Peer Review File

Manuscript Title: Spatial variations in aromatic hydrocarbon emission in a dust-rich galaxy

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

Following is my report on 'Emission from polycyclic aromatic hydrocarbons resolved in the early universe,' by Spilker et al., submitted for publication in Nature.

They present the discovery of emission by small, ring-like dust known as polycyclic aromatic hydrocarbons, from a galaxy at very high redshift (z = 6), using the JWST. The galaxy is strongly gravitationally lensed (30x mag). The line-to-continuum ratio implies that the dominant heating source is star formation, not a dust-obscured AGN, although the high resolution afforded by JWST, and the lensing angular maganification, allow them to show that the PAH emission, the FIR emission from large dust grains, and the stars, all have different spatial distributions on kpc-scales, indicating large variations in the local UV radiation field.

These results have pushed the use of PAH lines as diagnostics of the ISM in galxeis to the largest look-back times, and provide unique information on small-scale processes in the first galaxies. As such, it is a unique and important result, both as a science study in itself, and as a demonstration for potential future studies with the very new JWST. I can recommend this paper for publication in Nature. I have a few comments below that might help with the presentation.

1. I think they could mention in a little more detail the comparison to the next highest z source for which PAH emission has been detect, and that is GN20 (reference 7). I believe in that case star formation was also favored for dust heating?

2. Figure 2 caption should mention that the rest frame 160um image (2C) is from ALMA, I believe.

3. Figure 2: what is the green dot at the center? is that the lensing galaxy? if so, can that be used to derive the stellar mass of the lens, and hence mass to light ratio for the lens, or has that been done in previous work?

4. Figure 3: Can they state more explicitly the physical size scales implied by the variations on the PAH-FIR-star spatial variations?

5. Figure 3: is the variation of the PAH/FIR ratio due to a change in dust grain size distribution, or could it be due to a variation in dust temperature?

6. A few sentences would be useful that consider the possibility for future studies of PAH emission at z > 6. Will this become routine with JWST? The last sentence alludes to this, but is very strong lensing required for this? I guess the case of GN20 argues not, but that is also a hyper-luminous IR galaxy and hence rare. Are other PAH features observable with JWST (eg. the 6.2um feature)? perhaps this information may be too much for the paper, but some information in the supplemental material would be good for posterity.

Referee #2 (Remarks to the Author):

Dear Dr. Sage,

I have reviewed the manuscript "Emission from polycyclic aromatic hydrocarbons resolved in the early universe" and my comments are attached below. I think the detection of the 3.3 micron PAH feature in a galaxy at z~4 is very exciting and worthy of publication in Nature. That this is now possible with relatively short spectroscopic observations with JWST is remarkable. These new observational tracers of dust in distant galaxies provide important insights into how dust evolves with redshift and how it impacts ISM physics. The manuscript is clear and concise. The methodology, data analysis, statistics, and treatment of uncertainties are reasonable and well described.

I have a few comments on the interpretation of the variable 3.3/160 ratios that I hope the authors will consider, as well as two smaller points listed below.

The authors spend a fair amount of discussion on the variable 3.3/160 ratios (interpreted as L_PAH/L_TIR), arguing that they may be evidence for spatially distinct distributions of PAHs and other dust components. However, as they note in the Methods, this implicitly assumes that the dust temperature is the same everywhere. PAHs are generally in the stochastic heating regime, and so approximately linearly trace the radiation field intensity. The 160 micron emission is non-linearly dependent on the radiation field intensity. Figure 15 of Draine & Li (2007) has a very clear illustration of this.

The baseline assumption in the paper appears to be that the PAHs are failing as a TIR tracer, but it seems equally possible that there are dust temperature variations and the 160/TIR ratio is variable. To make the case that this is not simply variations of the equilibrium dust temperature, it seems important to make an argument about how large of variations are possible in 3.3/160 for reasonable radiation field intensities. This galaxy has a very large star formation rate, and it isn't clear in the text if that is distributed over a wide area or concentrated in a central starburst (or if that is known). If concentrated in a starburst, the dust temperature could vary substantially over the galaxy, leading to variations in 160/TIR. Over some ranges of radiation field intensity, it is possible that the PAH emission could be a more reliable tracer of TIR, revealing where a simple conversion of 160 to TIR is failing due to changes in the equilibrium dust temperature.

- One point that would help in understanding the 160/TIR ratios is to state what the approximate spatial scales that the lensed image resolves. If the resolution element of MIRI-MRS is capturing a very large part of the galaxy, it will be less likely to have resolved scales where the equilibrium dust temperature can vary substantially.

