



On the Connection between Supermassive Black Holes and Galaxy Growth in the Reionization Epoch

Junyao Li^{1,2,3}, John D. Silverman^{2,4}, Takuma Izumi^{5,6}, Wanqiu He⁵, Masayuki Akiyama⁷, Kohei Inayoshi⁸, Yoshiki Matsuoka⁹, Masafusa Onoue^{2,8}, and Yoshiki Toba^{9,10,11}

¹ CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei 230026, People's Republic of China; lijunyao@mail.ustc.edu.cn

² Kavli Institute for the Physics and Mathematics of the Universe (WPI), The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

³ School of Astronomy and Space Science, University of Science and Technology of China, Hefei 230026, People's Republic of China

⁴ Department of Astronomy, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan

⁵ National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

⁶ Department of Astronomical Science, Graduate University for Advanced Studies (SOKENDAI), 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

⁷ Astronomical Institute, Tohoku University, Aramaki, Aoba, Sendai 980-8578, Japan

⁸ Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People's Republic of China

⁹ Research Center for Space and Cosmic Evolution, Ehime University, 2-5 Bunkyo-cho, Matsuyama, Ehime 790-8577, Japan

¹⁰ Department of Astronomy, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan

¹¹ Academia Sinica Institute of Astronomy and Astrophysics, 11F of Astronomy-Mathematics Building, AS/NTU, No.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan

Received 2022 March 17; revised 2022 April 27; accepted 2022 May 8; published 2022 May 23

Abstract

The correlation between the mass of supermassive black holes (SMBHs; \mathcal{M}_{BH}) and their host galaxies (\mathcal{M}_*) in the reionization epoch provides valuable constraints on their early growth. High-redshift quasars typically have an $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ ratio significantly elevated in comparison to the local value. However, the degree to which this apparent offset is driven by observational biases is unclear for the most distant quasars. To address this issue, we model the sample selection and measurement biases for a compilation of 20 quasars at $z \sim 6$ with host properties based on ALMA observations. We find that the observed distribution of quasars in the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ plane can be reproduced by assuming that the underlying SMBH population at $z \sim 6$ follows the relationship in the local universe. However, a positive or even a negative evolution in $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ can also explain the data, depending on whether the intrinsic scatter evolves and on the strength of various systematic uncertainties. To break these degeneracies, an improvement in the accuracy of mass measurements and an expansion of the current sample to lower \mathcal{M}_{BH} limits are needed. Furthermore, assuming a radiative efficiency of 0.1 and quasar duty cycles estimated from the active SMBH fraction, significant outliers in $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ tend to move toward the local relation given their instantaneous BH growth rate and star formation rate. This may provide evidence for a self-regulated SMBH–galaxy coevolution scenario that is in place at $z \sim 6$, with active galactic nucleus feedback being a possible driver.

Unified Astronomy Thesaurus concepts: Quasars (1319); Supermassive black holes (1663); Galaxy evolution (594); AGN host galaxies (2017); Active galactic nuclei (16)

1. Introduction

Active galactic nuclei (AGNs) or quasars, powered by mass accretion onto supermassive black holes (SMBHs), produce an enormous amount of energy that has long been speculated to have profound impacts on galaxy evolution (e.g., King & Pounds 2015). In the local universe, the mass of SMBHs appears to be closely connected with the bulge properties (e.g., bulge mass $\mathcal{M}_{\text{bulge}}$, stellar velocity dispersion σ_*), which inspired the concept of “coevolution” in studies of SMBH and galaxy evolution (e.g., Kormendy & Ho 2013).

High-redshift studies have mainly focused on the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ relation, with mounting evidence showing that its evolution in massive systems has not been significant since $z \sim 2$ (e.g., Jahnke et al. 2009; Schramm & Silverman 2013; Sun et al. 2015; Ding et al. 2020; Li et al. 2021). In particular, its intrinsic scatter appears similar to the local value (e.g., Ding et al. 2020; Li et al. 2021). These results suggest that a physical

coupling between SMBHs and galaxies (e.g., through AGN feedback) is likely at work to keep $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ relatively constant. To decipher how the relationship was first established in the early universe, a key strategy would be to measure the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ relation in the reionization era ($z > 6$) where we are able to probe the first generation of accreting SMBHs (e.g., Inayoshi et al. 2020).

Many of the $z \sim 6$ quasars discovered so far are powered by extremely massive BHs with $\mathcal{M}_{\text{BH}} \sim 10^9 M_{\odot}$ (e.g., Fan et al. 2000; Shen et al. 2019) and are actively forming stars with star formation rates (SFRs) $\sim 100-1000 M_{\odot} \text{ yr}^{-1}$ (e.g., Wang et al. 2013). Their $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ (where \mathcal{M}_* is approximated by the dynamical mass \mathcal{M}_{dyn} measured from gas kinematics using ALMA) appears to be significantly offset from the local value by up to 2.0 dex, suggesting that the growth of SMBHs substantially precedes their hosts (e.g., Neeleman et al. 2021). However, these quasars are biased tracers of the underlying SMBH population because only the most-luminous quasars powered by the most massive BHs can be detected in shallow surveys (e.g., Lauer et al. 2007; Volonteri & Stark 2011; Schulze & Wisotzki 2014). Lower-luminosity quasars detected in deeper surveys (e.g., the SHELLQs survey; Matsuoka et al. 2016)

lie closer to the local relation, thus confirming this bias (e.g., Izumi et al. 2019, 2021).

Moreover, the mass measurements at high redshifts suffer from significant uncertainties with possibly systematic biases. For instance, the \mathcal{M}_{BH} of a flux-limited quasar sample might be statistically overestimated by the single-epoch virial estimator (e.g., Vestergaard & Osmer 2009, hereafter VO09) because of an uncorrelated variation between AGN luminosity and broad-line width (especially for Mg II and C IV). This gives rise to a luminosity-dependent virial BH mass bias (hereafter the SE bias) as detailed in Shen (2013). In addition, the hosts of $z \sim 6$ quasars are found to be gas rich (e.g., Decarli et al. 2022), thus approximating \mathcal{M}_* by \mathcal{M}_{dyn} is likely an overestimate.

In this Letter, we model the selection and measurement biases for a $z \sim 6$ quasar sample compiled in the literature in order to reveal the underlying connection between SMBH and galaxy growth in the early universe. This work assumes a cosmological model with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The stellar mass and SFR are based on the Chabrier (2003) initial mass function.

