Dear Referee,

Thank you very much for your useful comments. We have revised the manuscript taking your comments into account.

> The manuscript "Discovery of Massive, Mostly Star-Formation Quenched
> Galaxies with Extremely Strong Lyman-alpha Emission at z~3" presents
> a sample of nine massive galaxies at z~2.5-3.2 with Lyman-alpha
> equivalent width >100A and stellar mass log(M)>10.5
>
> As described below, I believe that they Lya emission in five out of
> the sample of nine galaxies can be easily explained by significant
> star formation or AGN activity. This reduces the sample of noteworthy
> candidates to four.
>
> The result that some quiescent-looking galaxies have strong Lya
> emission, but no hint of significant AGN activity, is somewhat curious
> and may be worthy of publication. However there are many aspects of
> this analysis that need to be greatly extended and clarified. There is
> also good reason to believe that the Lya size measurements, upon which
> much of the physical discussion in this letter are based, are
> completely unreliable. For all of these reasons I do not think this
> paper can be published in the ApJL in its current form.

First of all, we agree with the referee's opinion that the MAESTLOs detected with the Chandra X-ray Observatory can be regarded as AGNs. Before the submission of this paper, only a part of Chandra data are available for the COSMOS field. Therefore, we kept the two Chandra sources in the MAESTLO sample. However, now, we can use the full Chandra data for the entire COSMOS field. We have made cross-matching analysis and found that one more MAESTLO (previous ID, pID, of 3) is a Chandra source. Accordingly, we have removed the three Chandra sources (pID 1, 2, and 3) from the MAESTLO sample in the revised manuscript. Therefore, we can discuss observational properties of the six MAESTLOs without any evidence for AGN activity, unambiguously.

You also raised questions on the 3 MAESTLOs with pID 5, 8, and 9. As for these objects, we give our answers below.

To avoid confusion, we show the following list.

```
================================
Previous ID (pID)      New ID
        1              removed
        2              removed
        3              removed
        4                 1
        5                 2
        6                 3
        7                 4
        8                 5
        9                 6
================================
```

Sample selection:
=================

> Of the nine galaxies in this sample, two (ID numbers 1 and 2) are
> detected in the Chandra data. These galaxies can be classified as
> AGN, since star formation cannot cause an x-ray detection at these
> redshifts. These galaxies also have high star formation rates as
> estimated from the SED fitting; hence it is not at all surprising
> that they have Lya emission.

As mentioned above, we have removed the three Chandra source, i.e., pID 1, 2, and 3 from our sample. We then concentrate on the well-defined sample of galaxies without any evidence for AGN. Therefore, the observed Ly alpha emission in the current six MAESTLOs is related either to star formation or to shock heated gas driven by a superwind or to both.



> There is also a third galaxy (ID=9) that looks as if it may have a
> power-law SED, which would strongly suggest AGN activity. The
> authors claim that they used the Donley et al. (2012) criteria to
> select AGNs from the IRAC colors, but since this object is not
> detected significantly at 8um, perhaps the authors didn't apply any
> IRAC AGN selection criteria to this object?

Thank you for your careful check. As you pointed out, we did not apply the Donley+12 criteria for this MAESTLO pID 9 because of its non-detection in the IRAC Ch4. We have examined whether or not it satisfies the Donley+12 criteria if it may have the reddest color of Ch2-Ch4 by adopting the 3-sigma upper limit for Ch4. We have found that this object is located at nearly the edge of the criteria; see the figure uploaded as Supplemental Material. Since the rest-frame UV and optical SED appears to be well explained solely by the stellar radiation (i.e., little [basically no] evidence for non-thermal continuum from an AGN, it is reasonable to keep this object as a MAESTLO without any evidence for AGN. Of course, it is possible that a very weak AGN is present in its nucleus. However, such an object does not affect on our discussion given in this Letter.

> Of the remaining galaxies, two (ID numbers 5 and 8) have
> log(SFR)=1.8-2, so it is very difficult to describe these galaxies as
> "mostly quenched." In fact they are not all that far below the main
> sequence of star formation at these redshifts (see e.g. Fig. 10 of
> Schreiber et al 2010, and note that they use a Salpeter IMF as opposed
> to Chabrier as used in this work. Also, recall that the observational
> scatter that is typically quoted for the main sequence is 0.2-0.3
> dex.). Finding Lya emission in some fraction of typical, massive
> star-forming galaxies is not surprising and is not worthy of
> publication (even if the galaxies are rather red, Lya emission can be
> explained by radiative transfer effects; e.g. Finkelstein et
> al. 2009).