Small notes on phrasing given the suggestions above:
- On page 3, right column, the variation in 3.3/160 is described as a "deficit of PAH emission" - given the discussion above it is not clear yet that there is a deficit of PAH emission.
- Also on page 3, right "it is clear that the 3.3 PAH feature is a poor tracer of the total IR luminosity" - I do not think this is clear yet and it may in fact be that 160 is a poorer tracer of total IR luminosity if dust temperature varies strongly across the galaxy.

Minor points:

- page 1, paragraph 1 "… are sensitive diagnostics of the formation, growth, and destruction of larger millimeter-size dust particles" - I do not think this is true. It has not yet been clearly established how PAHs form, even at low redshift. So it is not at all clear that they can be considered a diagnostic of larger grains, let alone a sensitive one. I recommend removing or rephrasing this claim.

- page 2, right column - the suggestion that the continuum emission is from an AGN is somewhat confusing given that later in the paper that is ruled out fairly conclusively. It may be worth removing this note and forward referencing the discussion that rules out AGN as the continuum source.

Author Rebuttals to Initial Comments:

Response to Referees

We thank both referees for their careful reading of the paper and constructive reports.

Referee 1

Referee 1:
1. I think they could mention in a little more detail the comparison to the next highest z source for which PAH emission has been detect, and that is GN20 (reference 7). I believe in that case star formation was also favored for dust heating?

Reply:
We agree that those results should have been mentioned in more detail, and we now briefly discuss this source in the main text (paragraph beginning "On global scales…"). That source was found to be an obscured AGN based on the low 6.2um equivalent width, but the high opacity results in the AGN not being bolometrically dominant over the full IR wavelength range.

Referee 1:
2. Figure 2 caption should mention that the rest frame 160um image (2C) is from ALMA, I believe.

Reply:
Done.

Referee 1:
3. Figure 2: what is the green dot at the center? is that the lensing galaxy? if so, can that be used to derive the stellar mass of the lens, and hence mass to light ratio for the lens, or has that been done in previous work?

Reply:
Correct, the lens is visible in the continuum image, but not in the (continuum-subtracted) PAH image, as expected. The Fig. 2 caption now mentions that the lens is visible in this panel. This galaxy has now been observed at many wavelengths from the UV to mid-IR, so yes, its stellar mass could be robustly estimated with standard techniques. The TEMPLATES team has not yet done this using the new JWST imaging, but all the data is public to anyone who wishes to do so.

Referee 1:
4. Figure 3: Can they state more explicitly the physical size scales implied by the variations on the PAH-FIR-star spatial variations?

Reply:
This is somewhat complicated because some regions of the galaxy are more highly magnified than others. Given the 0.65" resolution of the data and a flux-weighted average magnification ~30x, we expect the data to probe ~800pc scales in the source (scaling as 1/sqrt(mu)). We now mention this typical size scale in the main text (paragraph beginning "As the first space-based…").

Referee 1:
5. Figure 3: is the variation of the PAH/FIR ratio due to a change in dust grain size distribution, or could it be due to a variation in dust temperature?

Reply:
We have no way to infer anything about the grain size distribution, unfortunately. This would require other PAH features – especially the longest-wavelength features at >10um – but these features are redshifted out of the JWST wavelength coverage. Changes in the size distribution are one possible explanation for our results, which we mention in the main text (paragraph beginning "The origin of the spatial variations…").

Variations in dust temperature are plausible, and we attempted to quantify their magnitude in the Methods based on the full sample of SPT-selected galaxies, all of which have very good 100-3000um photometric coverage. We found uncertainties on L_{IR} of a factor of ~1.8 due to the distribution of L_{IR}/F_{160} that results from the range in dust temperature across the sample. See also our reply to Referee 2 below, where we examine this assumption on resolved scales within SPT0418-47.

Referee 1:
6. A few sentences would be useful that consider the possibility for future studies of PAH emission at z > 6. Will this become routine with JWST? The last sentence alludes to this, but is very strong lensing required for this? I guess the case of GN20 argues not, but that is also a hyper-luminous IR galaxy and hence rare. Are other PAH features observable with JWST (eg. the 6.2um feature)? perhaps this information may be too much for the paper, but some information in the supplemental material would be good for posterity.

Reply:
Of course we cannot predict what future observations will be led by the community, but we certainly hope that this work motivates additional PAH studies! These observations required only ~1 hour on-source, and there are hundreds of lensed and unlensed IR-luminous galaxies known, so there should be plenty of potential targets. The 6.2um feature is observable until z~3.5, though sensitivity concerns may limit its detectability to z<1.9 to avoid the longest-wavelength MIRI channel. We feel it is not necessarily our place to speculate on these future observations here, but we point out that our observing design and results will now be publicly available to other groups to guide their observations.