2. Sample

We adopt the $z \sim 6$ quasar sample compiled in Izumi et al. (2021) as our parent sample. It contains 46 quasars with \mathcal{M}_{BH} , \mathcal{M}_{dyn} , quasar luminosity (L_{3000} and M_{1450}), and infrared (IR) luminosity measurements from the literature. The \mathcal{M}_{BH} of 22/46 objects are derived from the virial estimator using the VO09 calibration for the Mg II line as

$$\log\left(\frac{\mathcal{M}_{\text{BH}}}{M_\odot}\right) = 6.86 + 0.5 \log\left(\frac{\lambda L_{\lambda 3000}}{10^{44} \text{ erg s}^{-1}}\right) + 2 \log\left(\frac{\text{FWHM}_{\text{Mg II}}}{\text{km s}^{-1}}\right), \quad (1)$$

while Eddington-limited accretion is assumed to estimate the mass for the remaining 24 objects.

The total IR luminosity (L_{TIR}) is derived by fitting the 1.2 mm ALMA continuum with an optically thin graybody spectrum assuming a dust temperature of 47 K and a dust spectral emissivity index of 1.6 that have been regularly adopted for $z \sim 6$ quasars, then extrapolating to the total IR (8–1000 μm) range; although the actual dust temperature could vary from source to source (e.g., Venemans et al. 2016). The SFR is derived using $3.88 \times 10^{-44} L_{\text{TIR}}$ (Murphy et al. 2011), assuming that the cold interstellar medium is predominantly heated by star formation.

The dynamical masses \mathcal{M}_{dyn} of these quasars are derived through gas kinematics using the [C II] line. The standard rotating thin-disk approximation is assumed for all quasars except two (given as upper limits). In this work, we only consider the 20 objects whose \mathcal{M}_{BH} and \mathcal{M}_{dyn} are derived from the Mg II line and the rotating thin-disk assumption, respectively, to ensure relatively reliable mass measurements. However, we caution that the derived \mathcal{M}_{dyn} is highly sensitive to the assumptions made on galaxy geometry and inclination angle (e.g., Wang et al. 2013).

Ignoring the possibly small contribution of dark matter within the [C II]-emitting region (e.g., Genzel et al. 2017), we estimate the \mathcal{M}_* of these quasars by subtracting the expected molecular gas mass (\mathcal{M}_{gas}) from their total \mathcal{M}_{dyn} , assuming that quasar hosts have similar gas content to star-forming galaxies (e.g., Molina et al. 2021). We adopt the typical

$\mathcal{M}_{\text{gas}}/\mathcal{M}_*$ ratio (μ_{gas}) of $z \sim 6$ galaxies given by Tacconi et al. (2018):

$$\log \mu_{\text{gas}} = 0.12 - 3.62 \times (\log(1+z) - 0.66)^2 - 0.35 \times (\log \mathcal{M}_* - 10.7),$$

where we adopted their $\beta=2$ result with the Speagle et al. (2014) star formation main sequence (MS) and assumed $\delta\text{MS} = 0$. We then use this relationship to estimate the typical \mathcal{M}_{gas} at a given \mathcal{M}_* and derive the correlation between \mathcal{M}_* and \mathcal{M}_{dyn} , where \mathcal{M}_{dyn} is approximated by $\mathcal{M}_* + \mathcal{M}_{\text{gas}}$. The resulting \mathcal{M}_{gas} and \mathcal{M}_* of our quasars at their respective \mathcal{M}_{dyn} are shown in Figure 1 (left panel). The gas mass is distributed between $\sim 10^{10}$ – $10^{11} M_\odot$, which is consistent with recent direct measurements in $z \sim 6$ quasars (e.g., Decarli et al. 2022). The derived \mathcal{M}_* is typically 0.2–0.5 dex smaller than \mathcal{M}_{dyn} . We also show the distribution of our quasars in the SFR– \mathcal{M}_* plane in Figure 1 (right panel). As can be seen, most quasars are located near the Speagle et al. (2014) MS.

3. Evolution of the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ Relation

3.1. The Observed $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ Relation

Figure 2 shows the $z \sim 6$ quasars in the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ plane compared to the local $\mathcal{M}_{\text{BH}}-\mathcal{M}_{\text{bulge}}$ relation. We adopt the local relation given by Häring & Rix (2004, hereafter HR04) to be self-consistent with the VO09 virial estimator (see Section 6.2 in Ding et al. 2020 for the discussion on the choice of the local baseline). The local sample consists of massive ellipticals and bulge-dominated S0 galaxies, thus we adopt $\mathcal{M}_{\text{bulge}} \approx \mathcal{M}_*$ in the following analyses. This figure confirms that quasars at $z \sim 6$ typically lie above the local relation, and the offset at a given stellar mass ($\Delta \log \mathcal{M}_{\text{BH}}$) generally increases with quasar luminosity.

3.2. Estimating Expected Biases

As introduced in Section 1, the current $z \sim 6$ quasar sample suffers from strong selection biases. Following Li et al. (2021), we perform a simple Monte Carlo simulation to build a mock AGN sample that mimics the observational biases to account for such an effect in order to reveal the underlying connection between SMBHs and their host galaxies (e.g., Schulze & Wisotzki 2014; Volonteri & Reines 2016). In the following, we briefly introduce our simulation method. The details of each step and the choice of model parameters are described in the Appendix.

Our simulation starts with the galaxy stellar-mass function (SMF) at $z \sim 6$ given by Grazian et al. (2015) and the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ relation to generate a sample of mock galaxies and SMBHs. The $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ relation is assumed to have a Gaussian intrinsic scatter (σ_μ) with a mean that evolves as

$$\Delta \log \mathcal{M}_{\text{BH}} = \gamma \log(1+z). \quad (2)$$

Assuming that Type 1 AGNs follow the same $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ relation as the underlying galaxy population (see Schulze & Wisotzki 2014 and Li et al. 2021), we determine the bolometric luminosity by using \mathcal{M}_{BH} and adopting an intrinsic Eddington ratio (λ_{Edd}) distribution function (ERDF) of Type 1 AGNs. We use the ERDF at $z = 4.75$ given by Kelly & Shen (2013) who jointly constrained the intrinsic ERDF and the active BH mass function (BHMF) based on uniformly selected

