J-PLUS: Uncovering a large population of extreme [OIII] emitters in the local Universe

A. Lumbreras-Calle[1],[*], C. López-Sanjuan[1], D. Sobral[2], J. A. Fernández-Ontiveros[1], J. M. Vílchez[3],
A. Hernán-Caballero[1], M. Akhlaghi[1], L. A. Díaz-García[3], J. Alcaniz[4], R. E. Angulo[5,6], A. J. Cenarro[1],
D. Cristóbal-Hornillos[1], R. A. Dupke[4,7,8], A. Ederoclite[9], C. Hernández-Monteagudo[10,11], A. Marín-Franch[1],
M. Moles[1], L. Sodré Jr.[9], H. Vázquez Ramió[1], and J. Varela[1]

[1] Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Unidad Asociada al CSIC, Plaza San Juan 1, 44001 Teruel, Spain
[2] Department of Physics, Lancaster University, Lancaster, LA1 4YB, UK
[3] Instituto de Astrofísica de Andalucía (IAA-CSIC), P.O. Box 3004, 18080 Granada, Spain
[4] Observatório Nacional - MCTI (ON), Rua Gal. José Cristino 77, São Cristóvão, 20921-400 Rio de Janeiro, Brazil
[5] Donostia International Physics Centre (DIPC), Paseo Manuel de Lardizabal 4, 20018 Donostia-San Sebastián, Spain
[6] IKERBASQUE, Basque Foundation for Science, 48013, Bilbao, Spain
[7] University of Michigan, Department of Astronomy, 1085 South University Ave., Ann Arbor, MI 48109, USA
[8] University of Alabama, Department of Physics and Astronomy, Gallalee Hall, Tuscaloosa, AL 35401, USA
[9] Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, 05508-090 São Paulo, Brazil
[10] Instituto de Astrofísica de Canarias, La Laguna, 38205, Tenerife, Spain
[11] Departamento de Astrofísica, Universidad de La Laguna, 38206, Tenerife, Spain

ABSTRACT

Context. Over the past decades, several studies have discovered a population of galaxies undergoing very strong star formation events, called extreme emission line galaxies (EELGs).

Aims. In this work, we exploit the capabilities of the Javalambre Photometric Local Universe Survey (J-PLUS), a wide field multifilter survey, with 2000 square degrees of the northern sky already observed. We use it to identify EELGs at low redshift by their [OIII]5007 emission line. We intend to provide with a more complete, deep, and less biased sample of local EELGs.

Methods. We select objects with an excess of flux in the J-PLUS mediumband *J*0515 filter, which covers the [OIII] line at z<0.06. We remove contaminants (stars and higher redshift systems) using J-PLUS and WISE infrared photometry, with SDSS spectra as a benchmark. We perform spectral energy distribution fitting to estimate the physical properties of the galaxies: line fluxes, equivalent widths (EWs), masses, stellar population ages, etc.

Results. We identify 466 EELGs at z < 0.06 with [OIII] EW over 300 Å and *r*-band magnitude below 20, of which 411 were previously unknown. Most show compact morphologies, low stellar masses ($\log(M_\star/M_\odot) \sim 8.13^{+0.61}_{-0.58}$), low dust extinction ($E(B-V) \sim 0.1^{+0.2}_{-0.1}$), and very young bursts of star formation ($3.0^{+2.7}_{-2.0}$ Myr). Our method is up to $\sim$ 20 times more efficient detecting EELGs per Mpc^3 than broadband surveys, and as complete as magnitude-limited spectroscopic surveys (while reaching fainter objects). The sample is not directly biased against strong Hα emitters, in contrast with works using broadband surveys.

Conclusions. We demonstrate the capability of J-PLUS to identify, following a clear selection process, a large sample of previously unknown EELGs showing unique properties. A fraction of them are likely similar to the first galaxies in the Universe, but at a much lower redshift, which makes them ideal targets for follow-up studies.

Key words. galaxies: starburst - galaxies: star formation - galaxies: dwarf - galaxies: photometry - galaxies: ISM

1. Introduction

Galaxies undergoing strong events of star formation and presenting compact morphologies have been identified and analyzed since the middle of the XXth century, with pioneering works by Haro (1956), Zwicky (1966), and Markarian (1967). One of the most fruitful developments in the detection of this kind of galaxies was the use of an objective prism to identify strong emission lines (Markarian 1967). Early analysis of this class of galaxies noticed their usually compact morphologies and blue colors, with their spectra resembling those of galactic HII regions, thus the label "HII galaxies" (Sargent & Searle 1970; Melnick et al. 1985). Low metallicities were also identified, as well as recent enhancement of star formation (Searle & Sargent 1972).

With the advent of modern CCD detectors and wide field surveys, new windows into the analysis of these objects were opened. An observational alternative for the identification of emission line galaxies (ELGs) is the use of multi-band surveys, with some examples including CADIS (Meisenheimer et al. 1998), COMBO-17 (Wolf et al. 2001, 2003), HiZELS (Geach et al. 2008; Sobral et al. 2015), ALHAMBRA (Moles et al. 2008), MUSYC (Cardamone et al. 2010), SHARDS (Pérez-González et al. 2013), COSMOS mediumband (Taniguchi et al. 2015), PAU (Benítez et al. 2009), SC4K (Sobral et al. 2018), J-PLUS (Cenarro et al. 2019), S-PLUS (Mendes de Oliveira et al. 2019) and J-PAS (Benitez et al. 2014). Using a set of several narrowband or mediumband filters, sometimes including also broadband ones, these surveys detect easily emission lines by identifying an excess of flux in one of the bands (e.g. Hip-

pelein et al. 2003; Maier et al. 2003; Gronwall et al. 2007; Sobral et al. 2013, 2015; Cava et al. 2015; Lumbreras-Calle et al. 2019a; Spinoso et al. 2020; Vilella-Rojo et al. 2021). While usually providing smaller fields of view and less wavelength coverage than broadband or objective prism surveys, they can reach deeper magnitudes than the former and higher spectral resolution than the latter. These surveys have yielded large samples of ELGs, sometimes identifying simultaneously different emission lines and covering a wide redshift range.

Focusing on higher redshifts, the search for ELGs has been performed using deep, broadband surveys, specially those with data from the *Hubble* Space Telescope (HST), such as GOODS, COSMOS, or CANDELS (Kakazu et al. 2007; van der Wel et al. 2011; Maseda et al. 2018). They found an increase in the density of ELGs at higher redshifts, in coherence with the evolution of the cosmic star formation history (SFH, Madau & Dickinson 2014). In fact, at very high redshift ($z > 6$), typical low mass galaxies are expected to be extreme emission line galaxies (EELGs), playing a crucial role in the reionization of the Universe (Ouchi et al. 2009; Bouwens et al. 2015). From an observational point of view, space-based grism spectroscopy in some of the HST deep fields provides an even greater insight into the nature of EELGs (Pirzkal et al. 2013). This approach has only been feasible in small fields, which is enough for higher redshifts but insufficient for local Universe studies. Upcoming survey telescopes such as Euclid, the Nancy Grace Roman Telescope and the Chinese Space Station Telescope will extend this grism analysis to fields covering thousands of square degrees. It is therefore vital to gather samples of EELGs that can be used as a reference for these kind of surveys, testing their limitations and providing a readily available reference sample.

