CrossMark

# A Large Sample of Extremely Metal-poor Galaxies at $z < 1$ Identified from the DESI Early Data

Hu Zou[1,2], Jipeng Sui[1,2], Amélie Saintonge[3], Dirk Scholte[3], John Moustakas[4], Malgorzata Siudek[5], Arjun Dey[6], Stephanie Juneau[6], Weijian Guo[1], Rebecca Canning[7], J. Aguilar[8], S. Ahlen[9], D. Brooks[3], T. Claybaugh[8], K. Dawson[10], A. de la Macorra[11], P. Doel[3], J. E. Forero-Romero[12], S. Gontcho A Gontcho[8], K. Honscheid[13], M. Landriau[8], L. Le Guillou[14], M. Manera[15], A. Meisner[6], R. Miquel[15], Jundan Nie[1], C. Poppett[16], M. Rezaie[17], G. Rossi[18], E. Sanchez[19], M. Schubnell[20], H. Seo[20], G. Tarlé[21], Zhimin Zhou[1], and Siwei Zou[22]

[1] Key Laboratory of Optical Astronomy, National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
zouhu@nao.cas.cn
[2] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 101408, People's Republic of China
[3] Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[4] Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA
[5] Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, E-08913 Bellaterra, Barcelona, Spain
[6] NSF's NOIRLab, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[7] Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK
[8] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[9] Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA
[10] Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA
[11] Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México
[12] Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia
[13] Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA
[14] Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), F-75005 Paris, France
[15] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, E-08193 Bellaterra, Barcelona, Spain
[16] Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA
[17] Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, USA
[18] Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Republic of Korea
[19] Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT), Madrid, Spain
[20] Department of Physics & Astronomy, Ohio University, Athens, OH 45701, USA
[21] University of Michigan, Ann Arbor, MI 48109, USA
[22] Department of Astronomy, Tsinghua University, 30 Shuangqing Road, Haidian District, Beijing, 100190, People's Republic of China

## Abstract

Extremely metal-poor galaxies (XMPGs) at relatively low redshift are excellent laboratories for studying galaxy formation and evolution in the early universe. Much effort has been spent on identifying them from large-scale spectroscopic surveys or spectroscopic follow-up observations. Previous work has identified a few hundred XMPGs. In this work, we obtain a large sample of 223 XMPGs at $z < 1$ from the early data of the Dark Energy Spectroscopic Instrument (DESI). The oxygen abundance is determined using the direct $T_e$ method based on the detection of the [O III]$\lambda$4363 line. The sample includes 95 confirmed XMPGs based on the oxygen abundance uncertainty; the remaining 128 galaxies are regarded as XMPG candidates. These XMPGs are only 0.01% of the total DESI observed galaxies. Their coordinates and other properties are provided in the paper. The most XMPGs have an oxygen abundance of $\sim 1/34\,Z_\odot$, a stellar mass of about $1.5 \times 10^7 M_\odot$, and a star formation rate of 0.22 $M_\odot\,\mathrm{yr}^{-1}$. The two most XMPGs present distinct morphologies suggesting different formation mechanisms. The local environmental investigation shows that XMPGs preferentially reside in relatively low-density regions. Many of them fall below the stellar mass–metallicity relations (MZRs) of normal star-forming galaxies. From a comparison of the MZR with theoretical simulations, it appears that XMPGs are good analogs to high-redshift star-forming galaxies. The nature of these XMPG populations will be further investigated in detail with larger and more complete samples from the ongoing DESI survey.

*Unified Astronomy Thesaurus concepts:* Metallicity (1031); Scaling relations (2031); Dwarf galaxies (416); Redshift surveys (1378)

*Supporting material:* machine-readable tables

## 1. Introduction

The hierarchical theory of galaxy formation suggests that large galaxies form through the assembly of smaller ones (White & Frenk 1991; Cole et al. 2000). It is expected that there are a vast number of low-mass galaxies with low metal

content in the early universe. Extremely metal-poor galaxies (XMPGs), which are chemically unevolved galaxies, provide ideal laboratories for validating the chemical evolution theories of galaxies and studying the physical processes in the early stages of their evolution. These galaxies are quite common at high redshift, but difficult to observe due to their low masses. The metallicities of galaxies at high redshift are also hard to reliably measure without large telescopes (e.g., Sanders et al. 2016b; Gburek et al. 2019), with JWST also now making this possible (e.g., Curti et al. 2023). Thus, local metal-poor

galaxies as possible analogs of primeval high-redshift young galaxies in terms of mass and metallicity, provide rich details for understanding the galaxy properties in extreme conditions, mass assembly, and elemental abundance in the early universe. We should also note that high-redshift dwarfs are formed in extremely overdense environments and the star formation process could be quite different in these conditions (e.g., Dekel et al. 2023). XMPGs have been used to explore chemical abundances, dust content, star formation, stellar feedback, escape of ionizing radiation, and cosmic environment in low-metallicity conditions (Thuan et al. 1999; Herrera-Camus et al. 2012; Filho et al. 2015; Sacchi et al. 2016; Olmo-García et al. 2017; Plat et al. 2019; Xu et al. 2022). Because of the early stage of galactic chemical evolution, they have been used to determine the primordial He abundance, providing critical constraints on Big Bang nucleosynthesis (Izotov & Thuan 2004; Hsyu et al. 2020; Matsumoto et al. 2022).

XMPGs are defined as galaxies with a metallicity of $Z < 0.1 Z_\odot$ (hereafter the metallicity is referred to as the gas-phase metallicity in H II regions). The most reliable metallicity measurements come from the so-called $T_e$ method, which relies on the detection of the [O III]$\lambda4363$ auroral line. The line intensity ratio of [O III]$\lambda4363$/[O III]$\lambda5007$ is sensitive to the electron temperature and hence the estimated oxygen abundance (Aller 1984; Nicholls et al. 2014). XMPGs are extremely rare and they constitute only about 0.1% of all galaxies in the local volume (Morales-Luis et al. 2011). Enormous efforts have been made during the past decades to identify XMPGs based on the detection of the [O III]$\lambda4363$ line in optical spectra. However, the [O III]$\lambda4363$ line is usually very weak, making the discovery of XMPGs difficult. Before the review of Kunth & Östlin (2000), it was reported that only 31 galaxies reached the metallicity threshold of XMPGs. Morales-Luis et al. (2011) identified a set of 130 XMPGs after a thorough bibliographic search and discovered 11 new ones with negligible [N II] lines in the seventh data release of the Sloan Digital Sky Survey (SDSS). This is because the [N II]/H$\alpha$ ratio is highly sensitive and positively correlates to the metallicity (Denicoló et al. 2002). Using the same spectroscopic data set, Sánchez Almeida et al. (2016a) have produced by far the largest sample of 196 XMPGs with intense [O III]$\lambda4363$ with respect to [O III]$\lambda5007$. A few more XMPGs have been discovered in later SDSS data releases (e.g., Gao et al. 2018) and other spectroscopic surveys, such as DEEP2 (Ly et al. 2015), LAMOST (Gao et al. 2017) and VUDS (Amorín et al. 2014). In addition to directly detecting the [O III]$\lambda4363$ line in the spectroscopic data, some authors attempted to select XMPG candidates from multiwavelength photometric data and confirmed them by spectroscopic follow-ups (Brown et al. 2008; Hsyu et al. 2018; Kojima et al. 2020; Isobe et al. 2022). Several tens of XMPGs have been found in this way. There are a total of about 350 known XMPGs collected from the literature by J. Sui et al. (2023, in preparation), who will utilize deep-learning algorithms to identify XMPG candidates from Dark Energy Spectroscopic Instrument (DESI) Legacy Imaging Surveys (Zou et al. 2017; Dey et al. 2019).

The number of existing XMPGs remains small. Systematically searching for XMPGs in large-scale spectroscopic surveys will considerably increase the sample size, improve the completeness, and provide an opportunity to take a systematic census of such an extreme galaxy population. DESI is embarking on a 5 yr spectroscopic survey to explore the nature of dark energy (Levi et al. 2013; DESI Collaboration et al.

2016a, 2016b). It is observing about 40 million extragalactic objects, including bright galaxies, luminous red galaxies (LRGs), emission-line galaxies (ELGs), and quasars (QSOs) (DESI Collaboration et al. 2016a; Myers et al. 2023; D. Schlegel et al. 2023, in preparation). The DESI Bright Galaxy Survey (BGS) will produce a magnitude-limited galaxy sample down to $r < 19.5$ mag (Hahn et al. 2023), helping us obtain a uniform XMPG sample at low redshift, while the ELG survey (Raichoor et al. 2023) can be used for detecting XPMGs at higher redshift. It is expected that several thousand XMPGs will be identified in the DESI survey, which is estimated from our XMPG detection with the DESI early data in this paper. We will use the direct $T_e$ method to measure the metallicities of galaxies to identify XMPGs at $z < 1$. With this sample, we can further study their photometric and spectral properties, investigate their local and global environment and gas components, and explore the scaling relations, star formation activity, and gas flows that are associated with such low-metallicity conditions.

This work focuses on the analysis of DESI early data, including the Early Data Release (EDR; DESI Collaboration et al. 2023b), the detection of XMPGs, and the exploration of their mass–metallicity relation (MZRs). The paper is organized as follows. In Section 2, we introduce the DESI spectroscopic data and our galaxy sample selection. Section 3 presents the determination of oxygen abundance and measurements of stellar mass and star formation rate (SFR). The XMGPs identified in this work and their basic properties are shown in Section 4. The XMPG distribution on the plane of stellar MZR is also discussed in this section. The summary is given in Section 5. Throughout this paper, we assume a flat $\Lambda$CDM cosmology with $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, and $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$. Both photometric magnitudes and spectral line fluxes are corrected for the Galactic extinction. The solar Oxygen abundance ($12 + \log(\text{O/H})_\odot = 8.69$) adopted in this paper is from Asplund et al. (2021).

## 2. Galaxy Sample Selection

### 2.1. DESI Spectroscopic Data

The DESI experiment is a new generation ground-based cosmological survey, aiming to explore the expansion history and structure growth rate of the universe with unprecedented precision and to understand the nature of dark energy (Levi et al. 2013; DESI Collaboration et al. 2016a, 2016b, 2023a, 2023b). It will accurately measure the redshifts of about 40 million galaxies and quasars over a 14,000 deg$^2$ footprint. The instrument is a multi-object spectroscopic system installed at the prime focus of the Mayall 4 m telescope at Kitt Peak, in Arizona (DESI Collaboration et al. 2022; Miller et al. 2023; Silber et al. 2023). It features a 3.2° field of view and 5000 fiber robots that can position optical fibers to 5000 objects and collect their spectra simultaneously. The fibers (diameter of 1″5) are fed to 10 three-arm spectrographs, which provide a continuous wavelength coverage of 3600–9800 Å. The spectral resolution varies from $R \sim 2000$ in the blue end to 5000 in the red, which is high enough to resolve the [O II] doublet.

The DESI spectroscopic targets include approximately 50 million extragalactic and Galactic objects, mainly selected from the DESI Legacy Imaging Surveys (Zou et al. 2017; Dey et al. 2019; D. Schlegel et al. 2023, in preparation). The imaging data are about 2 mag deeper than the SDSS, covering a sky area of 20,000 deg$^2$ in the optical $grz$ bands and including the latest 6 yr near-infrared observations from the Wide-field
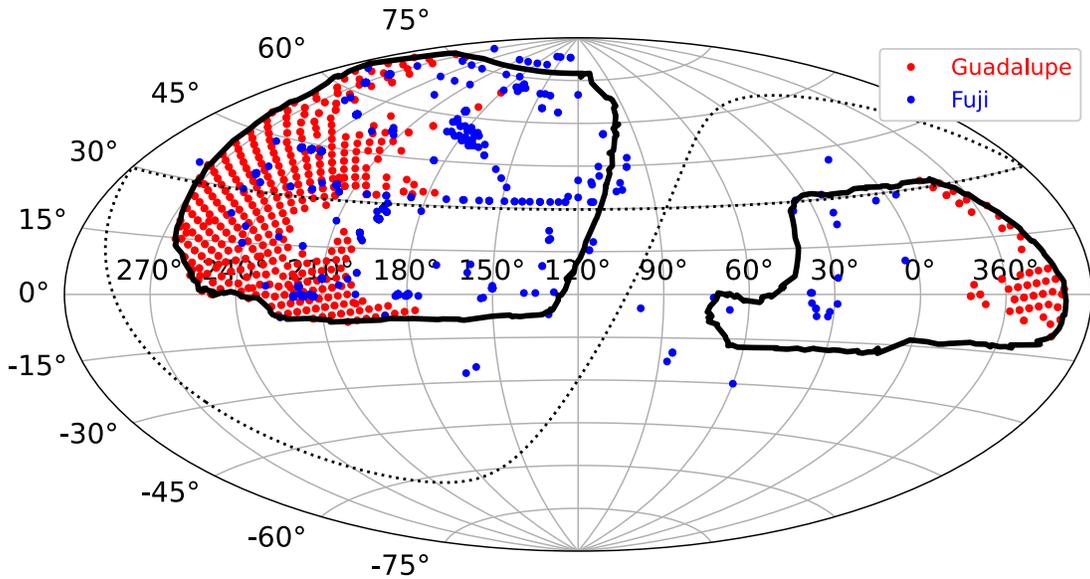
**Figure 1.** Sky coverage of the DESI early data in Aitoff projection. The black lines delineate the DESI baseline footprint of about 14,000 deg². The dotted line shows the Galactic plane. The circles represent the centers of 3°.2 DESI tiles (blue for Fuji and red for Guadalupe). Note that a few Fuji tiles are located outside of the DESI footprint.

Infrared Survey Explorer (WISE) satellite. The primary targets to be observed in dark time are LRGs at $0.4 < z < 1.1$, ELGs at $0.6 < z < 1.6$, and QSOs at $z < 3.5$ (Raichoor et al. 2020; Ruiz-Macias et al. 2020; Yèche et al. 2020; Zhou et al. 2020; Raichoor et al. 2023; Chaussidon et al. 2023; Zhou et al. 2023). During the bright time, DESI will carry the BGS of galaxies at $z < 0.6$ and the Milky Way survey of stars (Allende Prieto et al. 2020; Ruiz-Macias et al. 2020; Cooper et al. 2023; Hahn et al. 2023). There are also some secondary target programs to utilize the spare fibers and DESI capabilities, such as the LOW-Z Secondary Target Survey targeting low-redshift ($z < 0.03$) dwarf galaxies (Darragh-Ford et al. 2023).