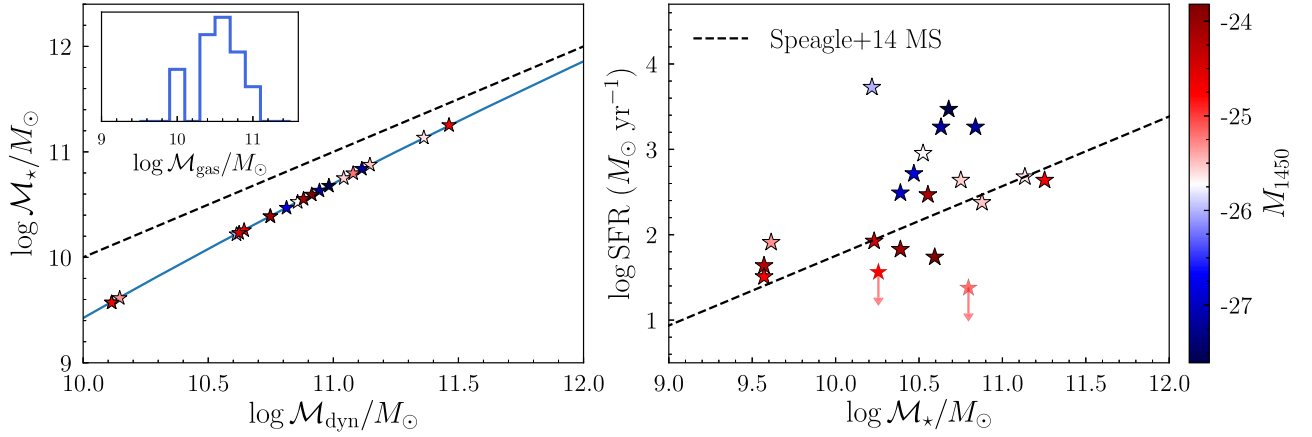


Figure 1. Left: estimating stellar masses of $z \sim 6$ quasars by subtracting the expected molecular gas masses (inset histogram) from their dynamical masses. Quasars are plotted as stars color-coded by their M_{1450} . The blue curve shows the M_* - M_{dyn} relation derived from the M_{gas}/M_* vs. M_* relation given by Tacconi et al. (2018). The black dashed line shows the one-to-one relation. Right: the SFR- M_* plane of our quasars compared to the Speagle et al. (2014) MS relation at $z \sim 6$. Two objects with upper limits in SFR are marked by red arrows.

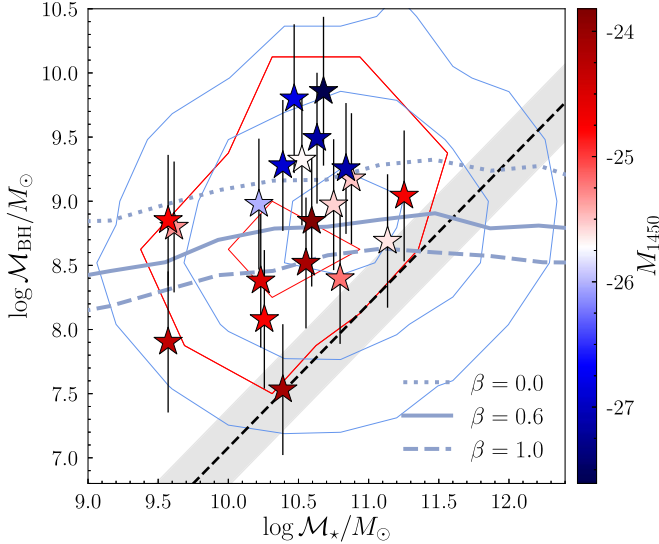


Figure 2. The $M_{\text{BH}}-M_*$ distribution at $z \sim 6$. The individual points and the red contours (1σ - 2σ levels) show the observed quasars color-coded by their M_{1450} . The blue contours (1σ - 3σ levels) show the mock AGN sample, which mimics the selection function and measurement uncertainties of the observations (i.e., Mock-Q) assuming $\beta = 0.6$. The average $M_{\text{BH}}-M_*$ relations of the mock AGN samples (incorporating observational biases) using $\beta = 0.0, 0.6$, and 1.0 are plotted as dotted, solid, and dashed blue curves, respectively. The local HR04 relation and its scatter are shown as a black dashed line and a gray shaded region.

SDSS quasars with the sample incompleteness being carefully corrected in a Bayesian framework.

We then derive a virial M_{BH} for each mock AGN using the VO09 virial estimator based on its true M_{BH} , luminosity, an assumed FWHM distribution (in a log-normal form; Shen 2013), and a parameter β ($0 \leq \beta \leq 1$) that describes the fraction of correlated response of line width to the variation of luminosity. The value of β is poorly constrained at present. We adopt $\beta = 0.6$ in this work, while $\beta \neq 1$ will give rise to the SE bias (Shen 2013). The resulting virial M_{BH} has a 0.4 dex scatter relative to the true M_{BH} , and it tends to overestimate the true M_{BH} if $L > \bar{L}$ (and vice versa), where L is the luminosity of an AGN with a true BH mass of $M_{\text{BH, true}}$, and \bar{L} is the average luminosity of all AGNs at the same $M_{\text{BH, true}}$ (see the Appendix for details). In addition, we add a random Gaussian error with a

dispersion of 0.5 dex to each true M_* to reflect the large uncertainties of estimating M_* from M_{dyn} .

In our framework, under the assumption of no evolution in the $M_{\text{BH}}-M_*$ relation (i.e., $\gamma = 0.0$, $\sigma_\mu = 0.3$) of the underlying SMBH population, one can estimate the expected bias (i.e., a positive $\Delta \log M_{\text{BH}}$ relative to the local relation) caused by the sample selection function by applying the same selection criteria of observations to mock AGNs. However, it is infeasible to define a selection function as our sample is a mixture of $z \sim 6$ quasars from various surveys with additional requirements of having near-IR spectroscopic and ALMA follow up to measure M_{BH} and M_{dyn} . Therefore, we assume a simplified scenario for which all selection biases come from the “effective” magnitude limit of different surveys, where effective means that these quasars are the relatively luminous ones selected from their parent samples for follow up observations. We simulate this selection function by producing randomly drawn samples of mock quasars that are matched to the observed M_{1450} distribution (hereafter the Mock-Q sample).