The precise definition of EELGs varies in different works in the literature. It is very often described as a threshold in the restframe equivalent width (EW) of the [OIII]5007 line, but values range from 100 Å in Amorín et al. (2015) and Pérez-Montero et al. (2020), to 300 Å in Jiang et al. (2019), with several surveys providing most objects above 500 Å (van der Wel et al. 2011; Yang et al. 2017; Maseda et al. 2018).

Recently, several projects have aimed at detecting samples of EELGs at low redshift using photometric surveys, such as Cardamone et al. (2009), Yang et al. (2017) and Senchyna & Stark (2019). They have demonstrated the power of broadband surveys to identify this kind of galaxies with relatively high purity, while covering much wider areas than previous surveys, allowing the study of rare populations. Studies using narrowband filters have also proven useful in the detection of ELGs over relatively modest fields (Kellar et al. 2012; Salzer et al. 2020) or brighter galaxies in very wide fields (Cook et al. 2019).

In addition to photometric surveys, the extensive database provided by the SDSS spectra has been thoroughly explored in the search of EELGs (Izotov et al. 2011) or extremely metal poor galaxies (Sánchez Almeida et al. 2016). The physical information for individual galaxies accessible through this database is unparalleled, showing for instance how these local samples can serve as analogs for the very high redshift systems responsible of the reionization of the Universe (Izotov et al. 2021). Nevertheless, a drawback of this data set is its poorly defined selection function, with galaxies often observed in surveys targeting other type of sources (such as quasars). As a result, no meaningful study regarding e.g. number densities can be performed, except for the very bright end, where completeness is high (Strauss et al. 2002). A very deep spectroscopic survey (with exposure times of ~ 10 h) that does not suffer from these selection biases is the

Table 1. J-PLUS filter system. The filters g, r, i, and z match their namesakes in SDSS.

| Filter	Central Wavelength [Å]	FWHM [Å]
u	3485	508
$J0378$	3785	168
$J0395$	3950	100
$J0410$	4100	200
$J0430$	4300	200
g	4803	1409
$J0515$	5150	200
r	6254	1388
$J0660$	6600	138
i	7668	1535
$J0861$	8610	400
z	9114	1409

MUSE *Hubble* ultra deep field (Bacon et al. 2017). Using the MUSE integral field unit they detected a large sample of ELGs, some extremely faint and undetected in the HST data. Nevertheless, this work covered a small area (~ 10 arcmin^2), insufficient for studies of uncommon sources at low redshift.

In this context, the Javalambre Photometric Local Universe Survey (J-PLUS, Cenarro et al. 2019) combines multiband observation with a very wide field, making it a unique tool for the identification of very rare emission line objects. In addition, as an imaging survey, it allows us to analyze the emission in the whole galaxy, not only the aperture where the spectra is extracted, a limitation found in fiber-fed spectroscopy surveys (such as SDSS, GAMA, or DESI). This photometric survey has already produced relevant results in the detection and analysis of extragalactic systems presenting emission lines in their spectra. In the local Universe, previous works by Logroño-García et al. (2019) and Vilella-Rojo et al. (2021) have analyzed Hα emitters and measured accurately the local star formation main sequence. In addition, Spinoso et al. (2020) performed a search for high-redshift quasars with Lyman α emission in J-PLUS, finding an unprecedented number of very high luminosity sources.

In the present work, we select low redshift EELGs in J-PLUS by identifying objects showing intense emission in the $J0515$ filter, as illustrated in Fig. 1. We cover a much wider area than previous multiband surveys, reaching deeper magnitudes than wide field spectroscopic surveys, and are able to identify emission lines with much more precision than broadband surveys. We aim at providing a new sample that will uncover many unclassified EELGs in the Local Univers, that can complement previous photometric and spectroscopic studies and serve as targets for follow-up observations.

The paper is structured as follows: In Sect. 2 we describe the J-PLUS survey and database, as well as our procedure for selecting the candidate sample and removing contaminants. In Sect. 3 we described the spectral energy distribution (SED) fitting analysis and its results, along with the final definition of the EELG sample and their visual morphologies. Finally, we compare these results with the available spectra for the sample. We discuss the results in Sect, 4, testing the sample selection, the number densities of EELGs we find, and reviewing the main physical properties of the EELG sample. Finally, in Sect, 5 we summarize our work and present the conclusions.

Throughout this paper, all quoted magnitudes are in the AB system, and all logarithms used are in base 10, and coordinates
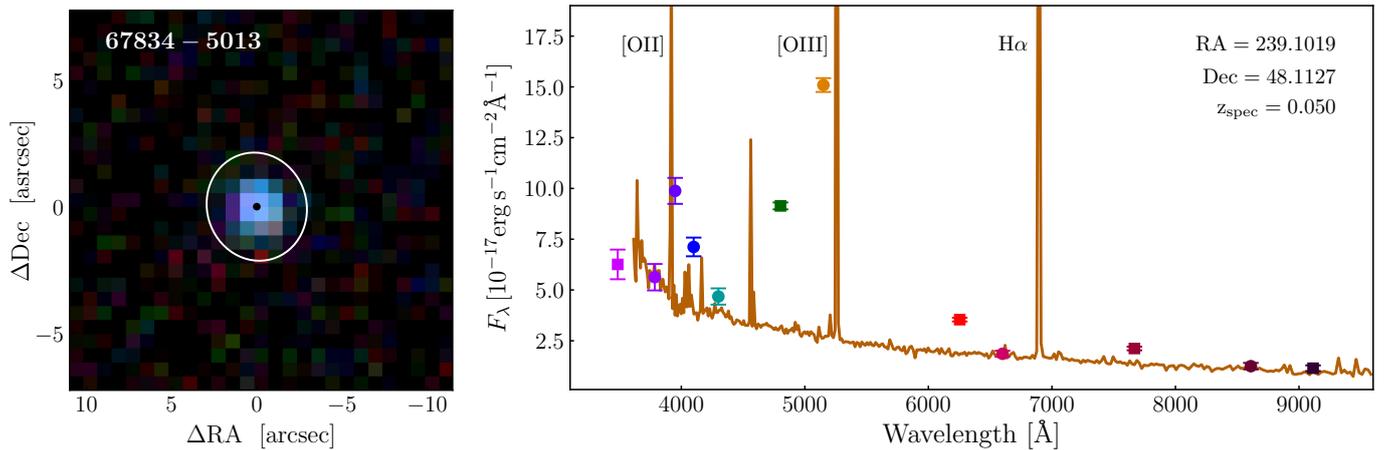
Fig. 1. Illustrative example of an extreme [OIII] emitter, the J-PLUS source 67834-5013 (RA = 239.1019, Dec = 48.1127, z_{spec} = 0.050). *Left panel:* Color composite of the galaxy, obtained from the *gri* J-PLUS images. The sky location of the source is shown as a black dot and the white ellipse indicate the three effective radius contour for the source. *Right panel:* J-PLUS twelve-band PSFCOR photometry of the galaxy. The squares show the five SDSS-like filters (*ugriz*), and circles the seven mediumband filters (*J*0378, *J*0395, *J*0410, *J*0430, *J*0515, *J*0660, and *J*0861). The solid line shows the spectra from SDSS with a downgraded resolution of $R \sim 180$ and normalized to the flux in the filter *J*0660. The location of the most prominent emission lines is marked: [OIII] (traced by *J*0515 and *g*), Hα (traced by *r*), and [OII] (traced by *J*0395).