The DESI project is expected to be a 5-year survey. Before the main survey started on 2021 May 14 there were several stages of survey validation (SV) for about 6 months (DESI Collaboration et al. 2023a), which aimed to refine the target selections and fiber assignment, improve the data reduction pipeline and observing efficiency, and validate the instrumental performances (DESI Collaboration et al. 2023a; A. Raichoor et al. 2023, in preparation; Alexander et al. 2023; S. Bailey et al. 2023, in preparation; D. Kirkby et al. 2023, in preparation; Lan et al. 2023; Schlafly et al. 2023), etc. An automatic spectroscopic data pipeline has been developed to reduce raw observations into wavelength- and flux-calibrated spectra, and to perform redshift measurements for all observed targets (Guy et al. 2023). The spectroscopic classifications and redshifts are determined using the `Redrock` pipeline[23] (S. Bailey et al. 2023, in preparation; Brodzeller et al. 2023), which fits a suite of templates of stars, galaxies, and quasars to the DESI spectra. There are several internal data releases generated by these data reduction and analysis pipelines, such as Fuji data assembly for the SV (732 observed tiles) and Guadalupe data assembly for the first 2 months of the main surveys (653 observed tiles). Although the EDR only includes the Fuji data DESI Collaboration et al. (2023b), we consider both Fuji and Guadalupe data assemblies as the DESI early data in this paper. Figure 1 shows the distribution of the observed DESI tiles in the full DESI footprint.

### 2.2. Flux Measurements of Emission Lines

The underlying stellar continuum is reproduced by using a full spectral fitting code of STARLIGHT Cid Fernandes et al. (2005), where the combination of 45 single stellar populations (SSPs) from Bruzual & Charlot (2003) and the Chabrier (2003) initial mass function (IMF) are adopted. The SSPs include 15 ages ranging from 1 Myr to 13 Gyr and three different metallicities (0.01, 0.02, 0.05). During the fitting, the main optical emission lines are masked out.

The modeled stellar continuum is subtracted from the observed spectrum and the fluxes of emission lines are then measured in the residual spectrum by Gaussian profile fitting. If not explicitly specified, all the spectral line measurements are calculated in the rest frame. We are interested in the following emission lines: [O II]$\lambda\lambda$ 3727,3729, H$\gamma$, [O III]$\lambda$4363, H$\beta$, [O III]$\lambda\lambda$ 4959,5007, [N II]$\lambda\lambda$ 6548,6583, H$\alpha$, and [S II]$\lambda\lambda$ 6716,6731. Due to their proximity, the following line groups are defined: [O II]$\lambda\lambda$ 3727,3729; H$\gamma$ and [O III]$\lambda$4363; [O III] $\lambda\lambda$ 4959,5007; [N II]$\lambda\lambda$ 6548,6583 and H$\alpha$; [S II]$\lambda\lambda$ 6716,6731. Emission-line fluxes in each group are fitted with Gaussian functions simultaneously in order to mitigate the contamination of neighboring lines. The steps of emission-line measurements are described below:

1. The local background is estimated by a linear fit to the spectrum within a 70 Å wavelength window around each line or line group, excluding regions affected by emission lines (15 Å width). The linear fit is performed with a robust $3\sigma$ clipping algorithm.
2. The fitted local linear background is subtracted from the spectrum and the background error is estimated from the rms of the linear fit.
3. The background-subtracted emission-line profiles (within a 15 Å window) are modeled with Gaussian functions, where the mean, standard deviation, and amplitude are free parameters to be fitted. The parameter uncertainties are provided by Gaussian fitting with the error spectrum as the weight.
4. In addition to the total flux, Gaussian width, and line centers, we also calculate the rest-frame equivalent width

---

23 https://github.com/desihub/redrock
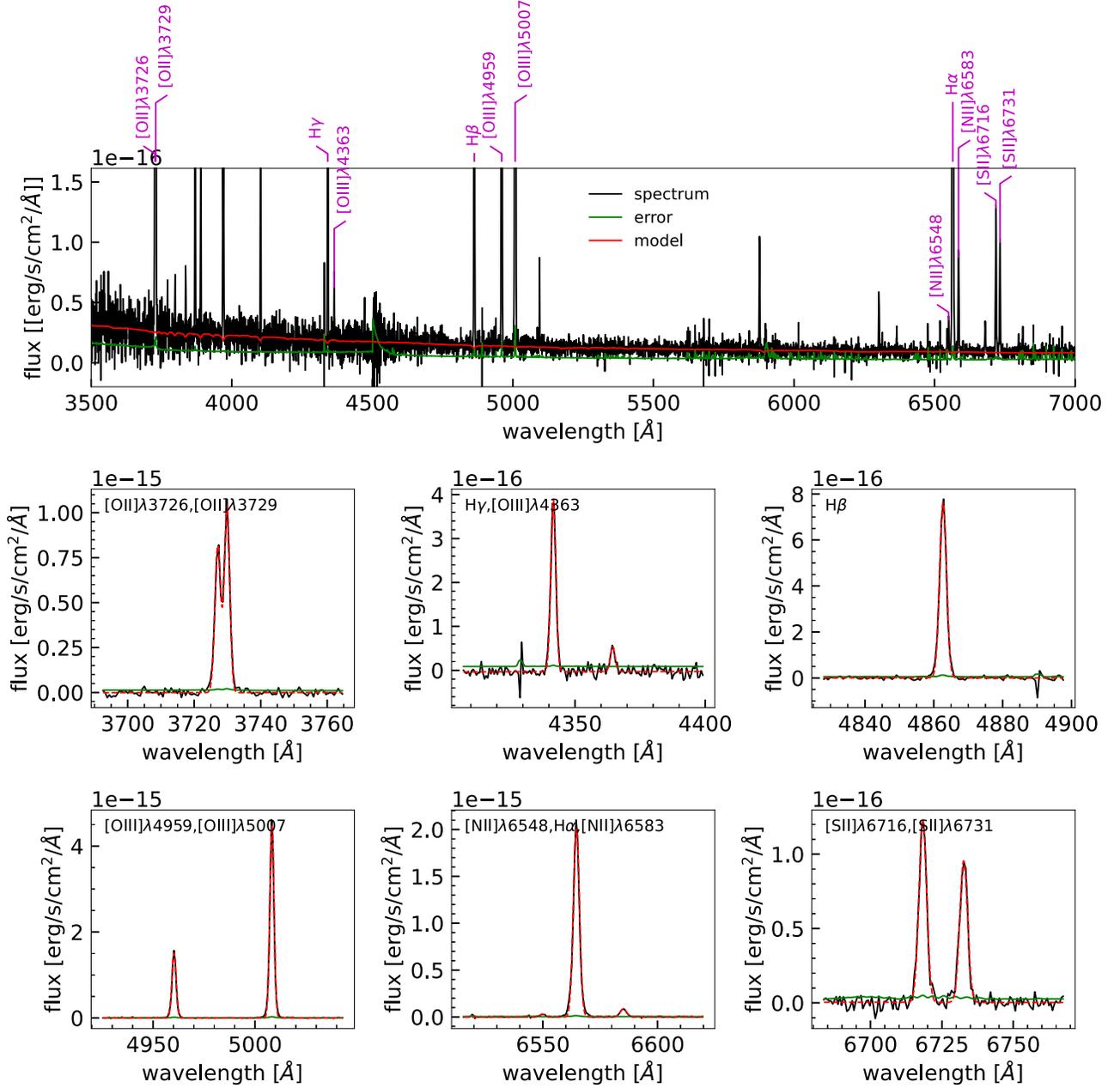
**Figure 2.** Example DESI galaxy spectrum and demonstration of the emission-line flux measurement technique. The top panel shows the spectrum of an arbitrarily selected galaxy at $z = 0.29$. The black line is the observed spectrum in the rest frame, the green line is the corresponding error spectrum, and the red line is the best-fitted model spectrum. The middle and bottom rows display local parts of the residual spectrum (observed—model) around a specified line or group of lines. Each line is fitted with a Gaussian function, and the grouped lines are fitted with multi-Gaussian functions simultaneously. The red dashed curve in each panel is the best-fitted line profile, which is a combination of the Gaussian function and a linear background.

for each emission line, whose error is estimated through error propagation.

Figure 2 shows a randomly selected DESI galaxy spectrum in rest-frame wavelength and the Gaussian fits to the emission-line profiles.

### 2.3. [O III]λ4363-detected Galaxy Sample

There are a total of 2,856,369 unique galaxies with successful measurements of spectral lines in the Fuji and Guadalupe data sets. These objects are selected from the redshift catalogs with SPECTYPE = "GALAXY" and ZWARN = 0, which means that they are spectroscopically confirmed galaxies with reliable redshift measurements.

Table 1 presents the sample selection steps and criteria. In order to eliminate spurious emission-line detections due to cosmic rays and large random noise, we only use emission lines with Gaussian FWHM > 1.2 Å, which is empirically set to be slightly larger than the pixel scale of 0.8 Å in observed wavelength. To obtain credible metallicity measurements as shown in Section 3, we require that the [O III]λ4363, [O II] λ3726, [O II]λ3729, [O III]λ4959, and [O III]λ5007 flux measurements have high enough signal-to-noise ratios (S/Ns) as shown in Table 1. Due to the lower limit of the S/N, ([O III] λ4363), the fitted FWHMs of this line might be abnormally large in some cases. In addition, the DESI data pipeline might mistakenly classify some broad-line active galactic nuclei (AGNs) as "GALAXY." So we set an upper FWHM limit of
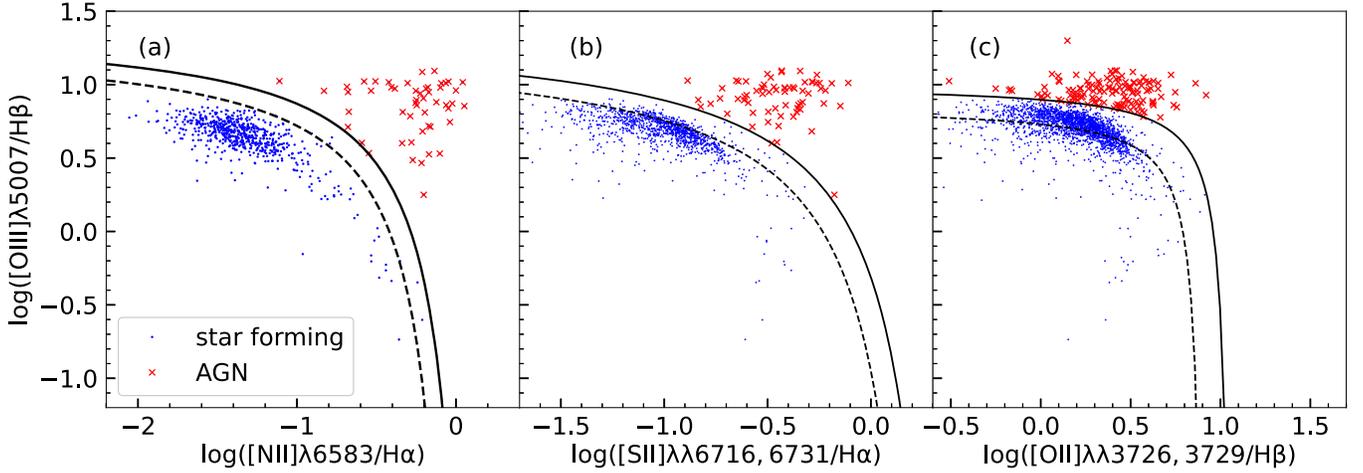
**Figure 3.** BPT diagnostic diagrams to discriminate star-forming galaxies (blue dots) from AGNs (red crosses). These galaxies shown are [O III]$\lambda$4363 detected. (a) The BPT diagram of [O III]$\lambda$5007/H$\beta$ vs. [N II]$\lambda$6583/H$\alpha$. The dashed line shows the empirical division between starbursts and AGNs obtained by Kauffmann et al. (2003). (b) The BPT diagram of [O III]/H$\beta$ vs. [S II]/H$\alpha$. The dashed line shows the theoretical division between starbursts and AGNs obtained by Kewley et al. (2001). The solid lines in these two panels display the demarcation curves with an additional uncertainty of 0.1 dex. (c) The galaxy distribution on the plot of [O III]$\lambda$5007/H$\beta$ vs. [O II]/H$\beta$. The dashed line presents the separation limit proposed by Lamareille (2010). The solid line displays the separation limit with the addition of an extra 0.15 dex uncertainty, which is used as our new proposed demarcation between starbursts and AGNs.

**Table 1**
Sample Selection of [O III]$\lambda$4363-detected Star-forming Galaxies

| Selection Steps | Selection Criteria | Number of Galaxies |
|---|---|---|
| Galaxies with reliable redshifts | SPECTYPE == GALAXY ZWARN == 0 | 2,856,369 |
| S/N and FWHM cuts | S/N([O III]$\lambda$4363) >3 S/N([O III]$\lambda$4959) >3 \|\| S/N([O III]$\lambda$5007) >3 S/N([O II]$\lambda$3726) >3 \|\| S/N([O II]$\lambda$3729) >3 S/N(H$\beta$) >3 S/N(H$\alpha$) >3 \|\| S/N(H$\gamma$) >3 FWHM([O III], [O II], H$\alpha$, H$\beta$, H$\gamma$, [S II], [N II]) >1.2 Å FWHM([O III]$\lambda$4363, H$\beta$, [O III]$\lambda$5007) <3 Å | 1633 |
| BPT diagnostic diagrams | Star-forming galaxies in BPT diagrams: [O III]/H$\beta$ versus [N II]/H$\alpha$ and [O III]/H$\beta$ versus [S II]/H$\alpha$\|\| [O III]/H$\beta$ versus [O II]/H$\alpha$ | 1623 |

3 Å for [O III]$\lambda$4363 and H$\beta$, although it might exclude some galaxies with galactic outflow that would be worth further investigation in the future. The inclusion of [O III]$\lambda$5007 limits our galaxies to the redshift range of $z \lesssim 0.96$. We obtain a total of 1633 [O III]$\lambda$4363-selected galaxies that satisfy all the above S/Ns and FWHM cuts.