The distribution of the Mock-Q sample in the $M_{\text{BH}}-M_*$ plane is plotted as blue contours in Figure 2. Their virial M_{BH} tends to overestimate the true M_{BH} by ~ 0.25 dex (see Figure 5(e) in the Appendix). Given the magnitude limit and the large uncertainties being added to both masses, the distribution is strongly modulated compared to the originally assumed HR04 relation. In Figure 2 we also show the average virial M_{BH} in bins of M_* for the Mock-Q sample as a blue solid curve. It represents the expected offset caused by selection effects and measurement uncertainties. To rephrase, any offset and large scatter in the observed $M_{\text{BH}}-M_*$ relation (red contours in Figure 2) that follows the blue curve and contours could be considered as lacking significant evolution in the mass relation of the underlying SMBH population, which appears to be true for our quasar sample. Note that the small offset between the observed quasars and the model predictions could be due to the different methods used to derive stellar masses in Grazian et al. (2015) and this work (spectral energy distribution fitting versus $M_{\text{dyn}}-M_{\text{gas}}$, where the latter may underestimate the total stellar mass; see Section 3.3).

We also show the impact of varying β in Figure 2. For reference, the choice of an extreme value (unlikely to be true; see the Appendix for details) for β will cause systematic shifts of the average $M_{\text{BH}}-M_*$ relation for the Mock-Q sample by

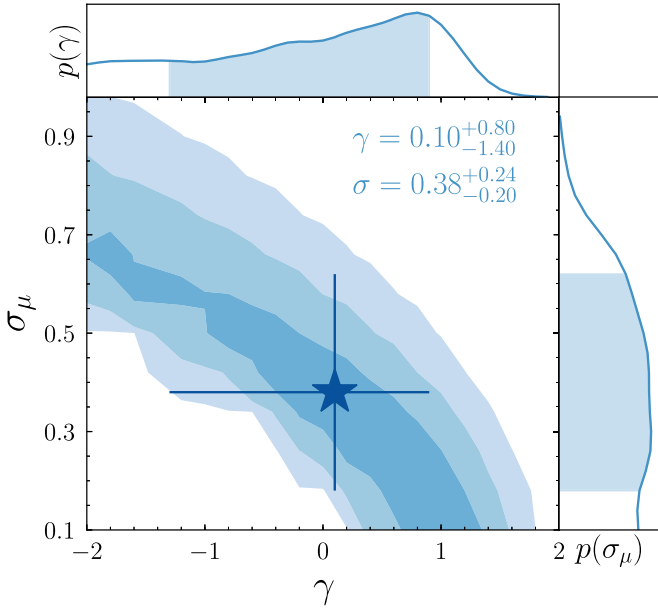


Figure 3. Constraining the evolution factor γ and the intrinsic scatter σ_μ based on a specific model ($\beta = 0.6$). The contours represent the 1σ – 3σ confidence regions of the posterior probability distribution. The histograms are the marginalized posterior distributions. The star indicates the posterior median and its 1σ error.

about -0.2 dex and $+0.4$ dex for the respective case of $\beta = 1.0$ (i.e., no SE bias) and $\beta = 0.0$ (i.e., the maximum SE bias for which the BH mass tends to be overestimated by ~ 0.65 dex). In both cases, the expected positions of $z \sim 6$ quasars (i.e., the dashed and dotted curves) are offset from the actual observed ones, thus a more positively evolving (for $\beta = 1.0$) or a more negatively evolving (for $\beta = 0.0$) $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ relation is required to explain the observations.

3.3. Constraining the Intrinsic $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ Relation

While the observed offset can be reproduced by an unevolving $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ relation, evolutionary models cannot be ruled out. Therefore, we determine the constraints on the intrinsic evolution by generating mock AGNs with a range of γ and σ_μ (assuming $\beta = 0.6$) and comparing the resulting $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ relation (incorporating observational biases) with the observed one to derive the likelihood (see Section 5 in Li et al. 2021 for details).

The posterior distributions of γ and σ_μ assuming a bounded flat prior ($-2.0 < \gamma < 2.0$, $\sigma_\mu > 0.1$) are shown in Figure 3. There is a strong degeneracy between γ and σ_μ : either a positive evolution (i.e., $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ increases with redshift) with a small σ_μ , or a negative evolution with a large σ_μ can reproduce the large apparent offsets. The peak of the posterior distribution is slightly skewed toward a positive evolution with a small scatter, as the sample only probes overly massive quasars that are clustered at the top left of the local relation. The best-fit values based on the 16th, 50th, and the 84th percentiles of the marginalized posterior distributions are $\gamma = 0.10^{+0.80}_{-1.40}$ and $\sigma_\mu = 0.38^{+0.24}_{-0.20}$, which are consistent with an unevolving $\mathcal{M}_{\text{BH}}/\mathcal{M}_*$ out to $z \sim 6$ (e.g., Volonteri & Stark 2011; Schulze & Wisotzki 2014).

However, given the various systematic uncertainties involved in mass measurements and model assumptions, it is improper to interpret the complex evolution with a single

number based on a certain model. For instance, the simplified thin-disk approximation will yield an underestimated \mathcal{M}_{dyn} (thus \mathcal{M}_*) if the galaxy contains a dispersion-dominated component (e.g., Pensabene et al. 2020). In addition, while we intended to use the total \mathcal{M}_* of these quasars in our analysis, which has been shown to correlate better with \mathcal{M}_{BH} than using $\mathcal{M}_{\text{bulge}}$ at high redshifts (e.g., Jahnke et al. 2009; Schramm & Silverman 2013; Li et al. 2021), the [C II] emission line mainly traces the inner galaxy region but not the entire galaxy (e.g., Venemans et al. 2017). These effects, together with the uncertain gas fraction and the choice of β , can induce systematic shifts of the mass measurements and affect the posterior distributions. Moreover, the bias estimates also depend on the assumed underlying distribution functions (e.g., the ERDF), which are not well constrained at $z \sim 6$ (see Section 6.3 in Li et al. 2021 for a detailed discussion). Therefore, it is clear that the current data set is insufficient to robustly constrain the intrinsic $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ relation of the underlying SMBH population at $z \sim 6$.