refer to the 2000 equinox. All mentions to the [OIII] line, unless specified, refer to the [OIII]5007+4959 doublet. The median is used to show the typical value of a magnitude, while the upper and lower limits presented refer to the 16th and 84th percentiles of the distributions. We assumed a ΛCDM cosmology with with Ω_{Λ} = 0.7, Ω_M = 0.3, and H_0 = 70 km s^{-1} Mpc^{-1}. In the SED fitting we use the Salpeter (1955) initial mass function (IMF).

2. Database and sample selection

2.1. J-PLUS second data release

J-PLUS[1] is being conducted at the Observatorio Astrofísico de Javalambre (OAJ; Cenarro et al. 2014) using the 83 cm Javalambre Auxiliary Survey Telescope (JAST80) and T80Cam, a panoramic camera of 9.2k × 9.2k pixels that provides a 2 deg^2 field of view, with a pixel scale of 0.55 arsec pix^{-1} (Marín-Franch et al. 2015). The J-PLUS filter system is composed of twelve passbands (Table 1) spanning the full optical range (3 500 − 10 000 Å), with seven mediumband and five broadband filters. The J-PLUS observational strategy, image reduction, and main scientific goals are presented in Cenarro et al. (2019). The J-PLUS photometric calibration is described in López-Sanjuan et al. (2019) and López-Sanjuan et al. (2021).

This work is based on the second data release (DR2) of J-PLUS. It covers 2 176 square degrees (1 941 deg^2 after masking), with 1 088 individual images of 2 deg^2. The survey reaches limiting magnitudes ranging from 21.8 mag in *r* to 20.5 mag in *z*, with 21.0 mag in the *J*0515 filter (considering 5σ detection).

In order to create the galaxy sample we will describe in this work, we run queries using the Astronomical Data Query Language (ADQL) interface in the J-PLUS database. In the queries we use the PSFCOR photometry of the catalog, which is designed to better capture the colors of the objects rather than their total flux (Molino et al. 2019). Briefly, PSFCOR photometry is measured considering elliptical apertures with semi-major axis equal to the Kron radius (Kron 1980) in the reference band (*r*), half the size of the AUTO aperture. Then, the photometry in the rest of the bands is measured, correcting for the different point

spread functions (PSFs), to produce accurate colors (see Molino et al. 2019 and Hernán-Caballero et al. 2021 for more details). This photometry, while optimal for comparing fluxes in different bands, underestimates the total flux of galaxies with respect to the AUTO photometry by 0.5 mag in average (González Delgado et al. 2021). Therefore, in the rest of the analysis of this work (unless specified), we re-scale the PSFCOR photometry to the AUTO one using the ratio between AUTO and PSFCOR fluxes in the *r* band. In addition, we apply a galactic extinction correction, as provided in the J-PLUS database.

2.2. Sample selection

2.2.1. Selection of intense *J*0515 emitters

The first step in the process is to query the J-PLUS database for objects that show a large excess of emission in the *J*0515 filter, since all extreme [OIII]4959+5007 emitters between redshift 0.007 and 0.06 should show strong emission in this band (Fig. 1). The precise definition of "strong emission" will be discussed in Section 3.1.1, but given the width of the *J*0515 filter, in order to secure clear detections we will aim at selecting objects with EW(J0515)>200 Å. First, in order to compute the excess of emission corresponding to the line flux, we need to estimate the continuum underneath it. To perform a simple query, we assumed that the continuum flux (in erg s^{-1} cm^{-2}) at the wavelength of *J*0515 is roughly equivalent to the flux of the *r* band. This obviously depends on the slope of the SED and relative intensity of the Hα line, but it is enough to draw a first sample of emitters, of which we will later select the most extreme ones. A possible bias against extreme Hα emitters introduced by this assumptions will be addressed in Sect. 3.2. Specifically, we imposed the following condition:

$$\frac{F(J0515) - F(r)}{F(r)} > 1, \qquad (1)$$

which, if we assume that *r* traces the continuum, would imply that the emission lines that lie within *J*0515 have EW $\gtrsim$ 200 Å. A precise estimation of the EW will be performed in Section 3.1.2, when accurate estimations of the continuum have been

computed. In addition to this cut, we imposed other constrains to ensure the quality of the sample. We selected objects with $r < 22$ mag and signal-to-noise ratio (SNR) larger than 3 in the r, g, and $J0515$ bands. We also remove objects with a FLAGS parameter indicating bad photometric quality for those bands (saturation, proximity to a bright star or the edge of an image, etc.), except when the flag indicates the existence of close companions or deblending. We prefer to draw a broad sample in order to only later apply further cuts, motivated by the properties and limitations of the sample. The full ADQL query is reproduced in Appendix A.

Nevertheless, in order to analyze the selected objects, we separate them into four groups, considering their CLASS_STAR value ($>=0.5$ and <0.5) and their deblending flag (either no deblending or some deblending). This allows us to better identify different types of spurious detections that only appear in some of the groups, and to develop techniques to address them specifically. This first rough selection yields a sample of 30 336 objects with flux excess in the $J0515$ filter.

2.2.2. Cross-match with SDSS spectra

Multiple types of sources can show excess in the $J0515$ filter, such as ELGs, quasi-stellar objects (QSOs), foreground stars, or any object with spurious photometric measurements. In addition, ELGs and QSOs at different redshifts can show emission in this filter, but due to different emission lines. Since we are only interested in low redshift ELGs ($z < 0.06$, where it is the [OIII] line that causes the excess emission), we need to purge the sample from all other sources. To test the accuracy of the methods for removing contaminants, we cross matched the excess-flux sample with the 16th data release of the SDSS spectroscopic database (Ahumada et al. 2020). This survey covers a large fraction ($\sim$ 80%) of the J-PLUS DR2 footprint (more than any other spectroscopic survey) and provides a spectral classification for the sources. We consider a 3 arcseconds maximum separation in the cross-match, identifying 2 561 objects.