The gas-phase extinction ($E(B - V)$) is calculated using the Balmer decrement, that is the change of the flux ratio of H$\alpha$/H$\beta$ or H$\beta$/H$\gamma$ relative to its intrinsic value. The intrinsic ratio (H$\alpha$/H$\beta$)$_0$ is 2.86 and (H$\beta$/H$\gamma$)$_0$ is 2.137, under the Case B recombination for $T_e = 10^4$ K and electron density of $N_e = 100$ cm$^{-3}$ (Hummer & Storey 1987). Here, we impose 3$\sigma$ cuts on H$\beta$ and either H$\alpha$ or H$\gamma$ emission lines. We preferentially use the H$\alpha$/H$\beta$ to calculate the extinction if the H$\alpha$ line is not redshifted out of the wavelength coverage ($z \lesssim 0.49$) and meets the S/N requirement. As a common practice, the $E(B - V)$ is set to zero when the observed line ratios are lower than the intrinsic values (about 32% of the galaxies). It could be due to the uncertainties of the line ratio or statistical deviations from the theoretical values. The median reddening is about 0.08 mag for the [O III]$\lambda$4363-selected galaxies. All the emission-line fluxes are corrected with the extinction using the attenuation law of Cardelli et al. (1989).

Only metal-poor star-forming galaxies are the focus of this paper, so AGNs should be also excluded. Their optical spectra are dominated by nonthermal emission, making it complicated to study the stellar population. Some diagnostic diagrams with emission-line ratios are usually used to discriminate different ionizing sources. Typically, the diagnostic diagrams of [O III]/H$\beta$ versus [N II]/H$\alpha$ (Baldwin et al. 1981) and [O III]/H$\beta$ versus [S II]/H$\alpha$ (Veilleux & Osterbrock 1987) are the standard tools to separate star-forming galaxies from AGNs (hereafter we refer to these diagrams as Baldwin–Phillips–Terlevich (BPT) diagrams). Figure 3 shows the method of using the BPT diagrams to select star-forming galaxies. The data points in this figure are all the [O III]$\lambda$4363-detected galaxies. In the diagram of [O III]$\lambda$5007/H$\beta$ versus [N II]$\lambda$6583/H$\alpha$ in Figure 3(a), the empirical starburst limit as shown in solid line is adopted from Kauffmann et al. (2003):

$$\log\left(\frac{[\text{O III}]\lambda5007}{\text{H}\beta}\right) = \frac{0.61}{\log([\text{N II}]\lambda6583/\text{H}\alpha) - 0.15} + 1.4,$$

(1)

where a modeling error of 0.1 dex (Kauffmann et al. 2003) is added to the original formula. The diagnostic diagram of [O III]/H$\beta$ versus [S II]/H$\alpha$ in Figure 3(b) provides a supplementary classification since the [N II]$\lambda$6583 line is not always detected and often blended with H$\alpha$. The separation

curve is proposed by Kewley et al. (2001):

$$\log\left(\frac{[\text{O\,{\footnotesize III}}]\,\lambda5007}{\text{H}\beta}\right)$$
$$= \frac{0.72}{\log([\text{S\,{\footnotesize II}}]\,\lambda\lambda6716,\,6731/\text{H}\alpha) - 0.42} + 1.40, \qquad (2)$$

where we also add an additional error of 0.1 dex (Kewley et al. 2001). The original separation line as shown in the dashed curve in Figure 3(b) is too close to the star-forming sequence. It would classify normal star-forming galaxies as AGNs. The above two BPT diagrams can only be used at redshifts $z < 0.49$. At higher redshift, we attempt to use the diagnostic diagram of [O\,{\footnotesize III}]$\lambda5007$/H$\beta$ versus [O\,{\footnotesize II}]/H$\beta$ first proposed by Lamareille et al. (2004) and revised in Lamareille (2010). Figure 3(c) shows the galaxy distribution on this diagram. The dashed line in this plot shows the updated demarcation curve of Lamareille (2010) to separate starbursts and AGNs. The solid line shows the separation limit plus an uncertainty of 0.15 dex. From this figure, we can see that a significant fraction of galaxies in the star-forming regions of the former two BPT diagrams lie above the curve, so the Lamareille (2010) separation is not optimal, especially for galaxies with small [O\,{\footnotesize II}]/H$\beta$ and large [O\,{\footnotesize III}]$\lambda5007$/H$\beta$. Hence, we adopt the demarcation curve with the addition of 0.15 dex uncertainty (Lamareille 2010), which is expressed as

$$\log\left(\frac{[\text{O\,{\footnotesize III}}]\,\lambda5007}{\text{H}\beta}\right)$$
$$= \frac{0.11}{\log([\text{O\,{\footnotesize II}}]\,\lambda\lambda3726,\,3729/\text{H}\beta) - 1.07} + 1.00. \qquad (3)$$

This new proposed separation limit is used to screen AGNs at higher redshifts in our sample via the diagram of [O\,{\footnotesize III}]$\lambda5007$/H$\beta$ versus [O\,{\footnotesize II}]/H$\beta$. We should note that the additional errors are added to both ordinate and abscissa in Figure 3 in order to set looser selection cuts of star-forming galaxies and make discrimination curves more suitable to our data. By using the above three BPT diagnostic diagrams, we get 1623 star-forming galaxies. All the following analyses in this paper are based on this sample.

## 3. Oxygen Abundance Determination and Other Property Measurements

### 3.1. Direct Measurement of Oxygen Abundance

The gas-phase metallicity can be estimated using the calibrations developed from either theoretical models or empirical methods. The direct method to measure the metallicity is the so-called $T_e$ method, which relies on the determination of the electron temperature ($T_e$) from auroral emission lines. $T_e$ can be measured by using the line intensity ratio of two different ionization species of the same element, such as [O\,{\footnotesize III}]$\lambda\lambda$ 4959,5007/[O\,{\footnotesize III}]$\lambda4363$ (Aller 1984). Although the direct $T_e$ method is subject to some caveats, such as metallicity underestimation caused by possible temperature fluctuation or gradients in high-metallicity star-forming regions (Kewley & Ellison 2008; Méndez-Delgado et al. 2023), it is currently the most reliable and accurate method to measure the metallicity.

In the direct $T_e$ method, the two-zone model is assumed, where the low- and high-ionization zones are traced by low- and high-ionization emission lines, respectively. The electron temperature in the high-ionization zone $T_e$([O\,{\footnotesize III}]) can be estimated using the line ratio of [O\,{\footnotesize III}]$\lambda4363$ and [O\,{\footnotesize III}]$\lambda\lambda$ 4959,5997, given the electron density ($N_e$) that can be calculated with the [O\,{\footnotesize II}] or [S\,{\footnotesize II}] doublets. The electron temperature in the low-ionization zone $T_e$([O\,{\footnotesize II}]) can be traced by [O\,{\footnotesize II}]$\lambda7320,7330$ lines or [N\,{\footnotesize II}]$\lambda5755$, which are too weak to be detected in the DESI spectra. An iterative method from Nicholls et al. (2014) is adopted to determine $T_e$([O\,{\footnotesize II}]) and the total oxygen abundance:

$$T_e([\text{O\,{\footnotesize II}}]) = T_e([\text{O\,{\footnotesize III}}])$$
$$\times (3.0794 - 0.086924Z - 0.1053Z^2 + 0.010225Z^3),$$
$$(4)$$

where $Z$ is the oxygen abundance. The iteration process starts with the O$^{++}$ abundance and the total oxygen abundance is the summation of the O$^+$ and O$^{++}$ abundances determined by the two-zone electron temperatures. Usually, the calculation converges within several iterations.

The [S\,{\footnotesize II}] doublet (6716 and 6731 Å) can be detected for galaxies at $z < 0.45$, while at higher redshift we can use the [O\,{\footnotesize II}]$\lambda\lambda$ 3726,3729 lines to calculate the electron density. We require that the flux S/Ns of these lines be greater than 4. If the S/Ns are not high enough or the line ratios are out of the valid ranges to determine $N_e$, we assume a constant $N_e = 100$ cm$^{-3}$ (Andrews & Martini 2013; Ly et al. 2014; Gao et al. 2017). The electron temperature is insensitive to $N_e$ in the relatively low-density regime and assuming $N_e = 100$ cm$^{-3}$ is consistent with the measurements from the [O\,{\footnotesize II}] or [S\,{\footnotesize II}] doublet. We use the Python package PYNEB[24] developed by Luridiana et al. (2015) to calculate the electron density, temperature, and oxygen abundance. As mentioned in Gao et al. (2017), the atomic recombination and collision strength data adopted in PYNEB are outdated, with the consequence that the electron temperature would be overestimated and thus the metallicity underestimated. We replace them with the updated data for O$^+$, O$^{++}$, and S$^+$ from Froese Fischer & Tachiev (2004), Kisielius et al. (2009), Tayal & Zatsarinny (2010), and Storey et al. (2014). If the [S\,{\footnotesize II}] or [O\,{\footnotesize II}] doublets have enough S/N, the module getCrossTemDen in PYNEB is used to calculate $N_e$ and $T_e$([O\,{\footnotesize III}]) simultaneously. If $N_e$ is set to 100 cm$^{-3}$, the PYNEB module getTemDen is used to calculate $T_e$([O\,{\footnotesize III}]). The abundance of O$^+$/H$^+$ is determined by the PYNEB module getIonAbundance with [O\,{\footnotesize III}]$\lambda\lambda$ 4959,5007/H$\beta$ and $T_e$([O\,{\footnotesize III}]). The O$^{++}$/H$^+$ abundance is obtained by the same PYNEB module with [O\,{\footnotesize II}]$\lambda\lambda$ 3726,3929/H$\beta$ and $T_e$([O\,{\footnotesize II}]) estimated by Equation (4). The total oxygen abundance 12+log(O/H) is calculated by adding the two abundances. The error estimation is implemented through Monte Carlo simulations. We add random errors to the fluxes of the measured emission lines according to their estimated uncertainties. This process is repeated 1000 times. Each combination of the fluctuated fluxes is fed to PYNEB to calculate $N_e$, $T_e$, and 12+log(O/H). The final value of each parameter is calculated as the median value and the error is estimated from the 68% confidence interval of the parameter distribution.

Figure 4 shows the distributions of $N_e$, $T_e$([O\,{\footnotesize III}]), and 12+log(O/H) and their error distributions. There are a total of

---

[24] http://www.iac.es/protecto/PyNeb/
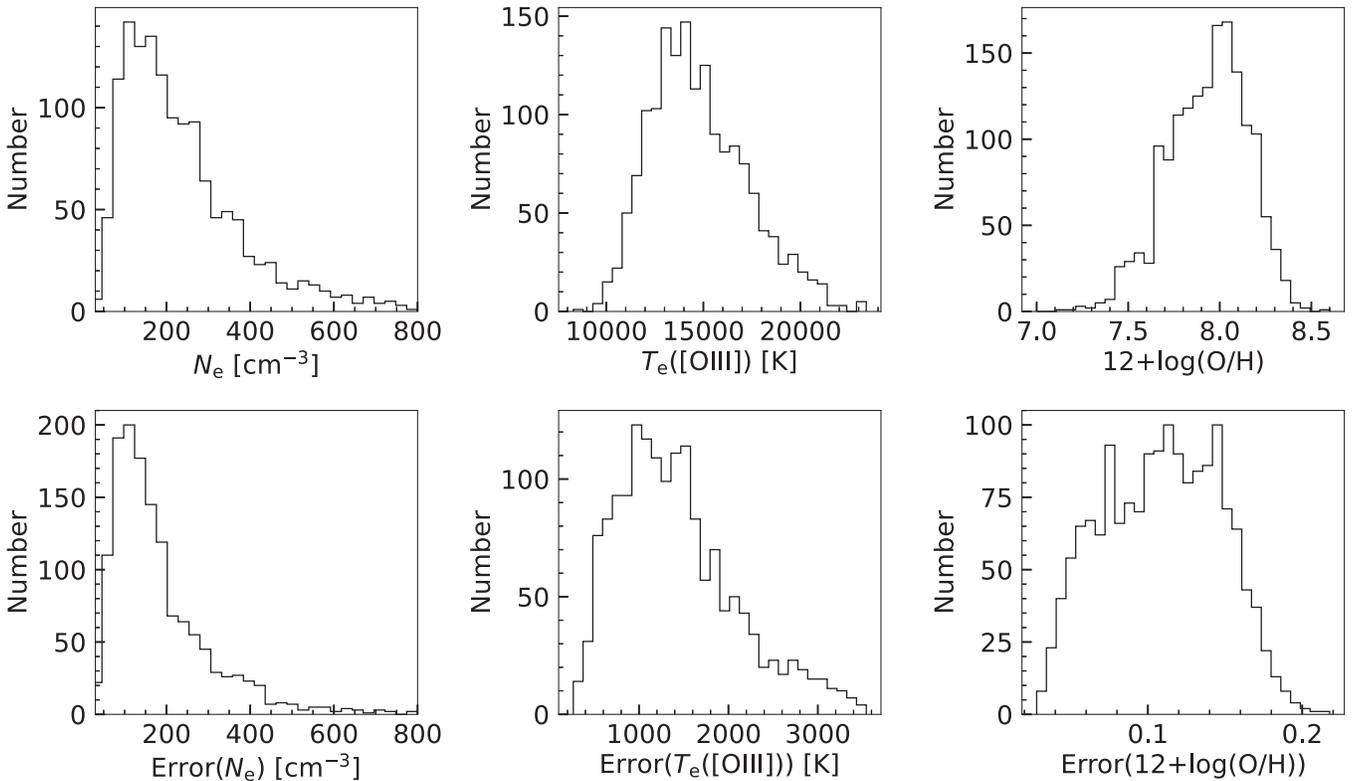
**Figure 4.** Top panels: distributions of $N_e$ (left), $T_e$([O III]) (middle), and 12+log(O/H) (right) of our [O III]$\lambda$4363-selected galaxies. Bottom panels: error distributions of $N_e$ (left), $T_e$([O III]) (middle), and 12+log(O/H) (right).

702 and 692 galaxies with $N_e$ calculated by [S II] and [O II], respectively. The remaining 229 galaxies have the assumed value of $N_e = 100$ cm$^{-3}$. About 90% of the galaxies have an electron density in the range of 76–629 cm$^{-3}$. The assumed $N_e = 100$ cm$^{-3}$ in this paper is in the valid range. There are about 1608 galaxies with effective $T_e$ measurement. The electron temperature of 90% of the galaxies ranges from 11,272–19,559 K and the median value is 14,421 K. The median error for $N_e$ and $T_e$([O III]) is 150 cm$^{-3}$ and 1326 K. Although the relative error of $N_e$ is large, it has little effect on the determination of the oxygen abundance due to its insensitivity to the electron density. The oxygen abundance of 90% of galaxies in our [O III]$\lambda$4363-selected galaxies ranges from 7.55–8.26 dex and the median value is 7.96 dex. The median metallicity error is 0.11 dex.