Thank you for this comment. Actually, the MAESTLOs, pID 5 and 8 have a large SFR and appear to be close to the domain of the so-called main sequence of star-forming galaxies at similar redshifts; note that their new IDs are 2 and 5 in the revised manuscript. However, since the MAESTLO pID 5 (i.e., new ID 2) has a ratio of age/tau larger than unity and shows a clear Balmer break, it is reasonable to consider this galaxy is in an quenching phase of star formation. On the other hand, as written explicitly in the footnote 16 in the manuscript, the MAESTLO pID 8 (i.e., new ID 5) shows a relatively weak Balmer/4000 A break compared to the other MAESTLOs and its tau is estimated to be highly uncertain. Therefore, to avoid any confusion, we have mentioned this explicitly in the revised manuscript. We also note that we do not include these galaxies for the discussion on the duration of the MAESTLO phase although this does not affect the discussion (i.e., only a factor of 4/6).

> This leaves a sample of four galaxies that are plausibly quenched,
> have no evidence of significant AGN activity, and have strong Lya
> emission (the sample increases to five if galaxy number 9 can be shown
> not to have IRAC colors consistent with an AGN). Such a very rare
> sample may be of marginal interest in itself, but it is impossible to
> put them into an evolutionary scenario.

Following your comment given in âM−^@M−^\Other commentsâM−^@M−^], we have investigated rest-frame UV-Opt-NIR colors of the six MAESTLOs comparing some star formation models. In this analysis, we have found that the four MAESTLOs with a smaller sSFR are located around the selection boundary for the quiescent galaxies in the rest-frame U-V vs. V-J diagram studied by Muzzin et al. (2013, ApJ, 777, 18). Therefore, we have confirmed that these four MAESTLOs are in a quenching phase. Thank you very much for your very useful suggestion.

We believe that the discovery of such quenching, massive galaxies at z ~ 3 is very important because such populations have not yet been directly detected to date. It is interesting to note that such quenching phase is manifested by the observed large EW_0(Lya) in the MAESTLOs. If this is the case, EW_0(Lya) will provide us a tool to identify the quenching phase in future observational studies.

Other comments:
=================

> -- Title: since the authors don't quote Lya fluxes (they only quote equivalent widths), I don't think it is justified to say that these galaxies have "extremely strong Lyman-alpha emission." In fact, the authors have not shown that the Lya emission is very strong; the EW is high, but that may only be because the rest-frame UV continuum is very faint. Indeed, the SFR estimates from Lya are quite low. For this reason I don't think it's fair to say that these galaxies have "extreme" Lya emission. I would suggest that they quote Lya fluxes, and compare them to the Lya luminosity function (e.g. Gronwall et al. 2007), if they wish to claim that the Lya emission is "extreme." Alternatively, a more accurate title would refer to "large Lyman-alpha equivalent widths"

Let us note that the term of "strong emission line" refers not to "large line flux" but to "large EW" usually. For "large line flux", we usually use "bright emission line". As you see in the original manuscript, we do not discuss Ly alpha luminosity of our sample galaxies explicitly.

However, in order to avoid any confusion to readers, we have changed the paper title to

"Discovery of Massive, Mostly Star-formation Quenched Galaxies with Extremely Large Lyman-alpha Equivalent Widths at z ~ 3".

> -- Also, regarding the title, at least four of the galaxies have significant, active star formation based on the SED fits. And, as explained above, the star formation rates are not strongly depressed compared to the main sequence. If these galaxies were in the local universe they would be called extreme starbursts. So it is not accurate to describe them as "mostly quenched."

Thank you for this comment. Now in the revised manuscript, we have 6 MAESTLOs; four galaxies have smaller sSFRs while the remaining two have typical SFR for galaxies at a concerned redshift. This means that the fraction of small sSFR galaxies is 4/6 = 67%. Therefore it seems reasonable to mention that the majority of MAESTLOs have a significantly smaller specific star formation rate, or one can conclude that they are "mostly quenched".

As mentioned already above, we have found that the two galaxies, pID 5 and 8 (their new IDs are 2 and 5 in the revised manuscript) can be interpreted as that the star formation in them ceased recently and started to experience their passive evolution. Although their observed SFRs are comparable to those in star-forming galaxies at the



similar redshift range (i.e., z ~ 3), we can give consistent interpretation on the all six galaxies in terms of the evolution of SF quenching.