2.2.3. Removal of contaminants via J-PLUS photometry

When performing a visual inspection of the samples with blending flag, we identified several objects that are classified as galaxies but in fact were stars measured with large apertures. These stars are included in our sample because of a spurious deficit in r emission, which implies a large $J0515$ over r ratio. Therefore, removing candidates with low r-band flux compared to redder filters (such as $J0861$ and z) allows us to eliminate these detections and keep all other candidates. Additionally, this removes several QSOs with red ($r - z$) and ($r - J0861$) colors. Specifically, we rejected the 4 491 objects that fulfill ($r - z$) < 1 and ($r - J0861$) < 1. The selection was confirmed by checking that none of the 148 objects with SDSS spectra removed in this step is a low redshift galaxy (as well as inspecting a subsample of their images and SEDs).

Another subsample of contaminants are QSOs at z $\sim$ 1.7 with a prominent C III]1909 emission line, which at that redshift creates an excess of flux in the $J0515$ filter, mimicking a low redshift ELG. In those cases, the object often shows also excess in the $J0430$ filter, corresponding to the strong C IV 1549 line. In fact, these emission lines have been used to detect AGNs in narrowband surveys, for example in Stroe et al. (2017a,b). We therefore removed the 255 objects with ($J0515 - J0430$) > 0. Out of those, 46 candidates have SDSS spectra, without any low redshift galaxy among them.
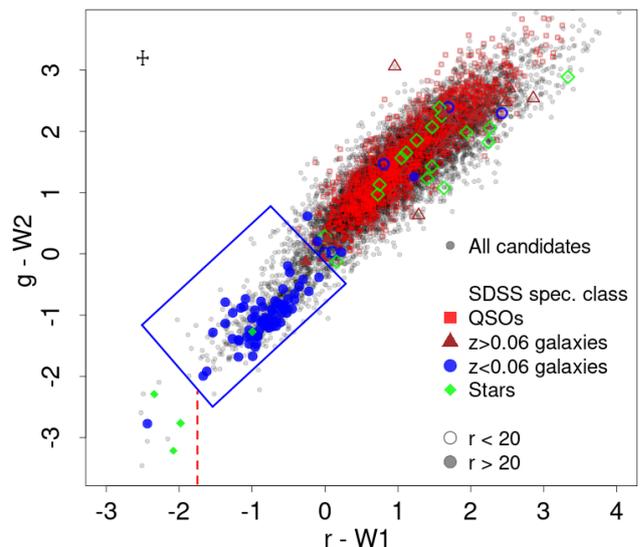
Fig. 2. Color-color diagram used to separate low redshift ELGs from other types of sources. The g and r data are taken from J-PLUS DR2, and the $W1$ and $W2$ from the *unWISE* catalog. All candidates from the sample with $J0515$ excess flux are shown in dark grey dots. Sources with SDSS spectra are shown in colors according to their nature (red squares for QSOs, brown triangles for galaxies at z>0.06, green diamonds for stars), and in filled symbols or outlines only if they are brighter or fainter than r= 20 mag. The objects selected as candidates to be EELGs are enclosed by the blue lines, while those to the left of the dashed red line are visually inspected to select only the extended ones as EELGs. In the top left corner the typical errors (for sources with r<20 mag) are shown.
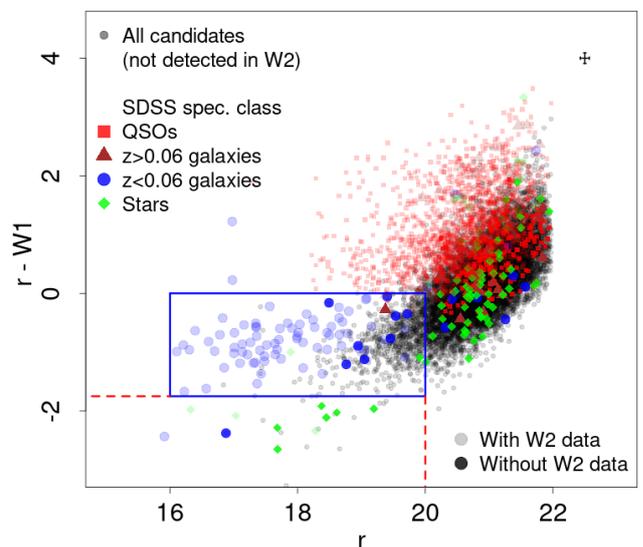
Fig. 3. Color-magnitude diagram used to separate low redshift ELGs from other types of sources, for those objects where the $W2$ flux was not available (grey dots). All sources with SDSS spectra are shown in colors. The color code for those is the same as the one used in Fig. 2, but in this case objects with $W2$ data are shown in a lighter shade. The objects selected as candidates to be EELGs are enclosed by the blue lines, while those to the left of the dashed red line are visually inspected to select only the extended ones as EELGs. In the top right corner the typical errors (for sources with r<20 mag) are shown.
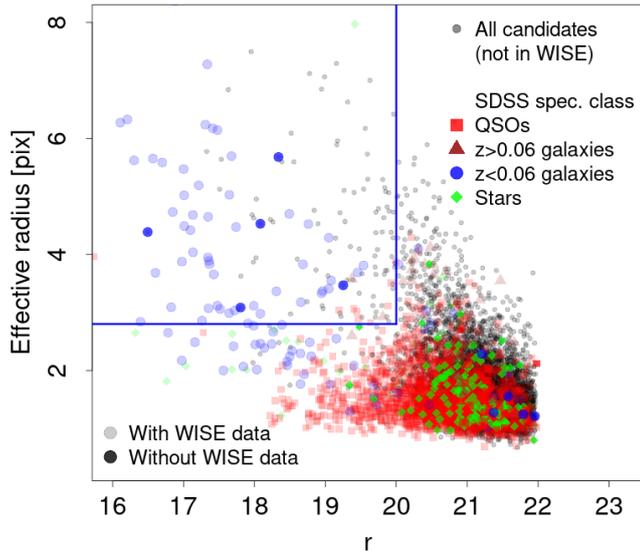
Fig. 4. Diagram showing the effective radius as a function of r magnitude, used to select candidates to be EELGs for those objects with excess flux in $J0515$ and no data in the $unWISE$ catalog (grey dots). All sources with SDSS spectra are shown in colors. The color code for those is the same as the one used in Fig. 2, but in this case objects with $W2$ data are shown in a lighter shade.

The sample with $J0515$ excess after this cleaning steps includes 25 590 sources.