Referee 2

Referee 2:
The authors spend a fair amount of discussion on the variable 3.3/160 ratios (interpreted as L_PAH/L_TIR), arguing that they may be evidence for spatially distinct distributions of PAHs and other dust components. However, as they note in the Methods, this implicitly assumes that the dust temperature is the same everywhere. PAHs are generally in the stochastic heating

regime, and so approximately linearly trace the radiation field intensity. The 160 micron emission is non-linearly dependent on the radiation field intensity. Figure 15 of Draine & Li (2007) has a very clear illustration of this.

The baseline assumption in the paper appears to be that the PAHs are failing as a TIR tracer, but it seems equally possible that there are dust temperature variations and the 160/TIR ratio is variable. To make the case that this is not simply variations of the equilibrium dust temperature, it seems important to make an argument about how large of variations are possible in 3.3/160 for reasonable radiation field intensities. This galaxy has a very large star formation rate, and it isn't clear in the text if that is distributed over a wide area or concentrated in a central starburst (or if that is known). If concentrated in a starburst, the dust temperature could vary substantially over the galaxy, leading to variations in 160/TIR. Over some ranges of radiation field intensity, it is possible that the PAH emission could be a more reliable tracer of TIR, revealing where a simple conversion of 160 to TIR is failing due to changes in the equilibrium dust temperature.

Reply:
We agree that this is the primary limitation of our current study. We attempted to address this issue in the Methods by looking at the distribution of F160/TIR across the full sample of SPT galaxies, but acknowledge this does not fully satisfy the referee's question since this comparison uses only global unresolved photometry. The TEMPLATES collaboration hopes to address this point in more detail in future work using all the available ALMA data for these sources. This analysis is more subtle than it sounds because the resolved data must be fit jointly with the unresolved photometry (which better samples the SED peak) and the pixels in the ALMA maps are correlated over the synthesized beam sizes.

Because we agree that this is the main limitation of our analysis, we have elected to add additional tests of the impact of changes in dust temperature on the conversion to TIR to the Methods (section "ALMA observations and analysis"). Briefly, we use ancillary ALMA rest-frame 120um data at matched resolution to estimate the variations in dust temperature across the source, mapping the 120/160 flux ratio to an approximate dust temperature (in similar fashion to the old IRAS color techniques).

We find remarkably (somewhat surprisingly) small variations in the 120/160 ratio map, implying small changes in dust temperature and only minor impacts to our TIR-F160 conversion. Importantly, this is even true in the brightest regions of the source where the S/N is high (and perhaps one would expect e.g. hot spots in the dust emission). Preliminary investigation suggests that some of this homogeneity may be due to the relatively low spatial resolution of the data – dust temperature variations may arise primarily on spatial scales below the 0.65" resolution of the MRS data we use. But because this analysis is ongoing, we prefer not to speculate too deeply here. In any case it is apparent that, at least at the spatial scales we are probing, variations in dust temperature / radiation field intensity within the galaxy are not the dominant source of uncertainty in our analysis.

Referee 2:
- One point that would help in understanding the 160/TIR ratios is to state what the approximate spatial scales that the lensed image resolves. If the resolution element of MIRI-MRS is capturing

a very large part of the galaxy, it will be less likely to have resolved scales where the equilibrium dust temperature can vary substantially.

Reply:
We agree; this was also suggested by Referee 1 and we now mention the approximate physical scale we expect to be probing here.

Referee 2:
Small notes on phrasing given the suggestions above:
- On page 3, right column, the variation in 3.3/160 is described as a "deficit of PAH emission" - given the discussion above it is not clear yet that there is a deficit of PAH emission.
- Also on page 3, right "it is clear that the 3.3 PAH feature is a poor tracer of the total IR luminosity" - I do not think this is clear yet and it may in fact be that 160 is a poorer tracer of total IR luminosity if dust temperature varies strongly across the galaxy.

Reply:
We hope that, given the tests described above now included in the Methods, that the referee finds these statements justifiable.

Referee:
Minor points:
- page 1, paragraph 1 "... are sensitive diagnostics of the formation, growth, and destruction of larger millimeter-size dust particles" - I do not think this is true. It has not yet been clearly established how PAHs form, even at low redshift. So it is not at all clear that they can be considered a diagnostic of larger grains, let alone a sensitive one. I recommend removing or rephrasing this claim.

Reply:
Agreed, rephrased.

Referee:
- page 2, right column - the suggestion that the continuum emission is from an AGN is somewhat confusing given that later in the paper that is ruled out fairly conclusively. It may be worth removing this note and forward referencing the discussion that rules out AGN as the continuum source.

Reply:
We apologize for the confusion. It is true that we have ruled out that an AGN dominates the continuum emission on global source-integrated scales, but cannot rule this out within smaller regions within the galaxy. We later speculate that AGN emission may dominate the continuum on scales smaller than the MRS resolution (main text paragraph beginning "On global scales…") since we do see evidence of lower PAH equivalent width in some regions. We prefer to keep our original wording here.