### 3.2. Stellar Mass Estimation

The stellar mass of each galaxy is estimated through stellar population synthesis fitting using the multiwavelength photometric data, as well as the DESI optical spectrum. The photometric data come from the data release DR9 of the DESI legacy imaging surveys (hereafter LS-DR9), which are mainly used for the target selections of the DESI spectroscopic survey (Zou et al. 2017; Dey et al. 2019). The imaging data in LS-DR9 provide both optical and infrared photometry, including three optical bands (g, r, and z) and four infrared bands (W1, W2, W3, and W4) from WISE. Due to much shallower photometric depths and lower spatial resolutions of W3 and W4 bands, we only use the W1 and W2 to construct the photometric SED. The limiting magnitudes in g, r, z, W1, and W2 are 24.7, 23.9, 23.0, 20.7, and 20.0 mag, respectively (Zou et al. 2022). Although the WISE imaging is relatively shallow and has a low resolution of

about 6″ (by contrast, 0.″9 for the z band), LS-DR9 provides consistent force-modeling photometry for both optical and infrared bands. Considering the importance of the near-infrared bands in estimating the stellar mass, we require that the two WISE bands have effective photometry with an S/N >2.

The DESI optical spectra, on the other hand, cover the wavelength range of 3650–9800 Å, providing complementary information about the stellar population. However, the continuum S/N especially for ELGs is low, so that we can hardly constrain the stellar mass by only using full spectral fitting techniques. Instead, we obtain artificial broadband spectrophotometry by convolving the optical spectra with self-defined 10 contiguous broad bands, which start from 3650 Å and have bandwidths of about 615 Å. Each self-defined filter is a boxcar filter with a flat transmission in a window of 615 Å width. Before convolution, we first scale each galaxy spectrum by multiplying by a scaling factor. The scaling factor is calculated as an aperture correction by comparing the total photometric fluxes in the r band with the spectrophotometric flux in this filter. Here we assume that the light collected by the fiber is proportional to the total light of the galaxy and the color of each galaxy is the same inside and outside of the fiber. The 10 customized broadband spectrophotometry as well as the five broadband photometry, are used for determining the stellar mass. Both five-band photometry and spectroscopic-based photometry are used together to infer stellar masses. We should also note that in some cases the DESI targets are not selected from the DESI imaging surveys, in which case the spectra are scaled to the the Gaia G-band photometry.

We use version 2022.1 of the Code Investigating GALaxy Emission (CIGALE; Boquien et al. 2019; Yang et al. 2020, 2022) to model the spectral energy distributions (SEDs) of galaxies. The CIGALE code can efficiently reconstruct the
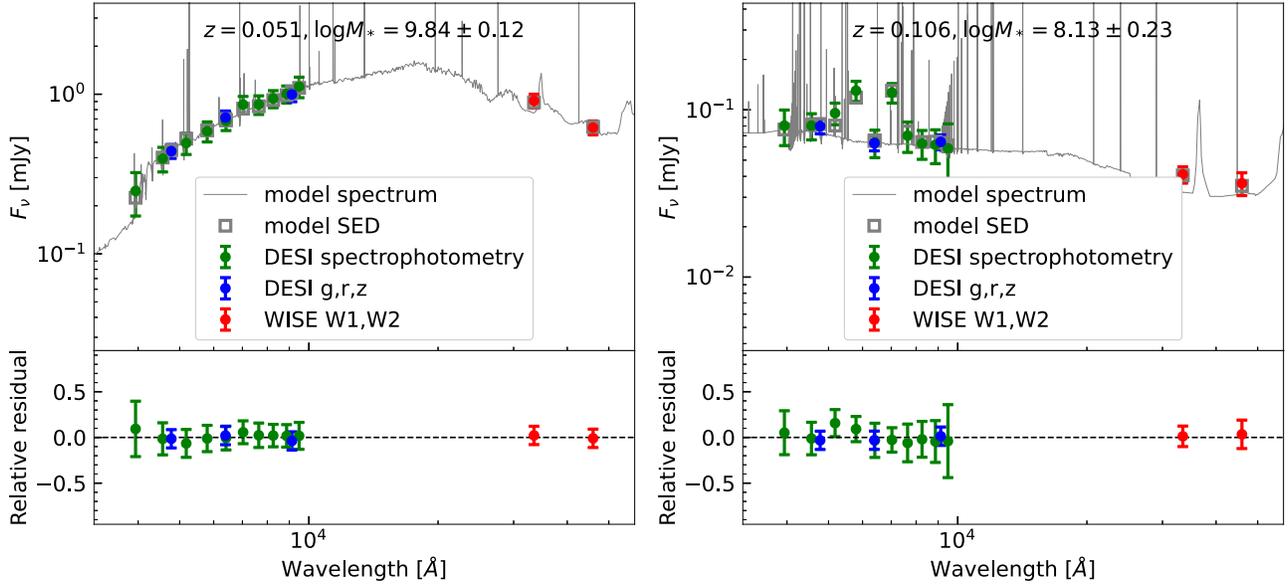
**Figure 5.** Two examples of the CIGALE stellar population synthesis fitting based on the DESI photometric and spectrophotometric data (left for an old quiescent galaxy at $z \sim 0.05$ and right for young star-forming galaxies at $z \sim 0.11$). The redshift and derived stellar mass for each galaxy are annotated in the top panel. The solid points with error bars represent the observed SEDs. The gray lines are the best-fit model spectra and the gray open squares display the model SEDs. The bottom panel shows the relative residual between observed and model SEDs, which is $(F_{\rm obs} - F_{\rm model})/F_{\rm obs}$, where $F_{\rm obs}$ and $F_{\rm model}$ are the observed and model fluxes for a given filter, respectively. The blue points with error bars show the DESI optical photometric fluxes and corresponding errors in the $g$, $r$, and $z$ bands, while the red ones display the WISE W1- and W2-band photometry. The green points denote the 10 spectrophotometric fluxes derived by convolving the DESI spectrum with 10 virtual broadband filters and scaled up to correct for aperture effects.

galaxy SED ranging from the far-ultraviolet to the infrared with highly flexible modeling of composite stellar population, gas emission, and dust attenuation and radiation. This modeling returns important physical properties of the galaxies, including stellar mass, SFR, dust extinction, and stellar age. In CIGALE, we adopt the simple stellar population (SSP) models of Bruzual & Charlot (2003), a Chabrier (2003) IMF, and a delayed star formation history (SFH). We select 21 stellar population ages (1 Myr to 13 Gyr), five stellar metallicities (0.0004, 0.004, 0.008, 0.02, 0.05), and 14 star formation timescales $\tau$ for the SFH ranging from 1 Myr to 80 Gyr. The Charlot & Fall (2000) extinction law and Draine et al. (2014) dust IR emission models are adopted to account for the dust attenuation of the interstellar medium and possible emission excess in WISE W1 and W2 bands, respectively. Due to strong emission lines in the spectra caused by the ionization of high-mass star radiation, the nebular emission is taken into account. However, no AGN emission is considered in the SED modeling, as we have already excluded those galaxies with AGN. The redshift is fixed to the spectroscopic redshift when constructing the models. Figure 5 shows two examples of the SED fitting with CIGALE to present two typical galaxies with old and young stellar populations, respectively.

To assess the quality of our stellar mass measurements, we crossmatch the entire DESI galaxy sample with the COS-MOS2020 catalog compiled by Weaver et al. (2022) for the Cosmic Evolution Survey (COSMOS). They have performed multiwavelength photometry (35 bands) and calculated photometric redshifts for galaxies down to the depth of $i \sim 27$ over a 2 deg$^2$ footprint. The *Farmer* catalog and stellar mass derived with the LePhare SED fitting tool is used in this paper. There are more than 6300 galaxies observed by DESI located in the COSMOS field. We limit the galaxies for comparison with the photometric redshift accuracy ($|\Delta z| = |(z_{\rm phot} - z_{\rm spec})/(1 + z_{\rm spec})|$, where $z_{\rm phot}$ and $z_{\rm spec}$ are photometric and
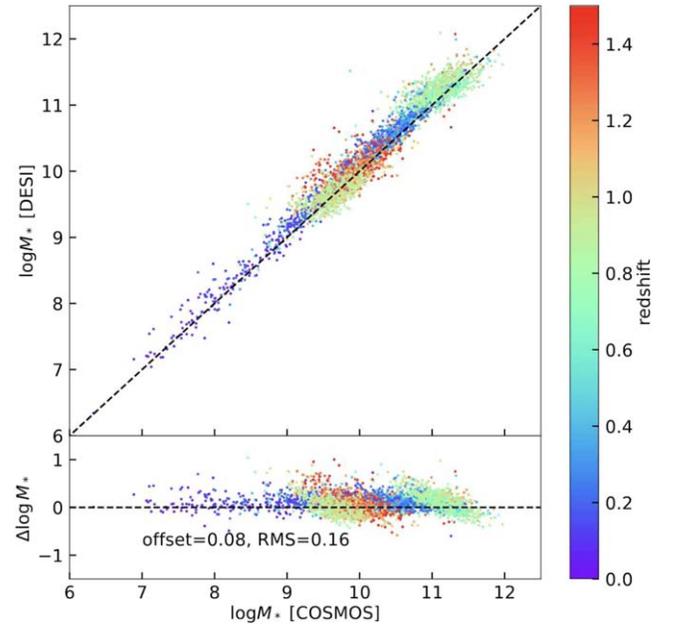


**Figure 6.** Comparison of the stellar masses from DESI and COSMOS2020. The top panel presents the stellar mass comparison and the diagonal line represents the exact match of the two measurements. The bottom panel shows the mass difference of these two surveys as a function of the stellar mass in the COSMOS2020 catalog. The horizontal line represents $\Delta \log M_* = 0$. The data points are color coded by the DESI spectroscopic redshift. The median difference and scatter of $\Delta \log M_*$ are also displayed in the bottom panel.

spectroscopic redshifts, respectively) less than 0.02 and the logarithmic stellar mass uncertainty less 0.1 dex (artificially set to relatively smaller values for better comparison), generating a set of about 3600 galaxies. Figure 6 shows the comparison of the stellar mass calculated using the DESI and COSMOS data. It can be seen that the stellar mass measurements of these two

surveys are consistent. There is a median bias of about 0.08 dex for $\Delta \log M_* = \log M_*[\mathrm{DESI}] - \log M_*[\mathrm{COSMOS}]$. It is mainly caused by different configurations in the SED fitting codes for constructing galaxy spectral models (e.g., different IMFs, model spectral libraries, SFHs). The dispersion $\sigma_{\Delta \log M_*}$ is about 0.16 dex, which is comparable to the estimated uncertainty with the DESI data (median of 0.18 dex). We also compare the stellar mass derived with only five-band photometry and find that it has a similar bias but a slightly larger dispersion of $\sigma_{\Delta \log M_*} \sim 0.17$ dex. Thus, the photometry plus spectrophotometry can indeed improve the stellar mass estimation. In particular, spectrophotometry is indispensable for those galaxies without DESI LS photometry.

### 3.3. SFR

The SFR is calculated from the de-reddened H$\alpha$ or H$\beta$ luminosity using the Kennicutt relation (Kennicutt 1998):

$$\mathrm{SFR} = 7.9 \times 10^{-42} L(\mathrm{H}\alpha) = 7.9 \times 10^{-42} F(\mathrm{H}\alpha) 4\pi d_L^2, \quad (5)$$

where $F(\mathrm{H}\alpha)$ is the extinction-corrected H$\alpha$ flux, $d_L$ is the luminosity distance, and $L(\mathrm{H}\alpha)$ is the H$\alpha$ luminosity. We preferentially use the H$\alpha$ luminosity to estimate the SFR. At $z > 0.49$ where the H$\alpha$ line is redshifted out of the wavelength coverage, the SFR from the de-reddened H$\beta$ luminosity is calculated assuming the intrinsic flux ratio of $(\mathrm{H}\alpha/\mathrm{H}\beta)_0 = 2.86$. Note that the emission-line luminosity is aperture corrected using the same scaling factor as the one when we estimate the stellar mass through SED fitting. The SFR error is mainly from the calibration uncertainty of Equation (5) (about 30%). Note that the uncertainty from the aperture correction is not included.

## 4. Our Identified XMGPs and Their Properties

### 4.1. XMPGs Identified from $T_e$-based Oxygen Abundance

From the oxygen abundance distribution in Figure 4, we can see that 902 [O III]$\lambda$4363-selected galaxies are metal-poor galaxies, which are defined to have $12+\log(\mathrm{O/H}) < 7.99$ ($Z < 0.2 \, Z_\odot$). No galaxy has an oxygen abundance larger than the solar metallicity. There are 223 XMPGs with $12+\log(\mathrm{O/H}) < 7.69$ ($Z < 0.1 \, Z_\odot$). These XMPGs account for 13.9% of all the [O III]$\lambda$4363-detected galaxies. They are about 0.01% of the total DESI observed galaxies, indicating their rarity. We define those galaxies with $Z + E(Z) < 0.1 Z_\odot$ as *confirmed XMPGs* and those with $Z < 0.1 Z_\odot$ but $Z + E(Z) \geqslant 0.1 Z_\odot$ as *XMPG candidates*, where $E(Z)$ is the abundance uncertainty. There are 95 confirmed XMPGs, 13 and two of which have metallicity less than $1/20$ and $1/30 Z_\odot$, respectively. The remaining 134 galaxies are XMPG candidates. Only six of our XMPGs are previously identified, if a crossmatching radius of $2''$ is considered.