2.2.4. Removal of contaminants: WISE photometry and object sizes

In an early analysis, it became clear that J-PLUS photometry alone was not enough to distinguish precisely between different types of sources, specially at fainter magnitudes. Following the method performed by Spinoso et al. (2020), we cross-matched our excess flux selection with the infrared (IR) data from the Wide-Field Infrared Survey Explorer (WISE) satellite. This mission performed a very wide survey in four IR bands, from 3.4 to 22 μm. We use the $unWISE$ catalog (Schlafly et al. 2019), which only has data for the bluest bands ($W1$ and $W2$), but provides a higher depth and spatial resolution. Out of the 25 590 sources with excess flux at this stage, 19 922 are present in the $unWISE$ catalog. We remove from that sample 71 stars that have been identified using Gaia DR2 data available in the J-PLUS tables.

After testing several alternatives, we found that the most precise way of separating the low redshift [OIII] emitters from the rest of the types of objects using $unWISE$ data was the $g - W2$ versus $r - W1$ diagram (Fig. 2). The separation in this diagram between low-redshift galaxies and quasars can be understood regarding the different components that dominate the infrared light emitted by these sources. In galaxies, the $W1$ and $W2$ filters cover a local minimum in their SEDs, at longer wavelength than the peak of emission from low-mass stars but at shorter wavelength than the peak of dust emission (see e.g. Boquien et al. 2019). This results in blue optical-$W1/W2$ colors. In low redshift quasars, the $W1/W2$ bands are dominated by dust emission, much brighter than their optical emission (dominated by the accretion disk), resulting in very red optical-$W1/W2$ colors. Even for higher redshift quasars, where the $W1/W2$ range is affected by a combination of disk and dust emission, the optical-$W1/W2$ colors remain slightly red or close to zero (see e.g. Hernán-

Caballero et al. 2016), thus different from low-redshift galaxies. We plot all sources with available $W1$ and $W2$ photometry in that diagram, indicating in different colors the type of source, for those that have available SDSS spectroscopy. It becomes clear that the low-z [OIII] emitters are clustered in a very specific region of Fig. 2, while the vast majority of stars, QSOs and higher redshift galaxies are excluded. We define a set of limits in Fig. 2 to select a sample of candidates to be EELGs, preserving a very high purity and completeness (a full account of those will be provided in Sec. 2.3). The selected objects are those located in the area that fulfills the following set of equations (blue lines in Fig. 2):

$$
\begin{aligned}
g - W2 &> 1.1 \cdot (r - W1) - 0.8 \\
g - W2 &< -1.3 \cdot (r - W1) - 0.19 \\
g - W2 &< 1.1 \cdot (r - W1) + 1.6 \\
g - W2 &> -1.3 \cdot (r - W1) - 4.5
\end{aligned}
\tag{2}
$$

Nevertheless, there remains a large fraction of objects without $W2$ photometry. From Fig. 2 it can be seen that the $r - W1$ color alone acts as a good discriminant on its own. When plotting $r - W1$ as a function of r magnitude (Fig. 3), we see that, for high values of r, the separation between AGNs and galaxies becomes less clear. In addition, there is a lack of spectroscopic data at $r > 20$ and $(r - W1) < 0$ which casts doubts about the nature of the objects with these characteristics.

Once again, we select the area in Fig. 3 where the low redshift galaxies are clustered, avoiding as many contaminants as possible. This results in the following conditions, that objects without $W2$ detection must fulfil:

$$
\begin{aligned}
(r - W1) &< 0 \\
(r - W1) &> -1.75 \\
r &< 20 \\
r &> 16
\end{aligned}
\tag{3}
$$

After applying both sets of criteria (from eq. 2 and 3), we are left with 1 447 candidates to be EELGs. If we applied to all sources with $unWISE$ data the selection in Eq. 3, the final sample would be slightly different: it would be missing 16 objects and including 78 new ones. Despite the small change, given the very clear separation and gap between spectroscopically confirmed QSOs and low redshift galaxies in Fig. 2, we consider that including the $g - W2$ color adds valuable information. Considering the uncertainties in the color and magnitude values, a small fraction of the selected galaxies could fall outside the selection areas (and vice versa). This effect is small in the color cuts, with ~ 10 % of selected galaxies less than 1σ away from the dividing line between low-z and quasar dominated regions in Fig. 2. Only 4% of the selected galaxies fulfil that condition for the $r - W1 = 0$ line in Fig. 3. The share of galaxies that could swap from the selected/rejected areas is higher for those with magnitudes close to $r = 20$ in Fig. 3, reaching values close to 20% of the selected sample. Nevertheless, since we do not expect a swift physical change at $r = 20$ mag, we consider that the contamination induced by this uncertainties is relatively small.

We also observe in both Figures 2 and 3 that there is an area with a small population of sources with very negative $r - W1$ color (marked in the figures with a dashed red line at $(r - W1)<1.75$, and in Fig. 3 also with a dashed line at $r=20$ mag). Some of the objects are spectroscopically confirmed to be low redshift galaxies, yet most are stars. In order to avoid missing a small but extreme population, we inspect the morphology
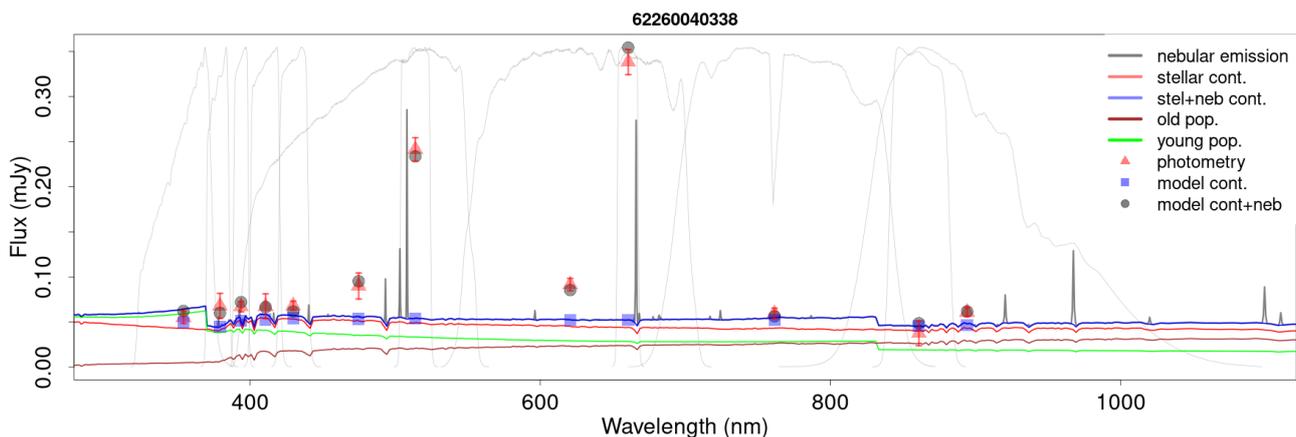
Fig. 5. Example of the J-PLUS SED and `CIGALE` fit of an EELG. Red triangles correspond to the observed photometry, grey dots to the `CIGALE` synthetic photometry, and blue squares to the `CIGALE` synthetic photometry only considering the stellar and nebular continua (not considering the emission lines). The thick lines represent the `CIGALE` synthetic spectrum, considering different components: the grey line takes into account the full model, while the blue does not take into account the emission lines. The red represents the stellar continuum, while the brown and green only consider the old and young stellar populations, respectively.

and SEDs of these galaxies (10 selected in Fig. 2 and 19 selected in Fig. 3). This subsample is heavily contaminated by stars, but still shows some promising candidates, so we add to the selected sample the 5 objects that show extended morphology (reaching 1 452 candidates up to this point).