> -- The authors discuss two possibilities for the cause of the Lya emission. One is winds that are blown out by AGN or supernovae, and the other is low-level star formation activity. The quoted conclusion that the galaxies may represent a brief transition phase between star-forming and quiescence is completely unsupported. If the Lya emission is caused by low level star formation, then it could suggest that the galaxies are not 100% quenched. It may also suggest that the galaxies are currently undergoing a minor merger with a gas rich dwarf, that the Lya emission is related to the mysterious Lya blobs, or that it is due to a projection effect with a nearby source; none of these possibilities are discussed by the authors, and neither can these possibilities be easily placed into an evolutionary scenario. Additionally, if the Lya emission is caused by AGN activity, then neither does that give any believable evidence that the galaxies are in a brief transition phase.

First, we have removed the discussion on the AGN winds/jets because we have rejected all AGN-related objects in this revised manuscript.

Since the rest-frame UV, optical NIR color analysis suggested by you have reinforced that the four MAESTLOs have quite similar colors of quiescent galaxies, we can conclude that the most MAESTLOs are in the final (quenching) phase of active star formation.

We would like to note that the presence of LABs at high redshift is not mysterious because a large number of LABs have been already found to date (e.g., Matsuda+04, AJ, 128, 569, Matsuda+11, MN, 410, L10, Saito+06, ApJ, 648, 54) and it is found that they have massive stellar components (Uchimoto et al. 2012, ApJ, 750, 116, Kubo et al. 2015, ApJ, 799, 38). Therefore, it is not plausible that an infalling gas rich dwarf causes Lya blob.

As for the projection effect and so on, we have newly added the following comments in the revised manuscript. We think that these comments provide enough information to readers.

"As for possible origins of the observed extended Lya emission, we may have a projection effect. However, we consider that such a projection effect cannot be the origin of most MAESTLOs because of the small number densities of both massive galaxies with relatively low sSFR and LAEs with extremely large EWs at z ~ 3 (e.g., ~ 1.3x10^{-5} Mpc^{-3} for galaxies with M_\star >= 10^{10.5} Msun and sSFR = 0.03--0.3 Gyr^{-1}, and ~ 5.4x10^{-6} Mpc^{-3} for LAEs with EW_0(Lya) >= 100 A). Namely, the projection probability of these objects with similar redshifts (e.g., Delta z < 0.5$) is expected to be extremely low (~ 0.003 such chance alignments within 1" in our survey volume). Another possibility is that a star-forming dwarf galaxy is going to merge onto these MAESTLOs. In this case, we have to explain why a merging dwarf galaxy experiences such active star formation. Finally, jets and/or outflows from AGN are also possible origins of such extended emission-line regions. However, our MAESTLO sample does not show any evidence for AGNs (Section~2), although weak AGN possibility is not completely rejected because of the limited depths of IRAC, XMM, Chandra, and radio data."

> -- Regarding the AGN activity, I found it very odd that the authors discuss how they removed AGN using radio/x-ray/IR indicators in Section 2, but then in Section 3 they explain that two of the sources are detected in the x-ray. An x-ray detection at z~3

> certainly means that the galaxies are AGN; it would take an extreme
> level of star formation to lead to an x-ray detection, and surely
> such star-formation can be ruled out by the SEDs and the lack of a
> 24um detection. It would make sense to remove all AGNs from the
> sample, or remove none of them. However if the AGNs are kept in the
> sample then the authors would have to clearly specify which galaxies
> are AGN in the table and figures.

We have already fixed this issue. Thank you very much for this comment again.

> -- Why have the authors not tried stacking the Chandra data in order
> to see if there is a weak detection?

As for the stacking analysis for the six MAESTLOs that are not detected in Chandra, we do not think that this works well because the number of sources is only 6. Such stacking analysis will work for much larger non-detection samples.

> -- If the galaxies do indeed represent a transition phase between
> star formation and quiescence, then there should be evidence for
> that in the SED fits and the colors. Do these galaxies have younger
> ages, or larger tau/age, then typical quenched galaxies at similar
> mass? However, given the very well-known degeneracies and
> systematic uncertainties in stellar population properties inferred
> from SED-fitting, it is strange that the authors don't at least
> mention this issue. The SFRs, in particular, are completely
> unreliable for "quenched" galaxies. So to complement the discussion
> of the SED fitting, I would also strongly suggest including a U-V
> vs. V-J color-color diagram, as this is a useful visual diagnostic
> of galaxy's colors, dust content, and age (e.g. Williams et
> al. 2009). In this diagram there is a clear separation between
> star-forming and quenched galaxies, and recently-quenched galaxies
> would be found at the lower-left edge of the sequence of quenched
> galaxies.