3.4. Predicted Evolution in the $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ Plane

Although the intrinsic evolution is unclear, overly massive quasars do exist. It is inevitable that their vigorous BH growth needs to be inhibited at some point in order to avoid unreasonably large \mathcal{M}_{BH} and to prevent them from further deviating from the local relation. The subsequent locations of these quasars in the $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ plane could be assessed by their instantaneous BH growth rate (BHGR) and SFR (e.g., Venemans et al. 2016). The BHGR can be estimated as

$$\dot{\mathcal{M}}_{\text{BH}} = \frac{(1 - \eta)L_{\text{bol}}}{\eta c^2}, \quad (3)$$

where c is the speed of light and $\eta = 0.1$ is the assumed radiative efficiency. The adopted constant η is suitable for our moderately accreting SMBHs that span $0.15 \lesssim \lambda_{\text{Edd}} \lesssim 3.0$ as expected from standard thin-disk theory (Shakura & Sunyaev 1973) or slim-disk theory (Abramowicz et al. 1988) for mildly super-Eddington quasars (e.g., Inayoshi et al. 2019, 2020). We assume that these $z \sim 6$ quasars can continue to form stars at their current SFR for a period (Δt). At the same time, the SMBHs keep accreting at the measured BHGR for a fraction of this time (i.e., the AGN duty cycle). The challenge is to constrain how long a quasar can sustain its high growth rate with its own feedback (e.g., Valentini et al. 2021); recent observations report a short quasar lifetime at $z \sim 6$ ($\sim 10^6$ yr on average; Eilers et al. 2021). We estimate the \mathcal{M}_{BH} -dependent duty cycle from the active SMBH fraction (see the inset in Figure 4), which is derived from the ratio between the BHMF of Type 1 AGNs at $z \sim 4.75$ (Kelly & Shen 2013) to the total BHMF scaled from the SMF (Grazian et al. 2015) at the same redshift and assuming an $\mathcal{M}_{\text{BH}}\text{--}\mathcal{M}_*$ relation with $\gamma = 0.1$ and $\sigma_\mu = 0.38$ (Section 3.3). This method likely yields an upper limit on the BH growth time at the quasar accretion rate as most active BHs at a given \mathcal{M}_{BH} are of much lower luminosities. We adopt Δt as the minimum value between the gas depletion timescale ($t_{\text{del}} = \mathcal{M}_{\text{gas}}/\text{SFR}$; $\sim 10\text{--}1000$ Myr) and 100 Myr (arbitrary value chosen to better visualize the evolution trend, which is shorter than t_{del}). The direction of \mathcal{M}_{BH} and \mathcal{M}_* during this period are illustrated by the dashed arrows in

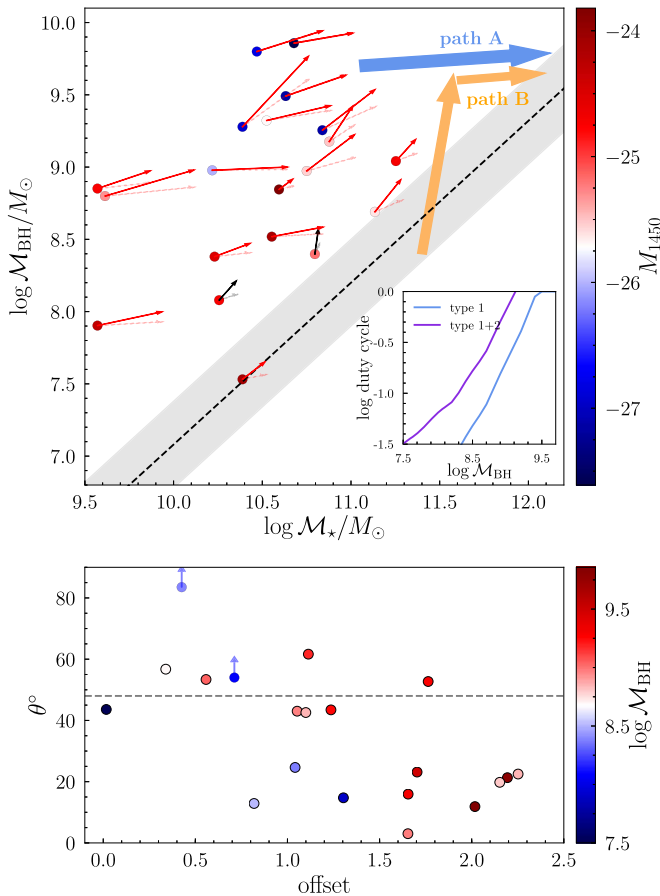


Figure 4. Top: predicted evolution of the growth of SMBHs and their host galaxies over the next $\min(t_{\text{del}}, 100)$ Myr. The dashed arrows represent the evolution vectors assuming the AGN duty cycle derived from the Type 1 BHMF, while the solid arrows show the evolution vectors after correcting the duty cycle for Type 2 AGNs. Two objects with upper limits in SFR are shown as black (gray) arrows. The adopted duty cycle– M_{BH} relations are shown in the inset. The blue and orange arrows show the possible evolution pathways for overly massive quasars. The black dashed line and the shaded region represent the local HR04 relation and its scatter. Bottom: the intersection angle as a function of offset. Each quasar is color-coded by their BH masses. Two objects with upper limits in SFR are marked by blue arrows. The dashed line at $\theta = 48^{\circ}$ represents the slope of the local relation.

the top panel of Figure 4, with the caveats that our estimation is a simplification of the complex physical processes (e.g., AGN feedback, gas accretion from the environment, merger) that could happen over the next ~ 100 Myr and the duty cycles for individual quasars are prone to significant uncertainties that are impossible to accurately constrain at present.

Interestingly, the predicted evolution exhibits a flow pattern, where quasars that are significant outliers in M_{BH}/M_{*} tend to converge to the local relation as previously seen at $z < 2$ (e.g., Merloni et al. 2010; Sun et al. 2015). We also show the evolution vectors (see the solid arrows) after correcting the active fraction for Type 2 AGNs (see the inset in Figure 4) assuming the luminosity-dependent obscured fraction ($\sim 60\%–80\%$ at $z \sim 6$) given by Vito et al. (2018) to account for the possible underestimate of the BH growth time that is not captured by the BHMF of Type 1s. Still, the general trend remains. The converging pattern also holds if we adopt the peak value of γ and σ_{μ} when estimating the total BHMF, which results in lower AGN duty cycles. However, the SFR derived from the single-band ALMA photometry may be overestimated

because quasars could also contribute to rest-frame $\sim 158 \mu\text{m}$ emissions (e.g., McKinney et al. 2021). Taking this effect into account will make the converging trend weaker or even disappear if the SFR is overestimated by a factor of $\gtrsim 2$. Multiband ALMA photometry is thus crucial to accurately constrain the dust temperature and subtract the quasar contamination when deriving the SFR.