There are nevertheless 5 668 objects that remain in the parent sample and do not have any counterpart in the *unWISE* catalog. We removed 60 stars identified with a cross-match to Gaia DR2. For this subsample without *WISE* data, we used a method of removing contaminants considering only the *r* band magnitude and the apparent effective radius of the object (R_{eff}), as measured with the `R_EFF` parameter in the J-PLUS database. This value is defined as the radius that encloses half of the total flux in the object in the *r* band, considering the run of the SExtractor software (Bertin & Arnouts 1996), performed in the J-PLUS data reduction pipeline (Cenarro et al. 2019). As shown in Fig. 4, at $R_{eff} > 2.8$ arcsec and $r < 20$ mag, the vast majority of sources with spectra are low redshift ELGs: all 5 sources with spectra but without *unWISE* counterpart fulfilling those conditions are in fact low redshift ELGs. Using R_{eff} to separate contaminants is less efficient than using *WISE* photometry (which is why we do not use it in the rest of the sample), but we still recover 92 extra candidates. At this stage, we have 1 544 candidates to be EELGs.

While this final selection biases our sample towards less compact EELGs, the effect is very small. Only 45 objects in the parent sample have no WISE data, $r < 20$ and $R_{eff} < 2.8$. Considering the selection of targets with WISE data, we would only expect around half of those objects to be low redshift galaxies. Therefore, missing those galaxies would mean a very small reduction in the sample of candidates ($\sim 1.4\%$), while the contamination rate would increase by ~ 30 %. Therefore no attempt is made to include objects without WISE detection and $R_{eff} < 2.8$.

2.2.5. Cosmic ray removal

In some cases, one of the (usually three) individual frames that are combined to form the final *J*0515 image of an object is affected by a cosmic ray, increasing significantly its flux. In a few of those instances, the data reduction process fails to remove the contaminated frame from the co-added frames, and the object appears in our selection since it apparently shows a strong ex-

cess in *J*0515. In order to remove these contaminants, we downloaded all frames that contribute to the *J*0515 image of each candidate, measured the object flux in each one, and checked if any individual frame deviates more than 20% from the median flux. This way we removed 21 objects, one of which is a spectroscopically confirmed star.

2.2.6. Photometric correction for deblended objects

In our sample, in order to reach the highest possible completeness, we have chosen to include objects with a blending flag. This allowed us to accurately analyze objects with relatively close companions without being contaminated by their emission, since the masks for each object have been separated. Nevertheless, this implies that for some extended galaxies, we only selected a small star-forming region, and the rest of the galaxy was considered as a separate object. This would bias our selection, identifying as "galaxies" what in fact are simply star-forming regions. In order to avoid this, we re-computed the photometry of all 375 selected objects that have a deblending flag. We used SExtractor in dual mode, with detection in the *r* band, tayloring the parameters of the code to adequate them to the sample of galaxies we were analyzing. In this case we did not perform any correction like the one applied to create the PSFCOR photometry, but given that the galaxies we are dealing with in this step are extended, such a correction would play a very minor role. After a visual inspection, 171 galaxies were confirmed to be better represented by our SExtractor run rather than the original J-PLUS photometry, and we kept these data during the rest of the analysis. During this stage, 30 galaxies were removed from the sample: 20 because they were spurious objects (almost all, spikes from bright stars), 1 was a repeated object, and in 9 cases neither the original photometry nor the new SExtractor run were able to properly measure them. For more details, see Appendix B.

2.3. Summary of the sample selection

With all these considerations, we have built a sample of 1493 sources that are candidates to be extreme [OIII] emitters at z < 0.06, by selecting objects with high *J*0515 over *r* ratio. We have removed the vast majority of contaminants from other types

of sources or higher redshifts galaxies, using J-PLUS and WISE photometry, and SDSS spectra. Out of 85 objects with available SDSS spectra in the candidate sample only three are not galaxies at z < 0.06: one is a star (likely selected due to an error in the photometric measurement) and two are galaxies at z = 0.072 and z = 0.123, with relatively high uncertainty in the J0515 flux. With those three interlopers, the purity our candidate selection is ~96%. Out of the 2560 objects with spectra in the excess flux sample, 89 are galaxies at $0.006 < z < 0.056$ with $r < 20$ mag. 82 of them were selected in the candidates sample, which translates into a completeness of ~ 92% in our selection, compared to the excess flux sample.

3. Results

To further confirm and characterize the nature of the sample of 1493 candidates to be EELGs at z < 0.06, we performed SED fitting using the `CIGALE` code (Noll et al. 2009; Boquien et al. 2016), which yields physical properties for the galaxies. We used the synthetic spectra to select only those systems with the highest EW values. Additionally, the galaxy images were inspected to classify their morphologies and the line fluxes estimated with `CIGALE` were compared to the available SDSS spectra.

3.1. SED fitting

To analyze in detail the physical properties of the sample of galaxies candidates to be EELGs, we performed SED fitting with `CIGALE`. It is a fast and flexible software implementing theoretical models for the different galaxy components. `CIGALE` creates a large grid of composite stellar populations based on single stellar population models and a variety of SFHs. It includes as well models for dust extinction (both for gas and stars), dust emission, and most importantly, nebular emission (both lines and continuum). The resulting grid of models are fitted to the photometric data, and the galaxy properties are estimated analyzing the posterior likelihood distribution, producing a best-fit model, and a bayesian estimate for each parameter.

3.1.1. CIGALE parameters

We use a simple model of two stellar populations with an exponentially declining SFH: an old population selected to represent the underlying galaxy, and a young one to reproduce the strong starburst causing the extreme emission lines. This type of modeling has been used in the literature, specially when dealing with galaxies with recent events of star formation (e.g. Nilsson et al. 2011; Catalán-Torrecilla et al. 2015; López-Sanjuan et al. 2017; Lumbreras-Calle et al. 2019a; Arrabal Haro et al. 2020). The very high EW of the emission lines detected in the galaxies in our sample implies a very short timescale for the event, therefore a very young population is more appropriate than a continuous star-formation model. The existence of an old, underlying stellar population has been demonstrated in systems similar to ours, such as blue compact dwarf galaxies (Amorín et al. 2007; Amorín et al. 2009), and it has to be included to account for the majority of the stellar mass in the galaxies. In order to simplify the modelling, we fix the τ value of the exponential to 50 Myr for the old population and 1 Myr for the young one.