Figure 7 shows the distributions of basic properties of all our XMPGs. About 74% of the XMPGs are located at a relatively low redshift of $z < 0.3$. There are 58 $z > 0.3$ XMPGs with the highest redshift up to 0.96. This is by far the largest sample of high-redshift XMPGs. The magnitude histogram in Figure 7(b) presents a bimodal distribution, which corresponds to the two types of DESI spectroscopic targets (brighter for BGS and fainter for ELG). Our XMPGs are almost dust-free, and the median $E(B-V)$ is about 0.01 mag. The median $N_e$ and $T_e$ are 182 cm$^{-3}$ and 18,927 K, respectively. The median [O III]$\lambda$5007

equivalent width (EW([O III]$\lambda$5007)) is about about 528 Å. The median SFR, stellar mass ($M_*$), and specific SFR (sSFR $= \mathrm{SFR}/M_*$) are 0.65 $M_\odot \, \mathrm{yr}^{-1}$, $4.65 \times 10^7 \, M_\odot$, and $1.14 \times 10^{-8} \, \mathrm{yr}^{-1}$. About 89% of our XMPGs are dwarf galaxies with $M_* < 10^9 M_\odot$.

For comparison, we also present the parameter distributions of all [O III]$\lambda$4363-detected samples in Figure 7. We notice that the XMPGs and the parent sample have similar distributions of redshift, magnitude, extinction, electron density, and EW([O III]$\lambda$5007). The XMPGs have lower metallicities and hence high $T_e$. From the SFR and mass distributions, we find that the XMPGs tend to present relatively lower SFRs than the parent sample, mostly due to their lower mass. Further investigations show that the sSFR distributions of these two samples are more similar. The median sSFR of XMPGs is around $1.14 \times 10^{-8} \, \mathrm{yr}^{-1}$, which is comparable to that of other [O III]$\lambda$4363-detected galaxies ($1.34 \times 10^{-8} \, \mathrm{yr}^{-1}$), but about eight times higher than that of normal star-forming galaxies (the control sample selected in Section 4.4), whose median sSFR $\sim 1.36 \times 10^{-9} \, \mathrm{yr}^{-1}$. Metal-poor galaxies are undergoing intense star formation. Tables 2 and 3 in the Appendix list the confirmed XMPGs and XMPG candidates identified in this work, respectively.

### 4.2. The Most Metal-poor Galaxies

The two most metal-poor galaxies in our sample, with an oxygen abundance $Z/Z_\odot < 1/30$, are DESIJ150535.89+314639.4 at $z = 0.054$ and DESIJ092331.28+645111.3 at $z = 0.005$. Figure 8 shows their color images composed with the $g$-, $r$-, and $z$-band imaging. The DESI photometric reduction pipeline implements forward modeling for each object with a point-spread function (PSF) profile or a galaxy profile (Dey et al. 2019). The intrinsic shape parameters used in this paper are from the Legacy Survey DR9 (D. Schlegel et al. 2023, in preparation). The model and residual images in Figure 8 show that the modeling works very well. The significant morphological discrepancy of these two XMPGs as shown in Figure 8 indicates possible different origins to their low metallicities. Figure 9 shows the observed spectra and photometric SEDs of these two XMPGs. Strong [O III]$\lambda$5007 and relatively weak [O II] emission lines in the spectra indicate that they have high-ionization environments. Strong H$\alpha$ emission lines suggest intense star formation.

*DESIJ150535.89+314639.4.* This is the most metal-poor galaxy (DESI TARGETID = 39628517750604871) identified by us. The oxygen abundance is $12+\log(\mathrm{O/H}) = 7.16 \pm 0.05$ ($Z \sim 2.9\% Z_\odot$). The lowest metallicity reported so far is about $1.6\% Z_\odot$ (Kojima et al. 2020). DESIJ150535.89+314639.4 is located at the redshift of $z = 0.054$. The $g$-band magnitude is 20.45 mag and the absolute magnitude is $M_g = -16.43$ mag. The optical colors are $(g-r) = -0.081$ and $(r-z) = -0.37$. The stellar mass and SFR are $1.51 \times 10^7 M_\odot$ and $0.22 \, M_\odot \, \mathrm{yr}^{-1}$, respectively, giving an sSFR of $1.43 \times 10^{-8} \, \mathrm{yr}^{-1}$. From Figure 8, we can see that DESIJ150535.89+314639.4 is a blue compact dwarf (BCD) galaxy with a half-light radius of $r_{\mathrm{eff}} = 640$ pc. The mean surface brightness $\mu_g$ within the half-light radius is about 21.22 mag arcsec$^{-2}$. It falls in the color cuts defined by Hsyu et al. (2018) to select BCD candidates and conforms to the BCD observational criteria listed in Morales-Luis et al. (2011).

*DESIJ092331.28+645111.3.* This is the second most metal-poor galaxy (DESI TARGETID = 39633440282250119). The
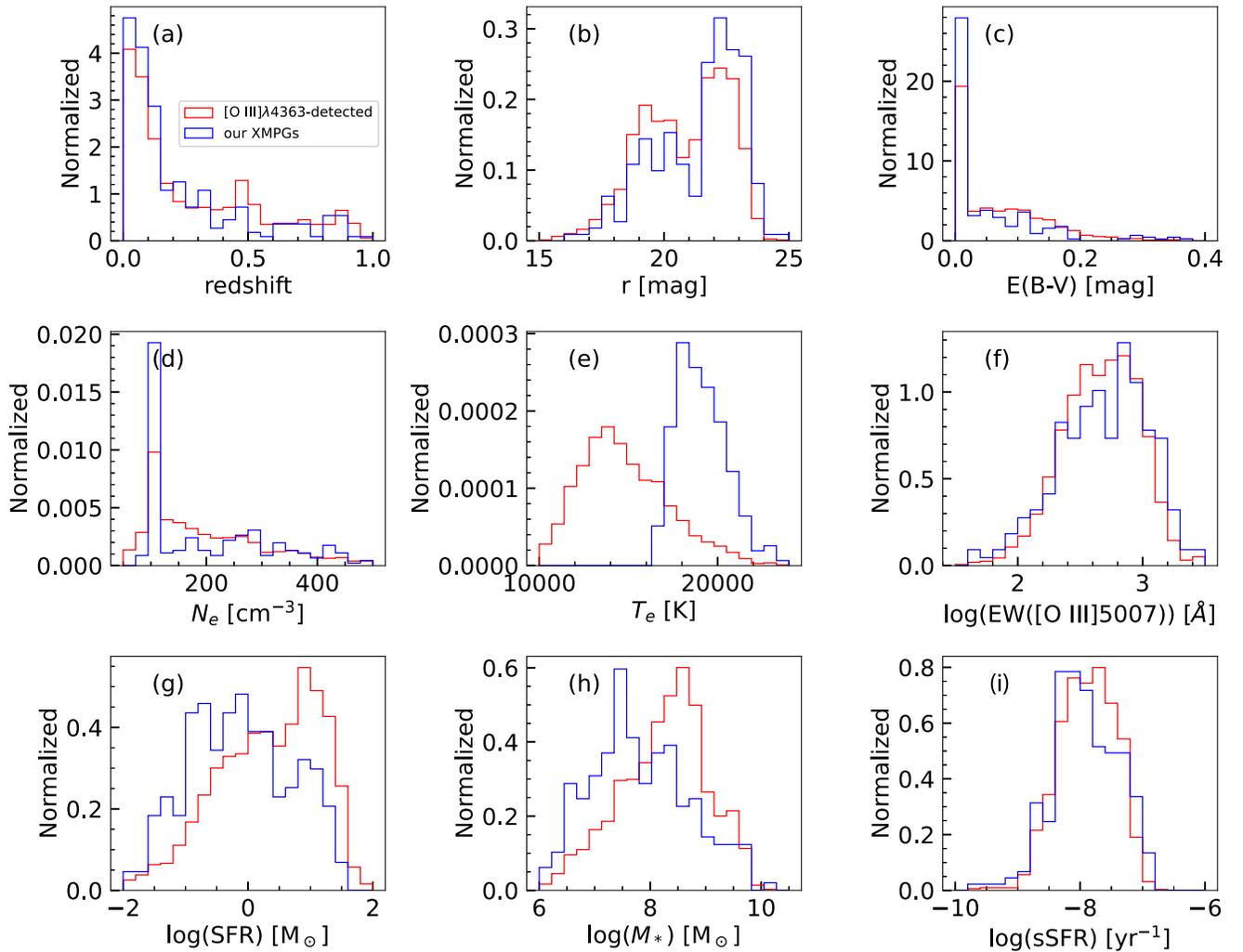
**Figure 7.** Distributions of basic properties for all our XMPGs (blue) and [O III]$\lambda$4363-detected galaxies (red), including redshift (a), $r$-band magnitude in mag (b), gas-phase extinction $E(B-V)$ in mag (c), electron density $N_e$ (d) in the unit of per cubic centimeter, electron temperature $T_e$ in K (e), logarithmic equivalent width of [O III]$\lambda$5007 in angstrom (f), logarithmic SFR in $M_\odot$ yr$^{-1}$ (g), logarithmic stellar mass in $M_\odot$ (h), and logarithmic sSFR per year (i).

metallicity is $12+\log(O/H) = 7.20 \pm 0.09$ ($Z \sim 3.2\%\,Z_\odot$). It is located at $z = 0.0054$ and $(\alpha, \delta) = (140°88032, 64°85315)$. We can see from Figure 8 that DESIJ092331.28+645111.3 has a complex morphology, exhibiting a tadpole-like or cometary shape. Such a head-tail morphology is common among XMPGs as shown in Morales-Luis et al. (2011), Sánchez Almeida et al. (2016a), and Kojima et al. (2020). The DESI fiber pointed to the head (i.e., the northwest part). The $g$-band apparent and absolute magnitudes of the head are $g = 20.46$ and $M_g = -11.38$. The optical colors are $(g - r) = 0.15$ and $(r - z) = -0.53$. The stellar mass and SFR are $1.38 \times 10^5 M_\odot$ and $0.0034$ $M_\odot$ yr$^{-1}$, respectively, giving an sSFR $= 2.42 \times 10^{-8}$ yr$^{-1}$. If we take the southeast part into account and assume the same mass–luminosity ratio ($M/L$) as the northwest one, the total stellar mass would be 1.67 times larger (may be larger due to possibly higher $M/L$ for the tail). The sSFR is slightly larger than DESIJ150535.89+314639.4, suggesting that its mass assembly is faster although lower mass. Nevertheless, this extremely metal-poor head shows a blue color and compact size ($r_{\rm eff} = 172$ pc). It connects to an extended fainter tail with a lower surface brightness. It was found that the XMPG heads and tails are dynamically connected and the tails have older stellar

populations and are more metal-rich than the heads (Sánchez Almeida et al. 2015).

DESIJ150535.89+314639.4 and DESIJ150535.89+314639.4. present two different morphologies, suggesting that they might have distinct physical origins. It is also possible that DESIJ150535.89+314639.4 might be in the stage of bursty star formation, so that the low surface brightness tail is outshone by the bright starburst clump. Very faint structures can be seen around this galaxy. To check these, we need higher-resolution and deeper images. We will systematically investigate the morphologies and stellar population properties of our XMPG sample in the future.

### 4.3. XMPG Distributions on the Color–Color and BPT Diagrams

The number of XMPGs identified here with DESI is almost double what is currently known. There are 338 XMPGs collected from the literature that are crossmatched with the DESI photometric catalogs (the photometric errors in $g$, $r$, and $z$ bands are limited to be less than 0.2 mag). Figure 10 compares the distributions of the DESI (this work) and literature XMPGs on the color–color diagram of $g - r$ versus $r - z$. The literature
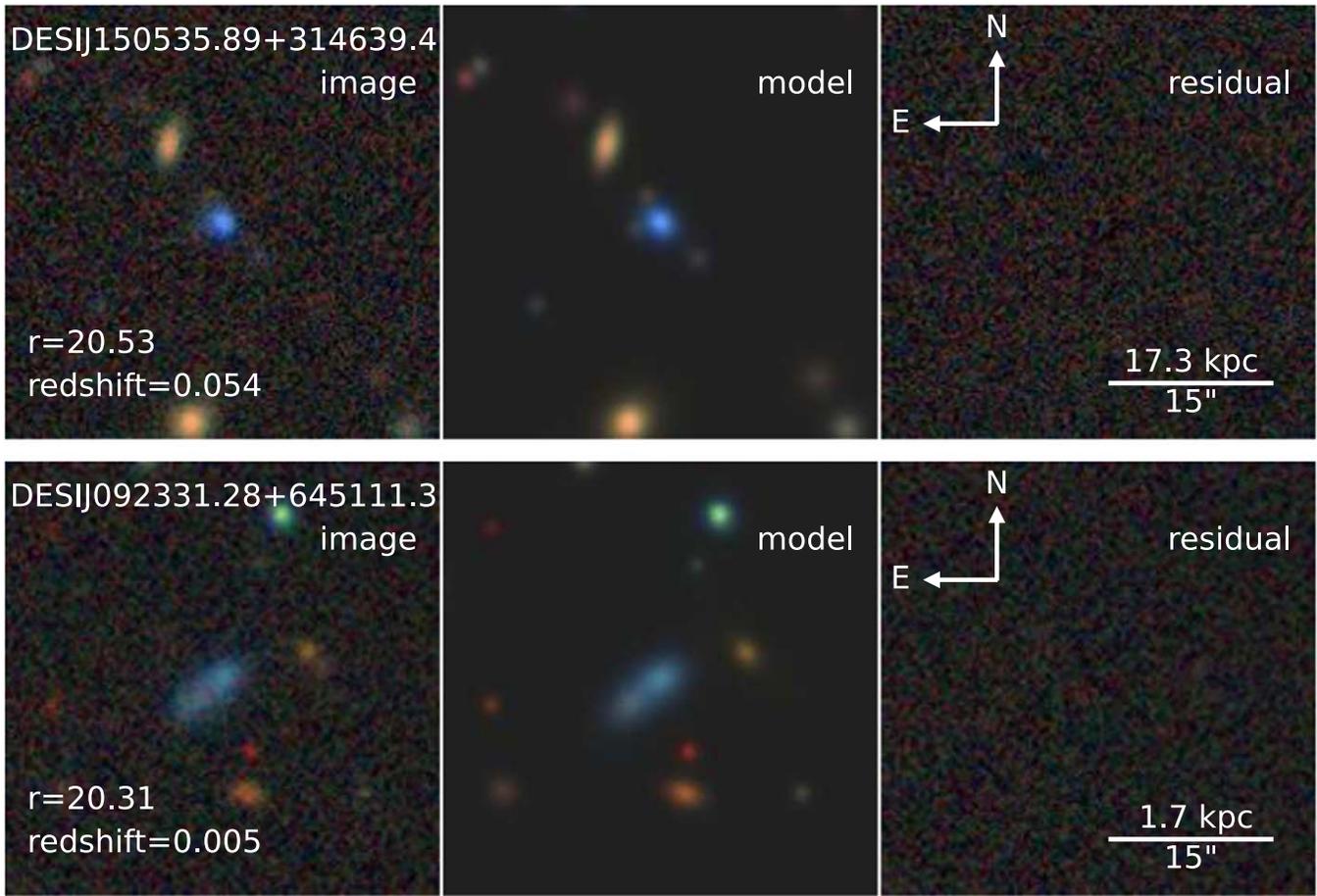
**Figure 8.** The two most metal-poor galaxies identified in this work. The left, middle, and right panels display the observed images, models, and residuals of the detected objects in the DESI Legacy imaging surveys, respectively. The residual is the subtraction of the model from the observed image, where the model is derived by the Tractor photometry code via forward modeling of the galaxy profile (Dey et al. 2019). The Tractor modeling provides the total integrated fluxes for galaxies. The top panels show the images of DESIJ150535.89+314639.4, and the bottom ones present those of DESIJ092331.28+645111.3. The *r*-band magnitude and redshift are marked in the left panels. The horizontal line denotes the scales in arcsec and kiloparsec at the distance of each galaxy. North is up and east is left.