In the bottom panel of Figure 4, we plot the intersection angle θ between the evolution vectors derived from the Type 1 + 2 duty cycle and the local relation as a function of offset (color-coded by M_{BH}) where a decreasing trend is evident. There are 12 of 14 quasars with an offset larger than 1.0 dex that have $\theta < 48^{\circ}$, where $\theta = 48^{\circ}$ corresponds to the slope of the local relation. The median intersection angle at offset > 1.0 dex is $\approx 23^{\circ}$, which is significantly smaller than that at offset < 1.0 dex ($\theta \approx 54^{\circ}$). It can also be seen that at similar BH masses, θ is smaller for quasars with larger offsets, thus the decreasing trend is not driven by less massive BHs with smaller offsets and shorter duty cycles. A natural explanation of the flow pattern and the decreasing trend could be AGN feedback (e.g., Valentini et al. 2021), which suppresses the growth of SMBHs once they deviate significantly from the local relation.

The converging pattern for the most-luminous and massive BHs may suggest that they have experienced rapid accretion episodes during seeding epochs and remain overly massive until they reach the local relation (i.e., path A in Figure 4) as shown by recent numerical simulations (e.g., Inayoshi et al. 2022). However, we cannot rule out the possibility that their progenitors are low-mass BHs moving upwards at similar stellar masses (i.e., path B in Figure 4). Such low- M_{BH} objects are undersampled in current surveys due to the detection limit, and their vigorous BH accretion may occur rapidly in a highly obscured or/and a radiative inefficient mode, which further reduces their apparent luminosity (e.g., Davies et al. 2019; Trebitsch et al. 2019). Therefore, it is essential to study the growth of low-mass systems and obscured quasars that have recently been discovered at $z \sim 6$ (e.g., Onoue et al. 2021).

4. Concluding Remarks

The $z \sim 6$ quasars typically have M_{BH}/M_{*} significantly larger than the local value. However, strong selection biases and significant measurement uncertainties severely limit the interpretation of the data. In this work, we account for these factors and demonstrate that the large apparent offsets and observed scatter could be reproduced by assuming that the underlying SMBH population at $z \sim 6$ follows the local $M_{\text{BH}}-M_{*}$ relation (Figure 2). However, a positive or even a negative evolution can also explain the data, depending on the evolution of the intrinsic scatter and various systematic uncertainties (Figure 3). It is thus crucial to emphasize that the evolution of the offset cannot be properly assessed without considering the scatter (see Li et al. 2021 for a similar issue at $z < 1$). Interestingly, quasars that are significant outliers in M_{BH}/M_{*} tend to have evolution vectors pointing toward the local relation (Figure 4). This may provide evidence that a self-regulated SMBH–galaxy coevolution scenario is already in place at $z \sim 6$, possibly driven by AGN feedback, although a robust conclusion can only be achieved with future observations that can accurately constrain the SFR (currently estimated from a single-band ALMA photometry assuming that the cold interstellar medium is mainly heated by star formation) and duty cycle for these quasars.

To break the degeneracy, expanding the current sample in both number statistics and to lower \mathcal{M}_{BH} limits is imperative (e.g., Habouzit et al. 2022). This is expected to be achieved by the ongoing SHELLQs survey (e.g., Matsuoka et al. 2016) and the forthcoming surveys by the Vera C. Rubin Observatory and Euclid, which will offer promisingly large and less-biased quasar samples with a more uniform selection function. It is also crucial to reduce the uncertainties (especially systematic effects) in the mass measurements. The James Webb Space Telescope will enable us to measure \mathcal{M}_{BH} using the more reliable $\text{H}\beta$ line and makes it possible to probe the stellar emissions of $z \sim 6$ quasars in the rest-frame optical bands, thus allowing direct measurements of their \mathcal{M}_* (e.g., Marshall et al. 2021). In the meantime, extending the current reverberation-mapped AGN sample to a wider parameter space is also important to validate and improve the virial \mathcal{M}_{BH} estimator, which will be achieved by the ongoing SDSS-V Black Hole Mapper survey. These efforts, together with a deeper understanding of the AGN accretion process (e.g., radiative efficiency, duty cycle), will allow us to better assess the connection between SMBH and galaxy growth in the reionization era.

We thank the referee for valuable suggestions that helped to improve the manuscript. J.Y.L. acknowledges support from the National Natural Science Foundation of China (12025303, 11890693). J.D.S. is supported by JSPS KAKENHI grant No. JP18H01251 and the World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan. K.I. acknowledges support from the National Natural Science Foundation of China (12073003, 12003003, 11721303, 11991052, 11950410493), the National Key R&D Program of China (2016YFA0400702), and the China Manned Space Project with No. CMS-CSST-2021-A04 and CMS-CSST-2021-A06.

Appendix Details of Generating a Mock AGN Sample

Here we outline the steps in our simulation by assuming $\gamma = 0.0$ and $\sigma_\mu = 0.3$ as an example. The simulation starts with the SMF at $z \sim 6$ given by Grazian et al. (2015) to generate a sample of mock galaxies ranging from $8.0 < \log \mathcal{M}_*/M_\odot < 11.0$. We assume that the $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ relation at $z \sim 6$ is the same as the local HR04 relation in terms of both mean and intrinsic scatter and randomly assign each mock galaxy a true \mathcal{M}_{BH} based on the HR04 relation. The resulting $\mathcal{M}_{\text{BH}}-\mathcal{M}_*$ distribution is shown in blue in Figure 5(a).

We convert \mathcal{M}_{BH} into bolometric luminosity by randomly sampling the intrinsic ERDF given by Kelly & Shen (2013) at $z = 4.75$ in the range of $-1.5 < \log \lambda_{\text{Edd}} < 0.5$ (e.g., Onoue et al. 2019; Shen et al. 2019). The bolometric luminosity is converted into rest-frame luminosity L_{3000} and absolute magnitude M_{1450} (Figure 5(b)) assuming the bolometric corrections to be 5.15 and 4.4, respectively (Richards et al. 2006). After these steps, we have the full knowledge of the true distribution of mock AGNs in the $\mathcal{M}_*-\mathcal{M}_{\text{BH}}-L$ plane.