We use Bruzual & Charlot (2003) stellar population models, while the nebular emission `CIGALE` uses is predicted based on the photo-ionization models by Inoue (2011). We assume the Calzetti et al. (2000) extinction law for dust extinction.

Table 2. CIGALE parameters

Parameter	Values
Redshift	0 - 0.06 in 0.0025 intervals
Age young pop. [Myr]	1, 2, 3, 4, 6, 7, 9, 12
Age old pop. [Myr]	200, 500, 1000, 2000, 5000
Burst ratio	0.0005, 0.0025, 0.005, 0.01
	0.03, 0.05, 0.075, 0.15, 0.3
Metallicity [Z]	0.0001, 0.0004, 0.004, 0. 008, 0.02
Escape fraction	0, 0.2
log (U)	-4.0, -3.5, -3.0, -2.5, -2., -1.5
$E(B-V)_{young}$	0, 0.1, 0.2, 0.3, 0.5
$E(B-V)_{old}/(B-V)_{young}$	0.44, 0.9

In Table 2 we summarize the main free parameters used to fit the galaxy SEDs, and present a representative example in Fig. 5. For a more detailed account of the SED fitting process, including the values used for all the different parameters, see Appendix C.

3.1.2. SED Results

In Fig. 6 and Table 3 we show the distribution of the main stellar and nebular parameters that `CIGALE` delivers for our sample of 1493 EELG candidates. We have chosen to use the best-fit parameters instead of the Bayesian estimates in the `CIGALE` output for consistency, since `CIGALE` does not provide Bayesian-estimated for individual components in the synthetic spectra, and we rely on them to compute EWs. In addition, the Bayesian-estimated photometric redshifts are less accurate than the best fitting ones (see section 3.4).

It is important to note the limits on the accuracy of these SED fits. The main driver of them is to obtain accurate estimations of the flux of the most intense emission lines, as well as the underlying continuum flux. Additionally, we rely on the stellar mass estimates for comparison purposes, and in other parameters (burst age, extinction, metallicity) for broad properties of the sample, given the assumptions in the `CIGALE` run. The results of the fits cannot be analyzed in extreme detail, specially in some sparsely sampled, highly degenerated parameters like metallicity, log(U), escape fraction, etc. For an additional discussion on this topic, see appendix C.

Considering the best-fit parameters, the age of the star formation burst is very low, with almost no galaxies showing values higher than 8 Myr. This is consistent with the extreme EW values measured, since that parameter decreases very rapidly in the first Myrs of a star-formation burst (see e.g. Leitherer et al. 1999).

The metallicity and ionisation parameter of the sample are low. This is typical for low mass galaxies with strong bursts of star formation, specially at high redshift (Khostovan et al. 2015; Tang et al. 2021; Matthee et al. 2021). Nevertheless, the selection process for our sample, imposing a very high EW of [OIII], may prevent the selection of some extremely metal poor galaxies (see Sect. 4.1.3 for more details).

The typical extinction values that `CIGALE` derives for the main sample are low, which is consistent with previous results obtained for samples of low-mass star-forming galaxies (Garn & Best 2010; Duarte Puertas et al. 2017; Lumbreras-Calle et al. 2019a).

Even considering the strong burst of star formation that these galaxies are enduring, the mass ratio between old and young population is typically low ($\log(M_{you}/M_{old}) = -2.00^{+0.70}_{-0.60}$). This is however not surprising, since the mass-luminosity ratios vary
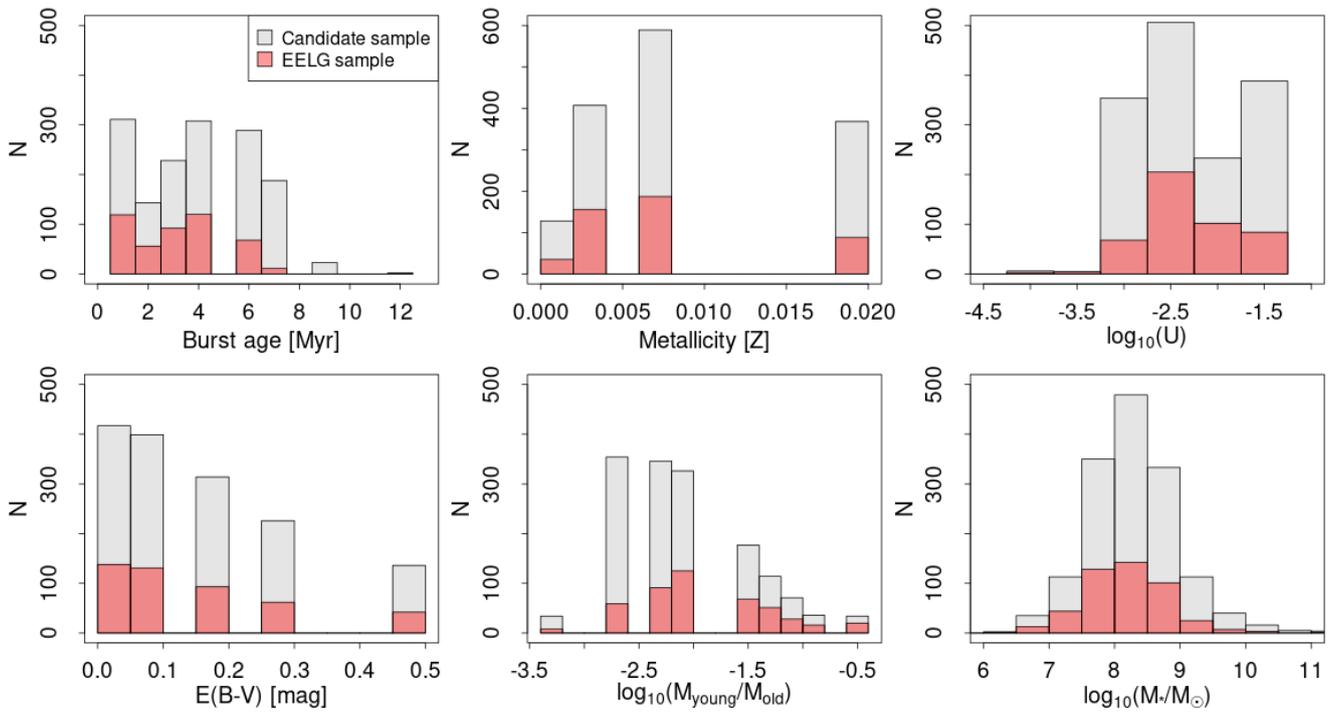
Fig. 6. Histograms showing the best fitting values for the J-PLUS galaxies, derived using the SED fitting software `CIGALE`. We show in grey the results for the 1493 galaxies in the candidate sample and in red the 466 in the EELG sample (EW[OIII]>300 Å). From left to right and top to bottom, we show the age of the young population, the stellar metallicity, the ionization parameter, the colour excess, the mass ratio between young and old populations, and the total stellar mass. The gaps observed in some of the histograms are due to the sampling in the parameters (see Table 2 and Appendix C).