XMPGs are collected by J. Sui et al. (2023, in preparation). The magnitudes have been corrected for the Galactic extinction. In the right panel of this figure, we can see that most XMPGs in the literature seem to lie in a low-redshift sequence. Only a few XMPGs are identified at redshift of $z > 0.3$. Our local XMPGs at $z < 0.1$ also follow this sequence. A large sample of relatively higher-redshift XMPGs are identified to the right and the top of the sequence. This is mainly caused by the strong H$\alpha$ and [O III]$\lambda$5007 lines redshifting into or out of the *r* and *z* bands, which is similar to the effect of green pea galaxies (Cardamone et al. 2009). These galaxies and their positions in the color space will help to identify XMPG candidates at $z < 1$ from future large-scale imaging surveys.

Figure 11 displays the distribution of our XMPGs on the BPT diagram of [O III]$\lambda$5007/H$\beta$ versus [O II]/H$\beta$. Other [O III]$\lambda$4363-detected galaxies are also overplotted in this figure. Almost all [O III]$\lambda$4363-detected galaxies except one galaxy are star-forming galaxies lying below the demarcation line. This galaxy is selected as a star-forming galaxy via the BPT diagrams of [O III]/H$\beta$ versus [N II]/H$\alpha$ and [O III]/H$\beta$ versus [S II]/H$\alpha$. Compared to other [O III]$\lambda$5007-detected galaxies, our XMPGs present slightly smaller [O III]/H$\beta$, but still have high-ionization conditions due to large [O III]/[O II] ratios. We also check the loci of these galaxies in the Mass-Excitation (MEx) diagnostic diagram (Juneau et al. 2011, 2014)

and find that only 2.5% of our XMPGs and 5.2% of [O III] $\lambda$4363-detected galaxies might host AGNs. Different diagnostic diagrams may have slightly different AGN predictions. In addition, there are parameter measurement uncertainties in these diagrams. So there could be some inconsistent classifications using different diagnostic diagrams.

### 4.4. Environment

To characterize the environment of our XMPGs, we adopt the statistics adopted in the analyses of the so-called blueberry galaxies (Yang et al. 2017), which is the distance (*D*) of the nearest galaxy. We construct a control sample from the DESI spectroscopic catalog for comparison. The comparison galaxy samples are selected through closest matching in both spatial distance and stellar mass with the total spectroscopically confirmed star-forming galaxies. This makes sure that the control sample has both the same redshift and mass distributions as those of our XMPGs. Figure 12 shows the cumulative distributions of *D* for these two samples. From this figure, we can see that there is an obvious difference between the two distributions. The Kolmogorov–Smirnov test gives a *p*-value close to zero, proving the significance. We conclude that our XMPGs inhabit relatively low-density environments. Nevertheless, we do not rule out that XMPGs might be mergers or
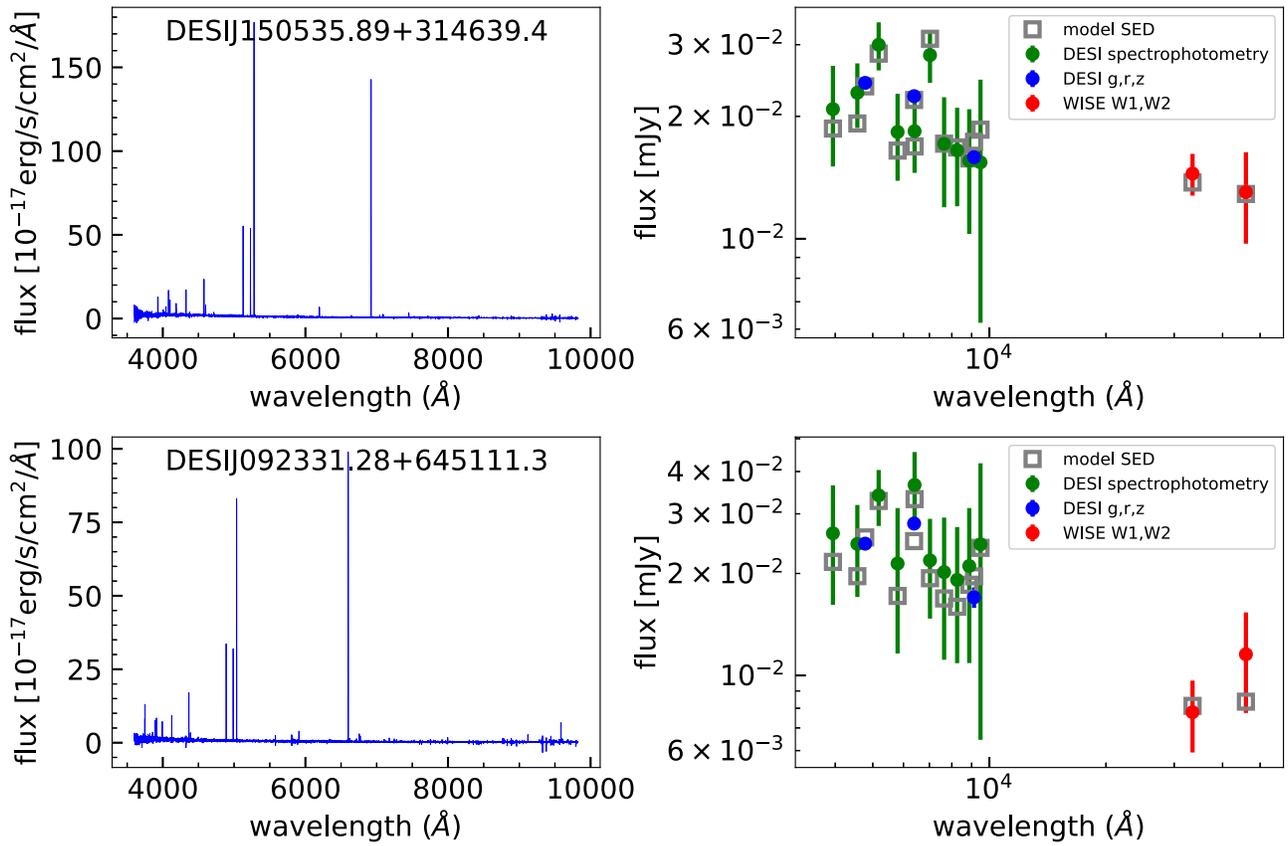
**Figure 9.** Observed spectra (left panels) and multiwavelength SEDs (right panels) for our two most metal-poor galaxies. The top panels are for DESIJ150535.89 +314639.4 and the bottom ones are for DESIJ092331.28+645111.3. The colored symbols are the same as those in Figure 5.
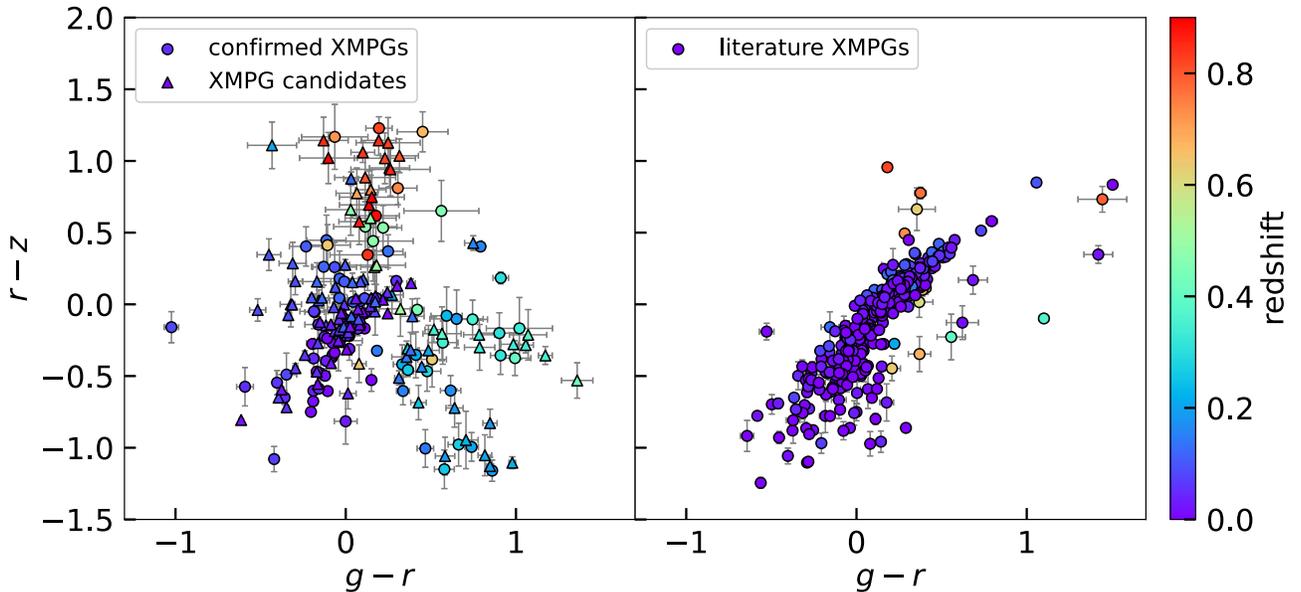


**Figure 10.** Observed color–color diagrams for our DESI XMPGs (left) and the XMPGs collected from the literature (right). The points in each panel are color coded by redshift. The gray error bars show the uncertainties of the colors. In the left panel, the circles represent our confirmed XMPGs, while the triangles represent the XMPG candidates.

closely interact with neighbor satellites because the statistics used in this paper are on a large scale.

As investigated by Cardamone et al. (2009) and Yang et al. (2017), both green pea and blueberry galaxies are strong [O III] and compact emitters at different redshifts. These galaxies are relatively metal-poor and are found to be situated in low-density environments. More than half of our XMPGs have effective radii less than 3 kpc and an equivalent width of [O III] λ5007 larger than 200 Å. These features are used to select green pea-like galaxies. It could be foreseeable that our XMPGs are also located in low-density regions. In addition, Filho et al. (2015) and Sánchez Almeida et al. (2016a) explored
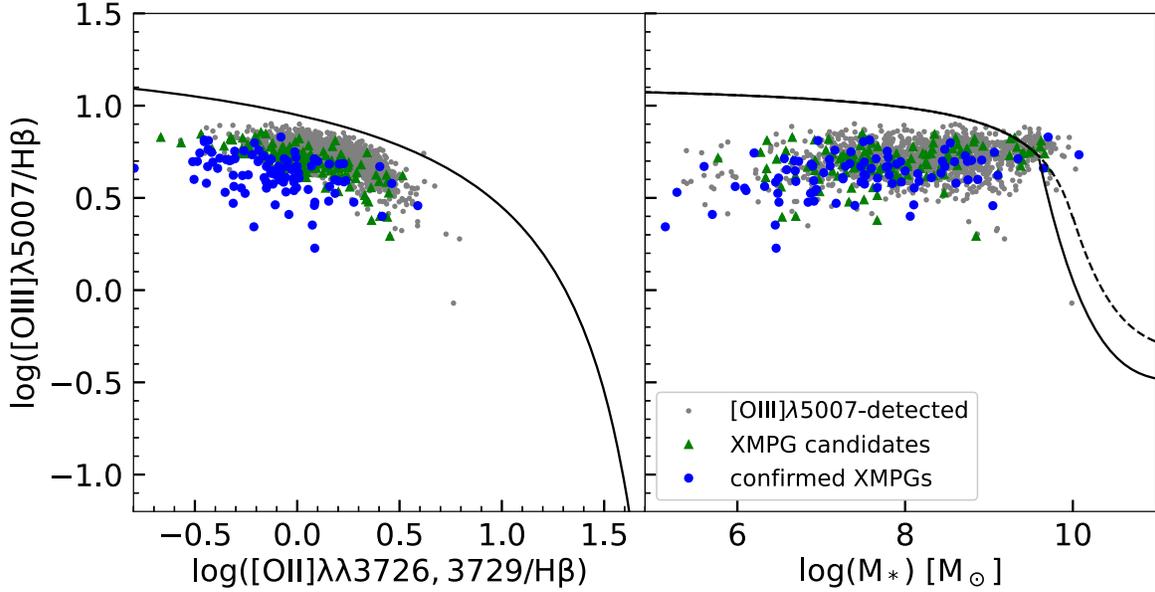
**Figure 11.** Left: BPT diagnostic diagram of [O III]$\lambda5007$/H$\beta$ vs. [O II]/H$\beta$ for all [O III]$\lambda4363$-detected galaxies (gray), confirmed XMPGs (blue), and XMPG candidates (green). The solid curve is the demarcation line as defined in Equation (3) to separate AGNs from star-forming galaxies. Right: MEx diagram for the same galaxies. The solid and dashed lines are the demarcation curves for starbursts (lower left) and AGNs (upper right), respectively, which come from Juneau et al. (2014).
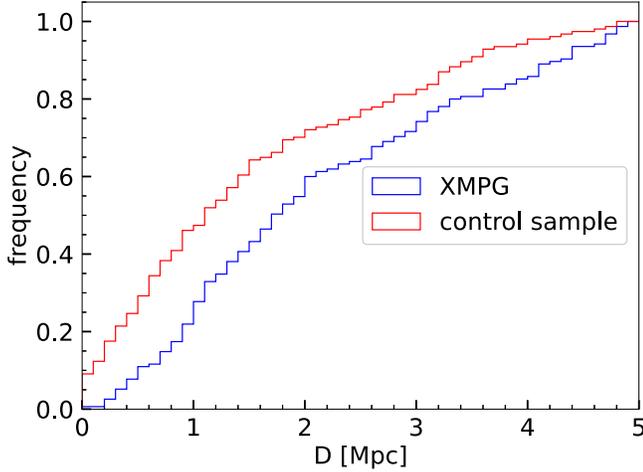


**Figure 12.** Cumulative distributions of the closest distance ($D$) in all DESI star-forming galaxies to each galaxy in our XMPG (blue) and control (red) samples.

the large-scale environment of XMPGs using the $N$-body cosmological numerical simulations. They found that XMPGs tend to reside in low-density environments similar to BCD galaxies. About 75% of their XMPGs have a strong tendency to reside in voids and sheets.

