Ar diffusion data from alunite in ten samples from the Martabe epithermal gold deposit in Indonesia. They use their closure temperature constraints to conclude that gold in the Purnama pit was the result of fluid rock interactions from ~ 2.25 and ~2.00 Ma.

This manuscript is filled with numerous misleading statements, but more importantly it is plagued by a lack of up to date knowledge of geochronologic methods and flawed data interpretations. For these reasons, which are explained in much more detail below, I strongly urge that this manuscript be rejected.

<u>Lack of knowledge of published studies and use language that could mislead a non-specialist:</u>
Line 56: There are numerous other examples of 40Ar/39Ar dating of alunite. All published data on alunite do not need to be cited here, but other work besides Arribas et al should be cited. Vasconcelos 1999 provides a nice overview of dating supergene minerals. It should be cited here.

Line 224: "These are the first detailed ultra-high-vacuum (UHV) furnace-step-heating results for alunite that have been reported in the literature." This is not true. Polyak et al 1998 Science paper used a furnace to date Carslbad alunite. See another paper by Lin et al 2017. There are numerous other examples. This study is not the first to use a furnace on alunite. The text repeatedly touts the detailed 30+ step heating experiments. However, upon closer investigation, in most of the experiments in Figs 3-10, the first 12-20 steps comprise only 5-10% of the 39Ar released and the remaining 12-15 steps makeup the rest of the experiment. They have essentially done 12-15 step experiments, which are the exact type of experiments that the authors denounce in this manuscript. Thus, the repeated reference to detailed step heating experiments needs to be downplayed. Moreover, there is a whole section of the manuscript (Discussion 4.5) that sheds a negative light on laser analyses. On line 145, the text states that the "detailed heating schedule was chosen so that only small amounts of gas are released in each step…..The resultant age spectrum allows vastly more information to be ascertained than any laser spot analysis can provide." This statement is extremely misleading. I encourage the authors to look at any of the recent work by Anthony Koppers and colleagues on Cretaceous seamounts. Those incremental heating experiments have 30-40 steps and show sufficient detail. There are numerous other examples of detailed laser experiments also. These experiments are typically done on way less than the 100-150mg of material used by the authors in this study. The entire section 4.5 discounting laser step heating experiments should be omitted because a non-specialist may read this and think that they must use a furnace to step heat alunite and that is simply not true. There are numerous examples of laser step heating of alunite that yield plateaus that make geologic sense.

Lines 124-126: "The method utilised allows protracted cleaning of the furnace between samples, so furnace blanks are consistently reduced to low levels. Corrections are done by interpolation, but in general the blanks are so low that this is not essential."
What is meant by non-essential? Has a blank correction been performed? What is the size of the 36Ar signals relative to the blanks? How many blanks are measured over the temperature range of a typical experiment. The disadvantage of furnace step heating is you cannot assess the blank variability during the experiment and must interpolate blanks before and after the experiment when the furnace is clean. If 100 to 150 mg of alunite is being step heated, the blanks are undoubtedly going to vary during the experiment. Significantly more detail needs to be provided on number of blanks performed, how they vary with temperature, how big they are compared to the signal for each isotope, etc.

Paragraph beginning on line 345: There is more text with a negative tone towards laser methods here. This whole paragraph is misleading and not needed. For example, the text says, “It is rare for a laboratory to have the technical capability to conduct step-heating experiments”. Almost all 40Ar/39Ar labs have the capability to conduct step heating experiments. These types of statements are potentially very dangerous to a non-specialist who may be interested in using geochronology for a project.

Flawed data interpretations:

There is no doubt that some samples produce complex apparent 40Ar/39Ar age spectra, which may be due in part to multiple episodes of alteration. The authors propose to see through this complexity by using the asymptotes and limits method of Forster and Lister 2004. When conducting incremental heating experiments, by far the most common convention is to use the plateau age, isochron age, or in some cases an integrated age with caution if the sample does not form a statistically acceptable plateau. The authors have chosen not to use any of these methods and have opted for their own method. The text needs to clearly state that they are not using any common data interpretation methods and have chosen to use the asymptotes and limits method, which is not a conventional approach.

Line 231: “All of the age spectra show evidence of contamination in the first steps, with high initial ages that decrease, often to a small plateau age segment.” First of all, these first few steps comprise 5-10% of the 39Ar released. It is by no means a plateau. The word plateau should be omitted throughout the manuscript when talking about these mini segments. The authors are interpreting a few overlapping steps for 5% of the released gas to have significance. There is no age significance to this. Just like there is no significance to the ages later on in most of the experiments that overlap but are sloping downward and are only two steps (blue steps). The probability analysis of all the apparent ages is deeply flawed as many of the steps do not reflect the actual age of the alunite. These apparent ages may be affected by processes such as excess Ar, recoil, K loss, all of which are minimally discussed.

The spectra for each sample look very similar. Almost all of them get younger in the last 20% of the experiment. Coincidentally the Ca/K increase during this time. Yet there is very little mention of recoil. The authors instead attribute the decrease in age at the end of the experiment to mixing with a KCl/CaCl2 gas reservoir derived from inclusions. It is unlikely that all of the samples would be plagued by inclusions. Images of inclusions from each sample would help this argument if they are available. 39Ar recoil needs to be discussed as an alternative.

The isochrons have no statistics (no intercepts with uncertainties, no MWSD, etc) and thus they have little value as presented. The authors state that isochrons are not preferred for these types of samples. The reader cannot evaluate if the isochrons are of value until all of the data associated with them are provided.

There is a whole section on how the authors don’t agree with plateau criteria. The authors are entitled to their own opinion. However, numerous published studies have show that alunite can produce plateaus (Corral, 2021, Pan et al 2019, Ren and Vasconcelos, 2019, Mote et al 2001, and many more). Coincidentally most of these studies used a laser. It may be that the Martabe alunite are indeed very complicated. If they are going to argue for multiple generations of fluid interaction, there should be some chemical or mineralogic evidence to support this (overgrowths, zones, etc). The argument cannot be made solely based on complicated Ar spectra.

Line 96-100: These sentences read like they are straight out of a textbook and should be omitted or drastically streamlined. Most readers of GChron are going to be aware of what a fluence monitor is and why it is used.

There are 14 Figures. Ten of them look exactly the same except the data within them are different. I urge the authors to present a few examples and put the others in a supplement. This will reduce the redundancy of the figures.