Thank you very much for this comment. Following your suggestion, we have examined the distribution of the MAESTLOs in the rest-frame two colors plane, U-V vs. V-J, which have been evaluated from the best-fit model SEDs. The result is shown in new Fig. 4. Comparing the quiescent galaxies studied by Muzzin+13 (ApJ, 777, 18), we have indeed found that our four MAESTLOs with a low sSFR (new IDs 1, 3, 4, and 6) are located around the selection boundary for quiescent galaxies. It is thus suggested that they have been recently quenched and are moving into the passive evolution phase. Therefore, this color analysis reinforces our scenario. Although the colors of the other two MAESTLOs are consistent with the star-forming models, their colors can also be interpreted as a galaxy that ceased its star formation recently. The larger dust contents in these two galaxies may be expected if they are in an early phase of the superwind activity; i.e., most of dust grains may have not yet been blown out by the superwind.

> -- If there is significant Lya emission, there may also be other lines
> with significant emission (Ha, OII, OIII). There is the potential
> that these lines are affecting the SED fits. Do the SED fitting
> results change at all if the filters that are potentially affected
> by these lines are removed from the fits?

As you concerned, some MAESTLOs may also have other emission lines; these lines may be detected in JHK bands. However, since the effect of such emission lines is not expected to be serious because of the very wide filter bandpasses in JHK, we use all JHK-band data even if



some strong emission lines are expected to be in these bands. Although we have considered such case carefully in the sample selection procedures, we have not described explicitly in the previous manuscript. We have now described this treatment explicitly in Section 2 of the revised manuscript.

> -- The size measurements in the IA bands are potentially quite very
> important, as they may help to shed light on the mechanism
> responsible for the Lya emission. For this reason I think that it
> is important to take great care in the measurements, in order to
> know whether the ground-based measurements are reliable. Instead, I
> find the quoted measurements very problematic. For instance, for
> galaxies 6 and 2 the ground-based i*-band size measurements
> disagree *very strongly* with the F814 measurements. This leads me
> to believe that ground-based size measurements are completely
> unreliable for small objects, and that the quoted uncertainties are
> grossly underestimated. It is only for a single galaxy (ID=1) that
> there is any reason to believe that the ground-based measurements
> are reliable. But note that this galaxy is both quite bright and
> very extended; hence this may be a special case. If we can't trust
> the i*-band size measurements, why should we trust the IA-band
> measurements? The authors would have to do careful tests to show
> that the measurements and the uncertainties are reliable and
> unbiased. There are a number of examples in the literature of such
> test; for just one particularly recent example, see Straatman et
> al. (2015).

Following your suggestion, we carried out the Monte Carlo simulation to estimate the uncertainty in the size measurements. We put the best-fit model profile at 200 random positions around the original position (in a 2' x 2' region) per observed MAESTLO and re-measured their sizes. The standard deviation of these 200 measurements is adopted as the uncertainty, where the systematic effects such as the background fluctuation are taken into account. The estimated uncertainty is systematically larger than the nominal errors from GALFIT as expected. We replace all the size errors in Table 1 with those estimated from the simulation. As a result, we found that all measurements in the Subaru i'-band data including old Nos. 2 and 6 cannot be distinguished from the point sources, which is consistent with the very small sizes measured in the HST F814W band. On the other hand, three out of new 6 MAESTLOs are significantly extended in the Lya (intermediate-band) images considering the uncertainty estimated from the simulation, while the other three cannot be distinguished from the point sources. Therefore, the continuum emission from all the MAESTLOs is very compact, while a half of 6 MAESTLOs show the extended Lya emission. We have added the description of the simulation to estimate the uncertainty in the size measurements in Section 2, and describe these results including the new error estimates in Section 3 in the revised manuscript.