We then add realistic uncertainties to the mass terms to resemble observations. Given the large uncertainties on \mathcal{M}_{dyn} and \mathcal{M}_{gas} , we first add a random Gaussian uncertainty with a standard deviation of 0.5 dex to each true \mathcal{M}_* (Figure 5(c)). We also derive a virial \mathcal{M}_{BH} for each mock AGN using their true \mathcal{M}_{BH} , L_{3000} , and an assumed Mg II FWHM distribution through Equation (1). In this step, we take into account the luminosity-dependent SE bias. The SE bias originates from uncorrelated scatter between AGN luminosity and broad-line width due to both the variability of an individual quasar and the object-by-object diversity in the broad-line region (BLR) properties at a fixed true

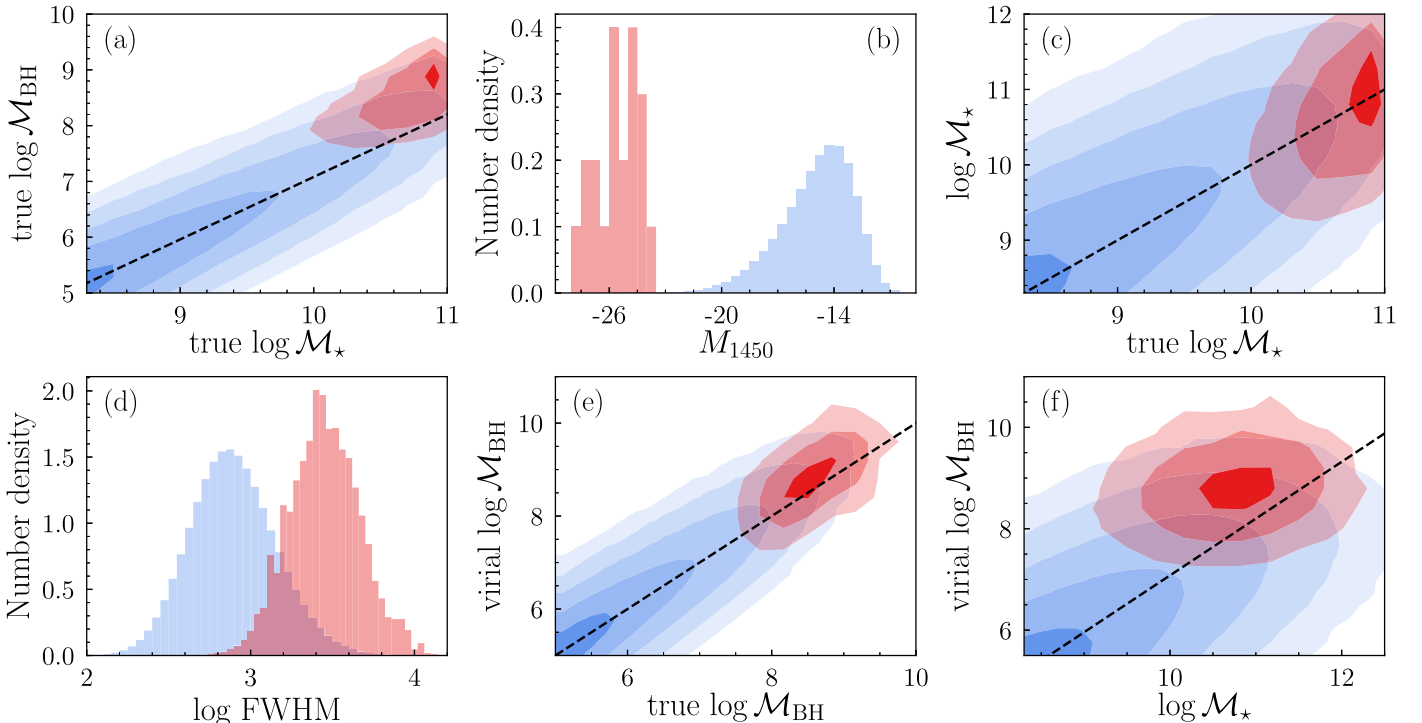


Figure 5. Parameter distributions of the mock AGN sample. The full mock sample is shown in blue with the contours representing the 1σ – 5σ levels. The mock sample matching in M_{1450} with the observed $z \sim 6$ quasars is shown in red with the contours representing the 1σ – 3σ levels.


\mathcal{M}_{BH} (Shen 2013). We consider the following cases to represent different levels of the SE bias:

1. Case A: virial \mathcal{M}_{BH} is an unbiased estimator of the true \mathcal{M}_{BH} regardless of AGN luminosity. This is done by assuming that the FWHM of Mg II follows a log-normal distribution with the mean value determined by the true \mathcal{M}_{BH} and L_{3000} for each mock AGN using Equation (1). We randomly sample the log-normal distribution with a dispersion of σ_{FWHM} to generate the FWHM for each source. The sampled FWHM is then combined with L_{3000} to derive the virial \mathcal{M}_{BH} . The dispersion σ_{FWHM} is chosen such that the resulting scatter of the virial \mathcal{M}_{BH} to the true \mathcal{M}_{BH} is 0.4 dex. By doing so, the variation of luminosity (relative to the mean) at a fixed true \mathcal{M}_{BH} is compensated by the concordant variation in FWHM.
2. Case B: Only part of the variation in luminosity can be compensated by line width. To simulate such a situation, for each mock AGN, we derive the difference (ΔL) between its luminosity (L) and the mean luminosity (\bar{L}) for all mock AGNs of the same true \mathcal{M}_{BH} . We assume that a fraction (β) of ΔL can be compensated by the line width by using $\bar{L} + \beta\Delta L$ to determine the mean of the log-normal FWHM distribution for each source. In this case, the higher (lower)-than-the-mean luminosity can only be partly compensated by the lower (higher)-than-the-mean line width, thus the resulting virial \mathcal{M}_{BH} tends to overestimate (underestimate) the true \mathcal{M}_{BH} except at $L = \bar{L}$.

It is currently unclear how strong β is. The nonbreathing effect of Mg II (i.e., the broad-line width does not respond to the continuum variability in individual quasar) suggests that β is not one (e.g., Yang et al. 2020). However, despite the lack of a BLR size (R)– L relation for individual quasars (in case of Mg II), a global R – L relation for a population of quasars spanning a broad range in BH masses and luminosities may still exist (e.g., Homayouni et al. 2020). This justifies the foundation of using the Mg II line as a single-epoch virial estimator, thus β is not likely to be zero. We adopt a relatively high response fraction ($\beta=0.6$) as our fiducial model. This assumption yields $\Delta \log \text{FWHM}_{\text{Mg II}} \propto -0.15\Delta L_{3000}$, as expected if the slope of the R – L relation for Mg II is ~ 0.3 (e.g., Homayouni et al. 2020). The resulting FWHM distribution and the virial \mathcal{M}_{BH} versus true \mathcal{M}_{BH} relation are shown in Figures 5(d) and (e), respectively.