### 4.5. MZR of XMPGs

It has been well established that the metallicity of star-forming galaxies correlates with the stellar mass (Tremonti et al. 2004), forming the so-called MZR. The MZR reveals that more massive galaxies appear to be more chemically enriched and low-mass galaxies for a steeper relation in this mass–metallicity plane. It was also found that the MZR evolves with redshift. The relation still holds at high redshifts out to 3 and beyond, while high-redshift galaxies are observed to be less enriched (Erb et al. 2006; Maiolino et al. 2008; Moustakas et al. 2011; Zahid et al. 2013; Ly et al. 2016). Analytic models and numerical simulations of galactic chemical evolution have
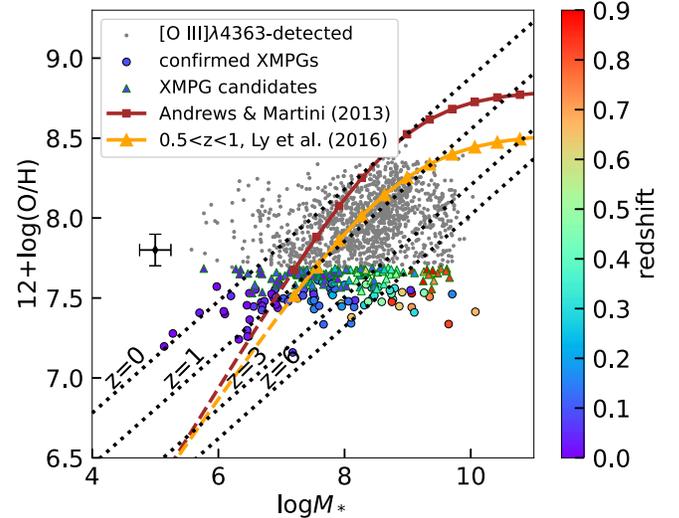


**Figure 13.** Stellar MZR of our XMPGs color coded by redshift. Smaller gray points are the [O III]$\lambda4363$-detected galaxies. The point with error bars shows the typical parameter measurement uncertainties. The local MZR from Andrews & Martini (2013) at $z \sim 0.08$ is plotted in brown with squares. The MZR at $0.5 < z < 1$ of Ly et al. (2016) is plotted in orange with triangles. The dashed extended lines correspond to the extrapolation of these two relations to the low-mass region. The dotted lines show the MZRs of theoretical simulations at $z = 0$, 1, 3, and 6 from Ma et al. (2016).

been developed to interpret the MZR (Davé et al. 2011; Lilly et al. 2013; Torrey et al. 2018). These theoretical models disclose that the existence of MZR is the consequence of the secular interplay among different physical processes including stellar formation, gas outflow, and gas recycling and accretion.

Figure 13 presents the MZR of our XMPGs (colored points) and all [O III]$\lambda4363$-detected galaxies (gray points). It can be seen that our XMPGs at higher redshift tend to be the galaxies with higher stellar mass, mainly due to the observational selection effect. In much MZR research, the metallicity determinations are based on strong-line calibrations. However, such determinations generate nearly 1 dex differences

depending on the choice of calibration method (Kewley & Ellison 2008; Moustakas et al. 2010). For this reason, the MZRs based on the $T_e$ method are chosen for consistent comparisons. Andrews & Martini (2013) (hereafter AM13) obtained an MZR at $z \sim 0.08$ via the direct method for normal star-forming galaxies using stacked SDSS spectra. This MZR, spanning a wide stellar mass range of $\log M_* = 7.4$–10.5, is plotted in Figure 13. We also display the MZR at $0.5 < z < 1$ derived by Ly et al. (2016), who used the direct $T_e$ method to measure the metallicity for several tens of galaxies at $z < 1$ with Keck/MMT spectra. From this figure, we can see that [O III] $\lambda 4363$-detected galaxies tend to lie below the local and Ly et al. (2016) MZRs, as expected since the [O III]$\lambda 4363$ line is more easily detected in low-metallicity galaxies.

Our XMPGs cover a wide range of stellar masses, from $M_* = \sim 10^5$ to $\sim 10^{10} M_\odot$. About 44% of our XMPGs at $z < 0.25$ are located below the AM13 MZR, and 100% of our XMPGs with $z > 0.5$ are located below the Ly et al. (2016) MZR, taking the MZR uncertainties (about 0.2 dex) into account. The deviations of our XMPGs from the MZRs of normal star-forming galaxies indicate that they might have distinct evolutionary pathways or reside in different environments. In addition, we also present the MZR evolution from the simulations of Ma et al. (2016). Their MZRs at $z = 0, 1, 3$, and 6 are shown in Figure 13. These simulations succeed in reproducing many observed properties (including MZR) of normal star-forming galaxies at $0 < z < 6$. It can be clearly seen that a considerable number of our XMPGs are located within or out of the MZR at $z = 3$–6. These XMPGs are probably the analogs of high-$z$ young galaxies, which are excellent samples for studying the evolution stage of the early universe.

In this analysis of the MZR relation, it is impossible to completely avoid the sample selection effect. Our XMPG galaxy sample comes from the DESI early data, which includes the observations from the SV and the first 2 months of the main surveys. The sample selection is extremely complicated because of different observing strategies during different SV stages and the main survey (DESI Collaboration et al. 2023a), different selection methods of galaxy targets, the incompleteness from the fiber assignment and redshift measurement, and the bias induced by the XMPG detection method. We use the [O III]$\lambda 4363$ line to identify XMPGs in this paper. This emission line is very weak and tends to be detectable in metal-poor galaxies, so it biases the galaxies with low metallicity. Nevertheless, it is obvious that the XMPGs deviate from the normal MZR, illustrating their peculiarity. It is also possible that the lower MZR of our samples relative to normal galaxies is a physical trend. Juneau et al. (2014) mentioned that there might be a relation between the emission-line luminosity and the metallicity trends. The low-redshift galaxies with high emission-line luminosity (like our metal-poor galaxy samples) follow a systematically lower MZR, similar to those galaxies at high redshift. It has been also observed that the MZR has a secondary parameter dependency on the SFR (Ellison et al. 2008; Mannucci et al. 2010). There is a more fundamental mutual relationship between $M_*$, metallicity and SFR, that is the so-called fundamental metallicity relation (FMR). It is found that there is no significant redshift evolution of FMR up to $z \sim 3$ (Mannucci et al. 2010). The existence of FMR means that galaxies with higher sSFR galaxies present lower metallicities.

## 5. Summary

In this work, we search for and identify a large sample of XMPGs at $z < 1$ from the early data of the DESI project and investigate their stellar mass–metallicity relation. It is expected that several thousands of XMPGs will be discovered from the full DESI survey in the future.

From the total number of 2.85 million galaxies in the DESI early data, we apply some S/N and FWHM cuts to select galaxies with good detections of emission lines and exclude galaxies with spurious line detection or broad-line AGNs. Diagnostic BPT diagrams are used to get rid of narrow-line AGNs. The final sample includes 1623 star-forming galaxies with significant [O III]$\lambda 4363$ detections. We adopt the direct $T_e$ method to measure the oxygen abundance, which relies on the detection of the [O III]$\lambda 4363$ auroral line. The stellar masses of the galaxies are estimated from broadband photometry and the DESI spectra.

A total of 223 XMPGs with oxygen abundance $Z < 0.1 Z_\odot$ are identified, including the largest sample of 58 XMPGs at $z > 0.3$. If taking the metallicity uncertainty into account, we get 95 confirmed XMPGs. The remaining 128 galaxies are XMPG candidates regardless of the uncertainty. Most of our XMPGs are dust-free dwarfs and have large EW([OIII]$\lambda 5007$) similar to green pea galaxies with strong [O III] emissions. The most XMPG has a metallicity of about $1/34$ of the solar value. A preliminary imaging examination of the two most metal-poor galaxies discovered in this study reveals two different morphologies, possibly indicating different evolution and physical origins. This will be investigated further in the future. The color–color diagram of $g - r$ versus $r - z$ shows a wide color distribution of our XMPGs in addition to the low-redshift sequence of the literature XMPGs. This is helpful in selecting XMPG targets at relatively high redshift for future spectroscopic follow-up. We characterize the local environment of our XMPGs by calculating the distance to the closest galaxy. In contrast with the control sample of normal star-forming galaxies, it is confirmed that our XMPGs tend to reside in relatively low-density environments.

A significant number of our XMPGs are located far below the local MZR and the MZR of normal star-forming galaxies in the similar redshift range. Compared with the simulations of Ma et al. (2016), our XMPGs are possibly low-redshift analogs of galaxies at high redshifts reaching $z = 6$ or beyond. They are excellent objects for exploring the evolution stage of the early universe. Future detailed investigations of these samples will provide critical constraints on this galaxy population.

## Appendix
## Basic Information and Properties of Our Confirmed XMPGs and XMPG Candidates

Tables 2 and 3 list the basic information and measured properties for our 95 confirmed XMPGs and 128 XMPG candidates, respectively. The data in these two tables and all data points shown in the published graphs are available in a machine-readable form in Zenodo doi:10.5281/zenodo.10276189.

**Table 2**
Properties of Our Confirmed XMPGs

| No (1) | Name (2) | TARGETID (3) | R.A. (deg) (4) | Decl. (deg) (5) | Redshift (6) | r (mag) (7) | E(B − V) (mag) (8) | EW([O III]) (Å) (9) | $N_e$ (cm$^{-3}$) (10) | $T_e$([O III]) (K) (11) | 12+log(O/H) (12) | $M_*$ log($M_\odot$) (13) | log(SFR) log($M_\odot$ yr$^{-1}$) (14) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | DESIJ150535.89 +314639.4 | 39628517750604871 | 226.39956 | 31.77763 | 0.0543 | 20.53 | 0.00 | 2.27 | 3.62e+02 ± 3.18e+02 | 2.33e+04 ± 1.30e+03 | 7.16 ± 0.05 | 7.18 ± 0.23 | −0.66 ± 0.13 |
| 2 | DESIJ092331.28 +645111.3 | 39633440282250119 | 140.88032 | 64.85315 | 0.0054 | 20.31 | 0.00 | 2.27 | 2.84e+02 ± 2.30e+02 | 2.07e+04 ± 1.89e+03 | 7.20 ± 0.09 | 5.14 ± 0.27 | −2.47 ± 0.13 |
| 3 | DESIJ213658.82 +041404.0 | 39627892803503091 | 324.24507 | 4.23445 | 0.0168 | 19.02 | 0.11 ± 0.03 | 2.93 | 4.92e+02 ± 4.18e+02 | 2.34e+04 ± 1.19e+03 | 7.24 ± 0.05 | 6.32 ± 0.09 | −0.61 ± 0.13 |
| 4 | DESIJ130103.92 +250558.2 | 39628373655292169 | 195.26632 | 25.09950 | 0.0257 | 20.68 | 0.00 ± 0.05 | 2.17 | 3.93e+02 ± 2.88e+02 | 2.02e+04 ± 3.15e+03 | 7.26 ± 0.17 | 6.46 ± 0.30 | −1.21 ± 0.13 |
| 5 | DESIJ120751.02 +002547.5 | 39627799857730585 | 181.96260 | 0.42989 | 0.0809 | 22.63 | 0.29 ± 0.19 | 2.86 | 100 | 2.20e+04 ± 2.42e+03 | 7.26 ± 0.11 | 6.48 ± 0.23 | −0.36 ± 0.13 |
| 6 | DESIJ144421.75 +072050.7 | 39627963238450509 | 221.09063 | 7.34742 | 0.0058 | 19.39 | 0.00 | 2.76 | 100 | 2.12e+04 ± 1.50e+03 | 7.28 ± 0.07 | 5.28 ± 0.11 | −1.83 ± 0.13 |
| 7 | DESIJ125640.82 +243304.2 | 39628362649440128 | 194.17009 | 24.55117 | 0.0488 | 21.80 | 0.02 ± 0.06 | 2.61 | 100 | 2.33e+04 ± 1.56e+03 | 7.30 ± 0.06 | 6.46 ± 0.25 | −1.11 ± 0.13 |
| 8 | DESIJ083738.65 +314918.2 | 39628516362290842 | 129.41106 | 31.82174 | 0.3502 | 22.38 | 0.00 | 2.62 | 3.54e+02 ± 3.24e+02 | 2.13e+04 ± 2.74e+03 | 7.32 ± 0.11 | 8.61 ± 0.25 | 0.36 ± 0.13 |
| 9 | DESIJ090136.94 +333611.4 | 39632950190410845 | 135.40392 | 33.60317 | 0.1162 | 22.00 | 0.05 ± 0.14 | 2.32 | 100 | 2.20e+04 ± 2.02e+03 | 7.34 ± 0.09 | 7.67 ± 0.22 | −0.46 ± 0.13 |
| 10 | DESIJ085602.66 +310445.1 | 39628500923056706 | 134.01108 | 31.07921 | 0.8221 | 22.90 | 0.02 ± 0.12 | 2.68 | 100 | 2.22e+04 ± 1.32e+03 | 7.34 ± 0.05 | 9.65 ± 0.39 | 1.37 ± 0.13 |

**Note.** Column (1): sequence number; column (2): defined XMPG name; column (3): DESI target ID; columns (4) and (5): R.A. and decl. in J2000, respectively; column (6): spectroscopic redshift; column (7): r-band magnitude; column (8): gas-phase extinction; column (9): equivalent widths of [O III]λ5007; column (10): electronic density; column (11): electronic temperature for [O III]; column (12): oxygen abundance; column (13): stellar mass; column (14): logarithmic SFR. Only the top 10 of the whole XMPGs sorted by increasing 12+log(O/H) are listed.