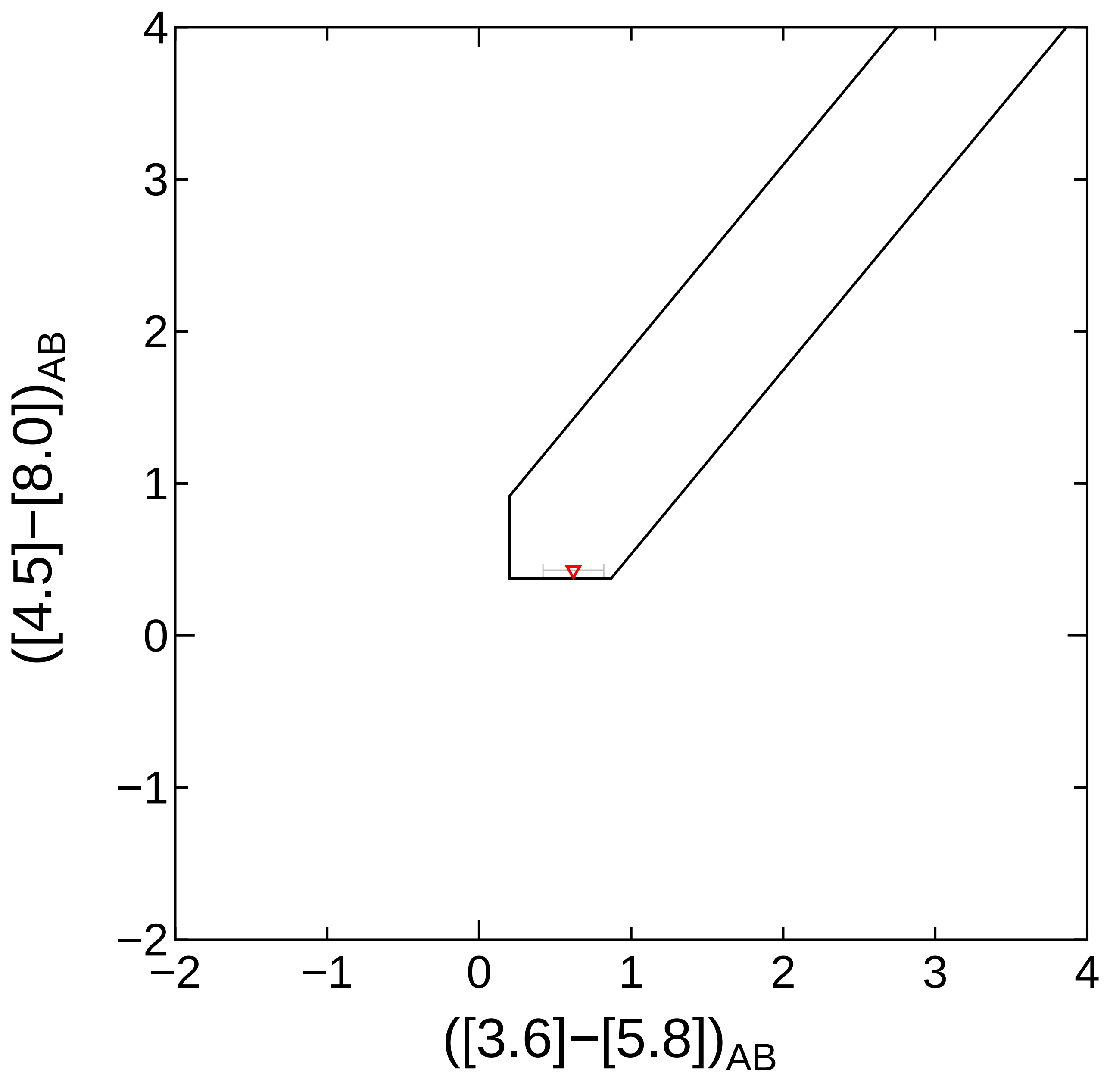
> -- The authors should state the FWHM of the IA and i* images.

We describe these PSF sizes in the form of the half light radius, which is used in the size measurements, in Section 2.

> -- At the end of the manuscript, the authors that that future studies
> will allow us to constrain the mechanisms that quench galaxies. I
> find this statement completely unsupported. Even if the MAESTLOs
> were transition objects, I am not sure how finding more such
> objects, or obtaining IFU observations, would shed any light. If
> the authors can specifically point out how studying MAESTLOs will
> clearly reveal the quenching mechanism, then they should explain
> it.

For example, observations to probe gas kinematics using ALMA will provide a useful constraint on the mechanism that quench star formation in galaxies. If outflowing gas is detected, the quenching mechanism is found to be strong stellar and/or supernova feedback. On the other hand, if gas is not detected at all, the mechanism should be strangulation, that is, the depletion of gas. We have added them at the end of the revised manuscript.

That’s all. Thank you very much.



4
3
2
1
0
−1
−2
−2
−1
0
1
2
3
4
([4.5]−[8.0])AB
([3.6]−[5.8])AB