With the aforementioned steps, we have generated the \mathcal{M}_{\star} and virial \mathcal{M}_{BH} for each mock AGN with realistic uncertainties (Figure 5(f)). In Figure 5 we show the distributions of a mock AGN sample matching in M_{1450} with the observed $z \sim 6$ quasars in red. The originally assumed underlying distributions are strongly modulated by the magnitude limit and the large uncertainties on both masses. In particular, the virial \mathcal{M}_{BH} tends to overestimate the true \mathcal{M}_{BH} by ~ 0.25 dex for this specific mock sample (Figure 5(e)), and the virial \mathcal{M}_{BH} versus \mathcal{M}_{\star} relation is significantly offset from the HR04 relation (Figure 5(f)).

ORCID iDs

Junyao Li  <https://orcid.org/0000-0002-1605-915X>
 Masayuki Akiyama  <https://orcid.org/0000-0002-2651-1701>
 Kohei Inayoshi  <https://orcid.org/0000-0001-9840-4959>

References

- Abramowicz, M. A., Czerny, B., Lasota, J. P., & Szuszkiewicz, E. 1988, *ApJ*, **332**, 646
- Chabrier, G. 2003, *PASP*, **115**, 763
- Davies, F. B., Hennawi, J. F., & Eilers, A.-C. 2019, *ApJL*, **884**, L19
- Decarli, R., Pensabene, A., Venemans, B., et al. 2022, arXiv:2203.03658
- Ding, X., Silverman, J., Treu, T., et al. 2020, *ApJ*, **888**, 37
- Eilers, A.-C., Hennawi, J. F., Davies, F. B., & Simcoe, R. A. 2021, *ApJ*, **917**, 38
- Fan, X., White, R. L., Davis, M., et al. 2000, *AJ*, **120**, 1167
- Genzel, R., Förster Schreiber, N. M., Übler, H., et al. 2017, *Natur*, **543**, 397
- Grazian, A., Fontana, A., Santini, P., et al. 2015, *A&A*, **575**, A96
- Habouzit, M., Onoue, M., Bañados, E., et al. 2022, *MNRAS*, **511**, 3751
- Håring, N., & Rix, H.-W. 2004, *ApJL*, **604**, L89
- Homayouni, Y., Trump, J. R., Grier, C. J., et al. 2020, *ApJ*, **901**, 55
- Inayoshi, K., Ichikawa, K., Ostriker, J. P., & Kuiper, R. 2019, *MNRAS*, **486**, 5377
- Inayoshi, K., Nakatani, R., Toyouchi, D., et al. 2022, *ApJ*, **927**, 237
- Inayoshi, K., Visbal, E., & Haiman, Z. 2020, *ARA&A*, **58**, 27
- Izumi, T., Matsuoka, Y., Fujimoto, S., et al. 2021, *ApJ*, **914**, 36
- Izumi, T., Onoue, M., Matsuoka, Y., et al. 2019, *PASJ*, **71**, 111
- Jahnke, K., Bongiorno, A., Brusa, M., et al. 2009, *ApJL*, **706**, L215
- Kelly, B. C., & Shen, Y. 2013, *ApJ*, **764**, 45
- King, A., & Pounds, K. 2015, *ARA&A*, **53**, 115
- Kormendy, J., & Ho, L. C. 2013, *ARA&A*, **51**, 511
- Lauer, T. R., Tremaine, S., Richstone, D., & Faber, S. M. 2007, *ApJ*, **670**, 249
- Li, J., Silverman, J. D., Ding, X., et al. 2021, *ApJ*, **922**, 142
- Marshall, M. A., Wyithe, J. S. B., Windhorst, R. A., et al. 2021, *MNRAS*, **506**, 1209
- Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, *ApJ*, **828**, 26
- McKinney, J., Hayward, C. C., Rosenthal, L. J., et al. 2021, *ApJ*, **921**, 55
- Merloni, A., Bongiorno, A., Bolzonella, M., et al. 2010, *ApJ*, **708**, 137
- Molina, J., Wang, R., Shangquan, J., et al. 2021, *ApJ*, **908**, 231
- Murphy, E. J., Condon, J. J., Schinnerer, E., et al. 2011, *ApJ*, **737**, 67
- Neeleman, M., Novak, M., Venemans, B. P., et al. 2021, *ApJ*, **911**, 141
- Onoue, M., Kashikawa, N., Matsuoka, Y., et al. 2019, *ApJ*, **880**, 77
- Onoue, M., Matsuoka, Y., Kashikawa, N., et al. 2021, *ApJ*, **919**, 61
- Pensabene, A., Carniani, S., Perna, M., et al. 2020, *A&A*, **637**, A84
- Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, *ApJS*, **166**, 470
- Schramm, M., & Silverman, J. D. 2013, *ApJ*, **767**, 13
- Schulze, A., & Wisotzki, L. 2014, *MNRAS*, **438**, 3422
- Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, **24**, 337
- Shen, Y. 2013, *BASI*, **41**, 61
- Shen, Y., Wu, J., Jiang, L., et al. 2019, *ApJ*, **873**, 35
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, **214**, 15
- Sun, M., Trump, J. R., Brandt, W. N., et al. 2015, *ApJ*, **802**, 14
- Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, *ApJ*, **853**, 179
- Trebtsch, M., Volonteri, M., & Dubois, Y. 2019, *MNRAS*, **487**, 819
- Valentini, M., Gallerani, S., & Ferrara, A. 2021, *MNRAS*, **507**, 1
- Venemans, B. P., Walter, F., Decarli, R., et al. 2017, *ApJ*, **837**, 146
- Venemans, B. P., Walter, F., Zschaechner, L., et al. 2016, *ApJ*, **816**, 37
- Vestergaard, M., & Osmer, P. S. 2009, *ApJ*, **699**, 800
- Vito, F., Brandt, W. N., Yang, G., et al. 2018, *MNRAS*, **473**, 2378
- Volonteri, M., & Reines, A. E. 2016, *ApJL*, **820**, L6
- Volonteri, M., & Stark, D. P. 2011, *MNRAS*, **417**, 2085
- Wang, R., Wagg, J., Carilli, C. L., et al. 2013, *ApJ*, **773**, 44
- Yang, Q., Shen, Y., Chen, Y.-C., et al. 2020, *MNRAS*, **493**, 5773