(This table is available in its entirety in machine-readable form.)

**Table 3**
Properties of Our XMPG Candidates

| No. (1) | Name (2) | TARGETID (3) | R.A. (deg) (4) | Decl. (deg) (5) | Redshift (6) | $r$ (mag) (7) | $E(B-V)$ (mag) (8) | EW([O III]) (Å) (9) | $N_e$ (cm$^{-3}$) (10) | $T_e$([O III]) (K) (11) | 12+log (O/H) (12) | $M_*$ log($M_\odot$) (13) | log(SFR) log($M_\odot$ yr$^{-1}$) (14) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | DESIJ140344.61 +484557.5 | 39633228671224911 | 210.93585 | 48.76599 | 0.0025 | 16.87 | 0.00 | 2.06 | 3.19e+02 ± 2.49e+02 | 1.96e+04 ± 3.27e+03 | 7.53 ± 0.17 | 6.53 ± 0.23 | −2.15 ± 0.13 |
| 2 | DESIJ072211.92 +415915.2 | 39633112598056268 | 110.54969 | 41.98758 | 0.0438 | 20.65 | 0.00 | 2.33 | 1.95e+02 ± 1.31e+02 | 1.88e+04 ± 3.40e+03 | 7.54 ± 0.19 | 6.70 ± 0.23 | −0.82 ± 0.13 |
| 3 | DESIJ140237.80 +354229.6 | 39632996071902273 | 210.65748 | 35.70824 | 0.1170 | 24.93 | 0.00 | 2.75 | 100 | 1.96e+04 ± 3.32e+03 | 7.54 ± 0.16 | 6.76 ± 0.51 | −1.46 ± 0.13 |
| 4 | DESIJ143510.37 +030408.5 | 39627860876461065 | 218.79323 | 3.06904 | 0.4578 | 23.01 | 0.01 ± 0.18 | 2.89 | 6.28e+02 ± 5.02e+02 | 1.92e+04 ± 3.06e+03 | 7.54 ± 0.15 | 8.27 ± 0.33 | 0.67 ± 0.13 |
| 5 | DESIJ140128.57 +041520.0 | 39627890895095223 | 210.36905 | 4.25556 | 0.1170 | 23.33 | 0.00 | 2.22 | 2.90e+02 ± 2.06e+02 | 1.86e+04 ± 2.74e+03 | 7.55 ± 0.16 | 7.36 ± 0.27 | −1.14 ± 0.13 |
| 6 | DESIJ084809.70 +313439.3 | 39628511245239168 | 132.04043 | 31.57760 | 0.8516 | 23.49 | 0.11 ± 0.19 | 3.00 | 6.30e+02 ± 4.87e+02 | 1.89e+04 ± 2.45e+03 | 7.56 ± 0.14 | 9.38 ± 0.33 | 1.37 ± 0.13 |
| 7 | DESIJ081428.78 +342638.3 | 39632970113354267 | 123.61991 | 34.44398 | 0.3236 | 22.32 | 0.00 | 2.85 | 8.51e+02 ± 7.19e+02 | 2.00e+04 ± 3.08e+03 | 7.56 ± 0.15 | 8.51 ± 0.24 | 0.28 ± 0.13 |
| 8 | DESIJ160521.36 +422914.4 | 39633123201254715 | 241.33899 | 42.48735 | 0.0775 | 18.71 | 0.10 ± 0.05 | 1.92 | 7.56e+01 ± 7.58e+01 | 1.84e+04 ± 3.42e+03 | 7.56 ± 0.22 | 8.85 ± 0.18 | 0.30 ± 0.13 |
| 9 | DESIJ144707.28 +062409.9 | 39627945244886188 | 221.78032 | 6.40276 | 0.0726 | 20.12 | 0.01 ± 0.06 | 2.25 | 100 | 1.94e+04 ± 2.95e+03 | 7.57 ± 0.15 | 7.94 ± 0.22 | −0.36 ± 0.13 |
| 10 | DESIJ122340.63- 060720.9 | 39627643062062197 | 185.91928 | −6.12249 | 0.2067 | 23.43 | 0.00 | 2.32 | 2.77e+02 ± 2.61e+02 | 2.07e+04 ± 2.81e+03 | 7.57 ± 0.14 | 8.25 ± 0.21 | −0.85 ± 0.13 |

**Note.** Only the top 10 of the entire XMPGs sorted by increasing 12+log(O/H) are listed.

(This table is available in its entirety in machine-readable form.)

## ORCID iDs

Hu Zou ⓘ https://orcid.org/0000-0002-6684-3997
Amélie Saintonge ⓘ https://orcid.org/0000-0003-4357-3450
Dirk Scholte ⓘ https://orcid.org/0000-0002-6867-1244
John Moustakas ⓘ https://orcid.org/0000-0002-2733-4559
Weijian Guo ⓘ https://orcid.org/0000-0001-9457-0589
S. Ahlen ⓘ https://orcid.org/0000-0001-6098-7247
A. de la Macorra ⓘ https://orcid.org/0000-0002-1769-1640
J. E. Forero-Romero ⓘ https://orcid.org/0000-0002-2890-3725
S. Gontcho A Gontcho ⓘ https://orcid.org/0000-0003-3142-233X
M. Landriau ⓘ https://orcid.org/0000-0003-1838-8528
L. Le Guillou ⓘ https://orcid.org/0000-0001-7178-8868
M. Manera ⓘ https://orcid.org/0000-0003-4962-8934
A. Meisner ⓘ https://orcid.org/0000-0002-1125-7384
R. Miquel ⓘ https://orcid.org/0000-0002-6610-4836
Jundan Nie ⓘ https://orcid.org/0000-0001-6590-8122
M. Rezaie ⓘ https://orcid.org/0000-0001-5589-7116
E. Sanchez ⓘ https://orcid.org/0000-0002-9646-8198
H. Seo ⓘ https://orcid.org/0000-0002-6588-3508
G. Tarlé ⓘ https://orcid.org/0000-0003-1704-0781
Zhimin Zhou ⓘ https://orcid.org/0000-0002-4135-0977
Siwei Zou ⓘ https://orcid.org/0000-0002-3983-6484

## References

Alexander, D. M., Davis, T. M., Chaussidon, E., et al. 2023, AJ, 165, 124
Allende Prieto, C., Cooper, A. P., Dey, A., et al. 2020, RNAAS, 4, 188
Aller, L. H. 1984, Physics of Thermal Gaseous Nebulae (Dordrecht: Reidel)
Amorín, R., Sommariva, V., Castellano, M., et al. 2014, A&A, 568, L8
Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140
Asplund, M., Amarsi, A. M., & Grevesse, N. 2021, A&A, 653, A141
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Brodzeller, A., Dawson, K., Bailey, S., et al. 2023, AJ, 166, 66
Brown, W. R., Kewley, L. J., & Geller, M. J. 2008, AJ, 135, 92
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, MNRAS, 399, 1191
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Chabrier, G. 2003, PASP, 115, 763
Charlot, S., & Fall, S. M. 2000, ApJ, 539, 718
Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2023, ApJ, 944, 107
Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, MNRAS, 358, 363
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Cooper, A. P., Koposov, S. E., Allende Prieto, C., et al. 2023, ApJ, 947, 37
Curti, M., D'Eugenio, F., Carniani, S., et al. 2023, MNRAS, 518, 425
Darragh-Ford, E., Wu, J. F., Mao, Y.-Y., et al. 2023, ApJ, 954, 149
Davé, R., Finlator, K., & Oppenheimer, B. D. 2011, MNRAS, 416, 1354
Dekel, A., Sarkar, K. C., Birnboim, Y., Mandelker, N., & Li, Z. 2023, MNRAS, 523, 3201
Denicoló, G., Terlevich, R., & Terlevich, E. 2002, MNRAS, 330, 69
DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, AJ, 164, 207
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2023a, arXiv:2306.06307
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2023b, arXiv:2306.06308
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv:1611.00036
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016b, arXiv:1611.00037
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172
Ellison, S. L., Patton, D. R., Simard, L., & McConnachie, A. W. 2008, ApJL, 672, L107
Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ, 644, 813
Filho, M. E., Sánchez Almeida, J., Muñoz-Tuñón, C., et al. 2015, ApJ, 802, 82
Froese Fischer, C., & Tachiev, G. 2004, ADNDT, 87, 1
Gao, Y., Bao, M., Yuan, Q., et al. 2018, ApJ, 869, 15
Gao, Y.-L., Lian, J.-H., Kong, X., et al. 2017, RAA, 17, 041
Gburek, T., Siana, B., Alavi, A., et al. 2019, ApJ, 887, 168
Guy, J., Bailey, S., Kremin, A., et al. 2023, AJ, 165, 144

Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023, AJ, 165, 253
Herrera-Camus, R., Fisher, D. B., Bolatto, A. D., et al. 2012, ApJ, 752, 112
Hsyu, T., Cooke, R. J., Prochaska, J. X., & Bolte, M. 2018, ApJ, 863, 134
Hsyu, T., Cooke, R. J., Prochaska, J. X., & Bolte, M. 2020, ApJ, 896, 77
Hummer, D. G., & Storey, P. J. 1987, MNRAS, 224, 801
Isobe, Y., Ouchi, M., Suzuki, A., et al. 2022, ApJ, 925, 111
Izotov, Y. I., & Thuan, T. X. 2004, ApJ, 602, 200
Juneau, S., Bournaud, F., Charlot, S., et al. 2014, ApJ, 788, 88
Juneau, S., Dickinson, M., Alexander, D. M., & Salim, S. 2011, ApJ, 736, 104
Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055
Kennicutt, R. C. J. 1998, ARA&A, 36, 189
Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183
Kewley, L. J., Heisler, C. A., Dopita, M. A., & Lumsden, S. 2001, ApJS, 132, 37
Kisielius, R., Storey, P. J., Ferland, G. J., & Keenan, F. P. 2009, MNRAS, 397, 903
Kojima, T., Ouchi, M., Rauch, M., et al. 2020, ApJ, 898, 142
Kunth, D., & Östlin, G. 2000, A&ARv, 10, 1
Lamareille, F. 2010, A&A, 509, A53
Lamareille, F., Mouhcine, M., Contini, T., Lewis, I., & Maddox, S. 2004, MNRAS, 350, 396
Lan, T.-W., Tojeiro, R., Armengaud, E., et al. 2023, ApJ, 943, 68
Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv:1308.0847
Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119
Luridiana, V., Morisset, C., & Shaw, R. A. 2015, A&A, 573, A42
Ly, C., Malkan, M. A., Nagao, T., et al. 2014, ApJ, 780, 122
Ly, C., Malkan, M. A., Rigby, J. R., & Nagao, T. 2016, ApJ, 828, 67
Ly, C., Rigby, J. R., Cooper, M., & Yan, R. 2015, ApJ, 805, 45
Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2016, MNRAS, 456, 2140
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, MNRAS, 408, 2115
Matsumoto, A., Ouchi, M., Nakajima, K., et al. 2022, ApJ, 941, 167
Méndez-Delgado, J. E., Esteban, C., García-Rojas, J., Kreckel, K., & Peimbert, M. 2023, Natur, 618, 249
Miller, T. N., Doel, P., Gutierrez, G., et al. 2023, arXiv:2306.06310
Morales-Luis, A. B., Sánchez Almeida, J., Aguerri, J. A. L., & Muñoz-Tuñón, C. 2011, ApJ, 743, 77
Moustakas, J., Kennicutt, R. C. J., Tremonti, C. A., et al. 2010, ApJS, 190, 233
Moustakas, J., Zaritsky, D., Brown, M., et al. 2011, arXiv:1112.3300
Myers, A. D., Moustakas, J., Bailey, S., et al. 2023, AJ, 165, 50
Nicholls, D. C., Dopita, M. A., Sutherland, R. S., et al. 2014, ApJ, 786, 155
Olmo-García, A., Sánchez Almeida, J., Muñoz-Tuñón, C., et al. 2017, ApJ, 834, 181
Plat, A., Charlot, S., Bruzual, G., et al. 2019, MNRAS, 490, 978
Raichoor, A., Eisenstein, D. J., Karim, T., et al. 2020, RNAAS, 4, 180
Raichoor, A., Moustakas, J., Newman, J. A., et al. 2023, AJ, 165, 126
Ruiz-Macias, O., Zarrouk, P., Cole, S., et al. 2020, RNAAS, 4, 187
Sacchi, E., Annibali, F., Cignoni, M., et al. 2016, ApJ, 830, 3
Sánchez Almeida, J., Elmegreen, B. G., Muñoz-Tuñón, C., et al. 2015, ApJL, 810, L15
Sánchez Almeida, J., Pérez-Montero, E., Morales-Luis, A. B., et al. 2016a, ApJ, 819, 110
Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2016b, ApJL, 825, L23
Schlafly, E. F., Kirkby, D., Schlegel, D. J., et al. 2023, AJ, 166, 259
Silber, J. H., Fagrelius, P., Fanning, K., et al. 2023, AJ, 165, 9
Storey, P. J., Sochi, T., & Badnell, N. R. 2014, MNRAS, 441, 3028
Tayal, S. S., & Zatsarinny, O. 2010, ApJS, 188, 32
Thuan, T. X., Sauvage, M., & Madden, S. 1999, ApJ, 516, 783
Torrey, P., Vogelsberger, M., Hernquist, L., et al. 2018, MNRAS, 477, L16
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295
Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
Xu, Y., Ouchi, M., Rauch, M., et al. 2022, ApJ, 929, 134
Yang, G., Boquien, M., Brandt, W. N., et al. 2022, ApJ, 927, 192
Yang, G., Boquien, M., Buat, V., et al. 2020, MNRAS, 491, 740
Yang, H., Malhotra, S., Rhoads, J. E., & Wang, J. 2017, ApJ, 847, 38
Yèche, C., Palanque-Delabrouille, N., Claveau, C.-A., et al. 2020, RNAAS, 4, 179
Zahid, H. J., Geller, M. J., Kewley, L. J., et al. 2013, ApJL, 771, L19
Zhou, R., Dey, B., Newman, J. A., et al. 2023, AJ, 165, 58
Zhou, R., Newman, J. A., Dawson, K. S., et al. 2020, RNAAS, 4, 181
Zou, H., Sui, J., Xue, S., et al. 2022, RAA, 22, 065001
Zou, H., Zhou, X., Fan, X., et al. 2017, PASP, 129, 064101