strongly between old and very young populations. Therefore, while the old population dominates the total mass of the galaxy, a relatively small (in mass) young population can have strong effects in the integrated photometry of a low-mass galaxy. The total stellar masses of the galaxies in the sample are not constrained by input parameters, and they span most of the range of what is usually considered dwarf galaxies, with $\log(M_\star/M_\odot) \sim 6.5 - 9.5$.

Two of the main parameters in the analysis of ELGs are the flux and EW of the emission lines (in our case specially the [OIII] line, used for the sample selection). We focus in the $J0515$ filter, where most of the line flux comes from the [OIII]4959+[OIII]5007 lines. To comput them, we first convolve the synthetic stellar and nebular continuum derived by `CIGALE` (blue line in Fig. 5) through the transmission curve of the $J0515$ filter, to obtain an estimation of the continuum in that filter ($F_{\mathrm{cont.}}$, blue squares in Fig. 5). Then, using the measured flux in that filter (F_{J0515}, red triangles in Fig. 5) and the classical formula (assuming a flat continuum within the filter, and considering fluxes in units of erg/s/cm²/Å), we obtained the EW as

$$\mathrm{EW[\mathring{A}]} = \frac{F_{J0515} - F_{\mathrm{cont.}}}{F_{\mathrm{cont.}}}\Delta = \frac{F_{\mathrm{em}}}{F_{\mathrm{cont.}}}\Delta, \qquad (4)$$

where Δ is the width of the filter, 200 Å, and ΔF_{em} is the line flux. In Fig. 7 we show the histogram of EW for the $J0515$ filter, separating the galaxies with SNR > 3 in the $J0515$ EW.

3.1.3. Comparison between `CIGALE` and J-PLUS fluxes

Given the redshift range covered, $0.0075 < z < 0.06$, for the majority of the sample the $J0515$ filter is affected by both the
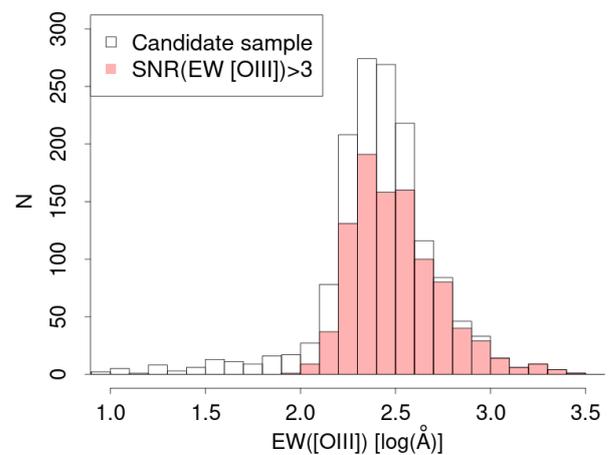
Fig. 7. Histogram of EW([OIII]), measured using the J-PLUS $J0515$ photometry and the `CIGALE` fits for the continuum. The white filled histogram represents the whole sample of candidates, while the red filled one shows only the galaxies with SNR>3 in the EW([OIII]) value.

[OIII]5007 and [OIII]4959 emission lines. According to the photometric redshifts obtained from `CIGALE`, for a small fraction of galaxies (~ 18 %, at the higher redshift end), the [OIII]5007 line lies in a low transmission region of the $J0515$ filter (or even entirely out of it). The Hβ line enters in the filter wavelength range in almost half of the sample, while at the very low redshift end, some galaxies (~ 5 %) lack the contribution of the [OIII]4959 line. Nevertheless, the limited photometric redshift precision (see Sect. 3.4) prevents us from providing a detailed account of this distribution of emission lines, galaxy by galaxy.

Table 3. Main `CIGALE`-derived properties of the J-PLUS selected EELGs. The complete table is available online; only the first row is shown here as guidance.

ID	Redshift	Mass	Mass ratio	Age young	Metallicity	log(U)	E(B-V)	$F_{[OIII]5007}$	$F_{H\alpha}$
		$[log(M_\odot)]$	$[log(\frac{M_{you}}{M_{old}})]$	[Myr]	[Z]		mag	$log(erg/s/cm^2)$	
63414026530	0.0075	6.96	-1.3	4.0	0.004	-2.0	0.10	-12.88	-13.16

Table 4. Main photometric properties of the J-PLUS selected EELGs. The complete table is available online; only the first row is shown here as guidance.

ID	r	EW [OIII]	EW $H\alpha$
	mag	Å	Å
63414026530	17.80 ± 0.04	505 ± 12	480 ± 100

We can still study the whole sample comparing the emission line flux in the whole filter, estimated using the $J0515$ photometry and the `CIGALE`-derived continuum value, with the fluxes computed by `CIGALE` for each emission line. The same can be done for the r broadband filter, which is significantly contaminated by the $H\alpha$ emission line. In Figs. 8 and 9 we show the logarithm of line fluxes derived directly from the J-PLUS photometry, as a function of the `CIGALE`-estimated individual line fluxes. The integrated fluxes are clearly dominated by the brightest emission lines in several filters. In the r filter (Fig. 9), only taking into account the $H\alpha$ line is enough to recover a very good one-to-one relationship, with a small offset ($\sim$ -0.04 dex) and low scatter ($1\sigma \sim 0.13$). If we include in the analysis the [NII]6584 flux, the offset is greatly reduced (down lo -0.008 dex), with a similarly low scatter ($1\sigma \sim 0.12$ dex). Nevertheless, given the small impact of [NII]6584, we consider that the line emission in the r filter is mostly dominated by $H\alpha$. For the $J0515$ filter (Fig. 8), taking into account the [OIII]4959 and [OIII]5007 lines for all galaxies provides a reasonably good fit, with a negligible offset ($\sim$ -0.007 dex) and and very low scatter ($\sim$ 0.096 dex). Nevertheless, in the highest redshift range of our sample (z≥0.05), the [OIII]5007 line falls in a very low transmission region of the $J0515$ filter. This translates into the parallel subsample, $\sim$ 0.35 dex below the one-to-one relationship, clearly noticeable in Fig. 8. Only $\sim$ 7.5% of the sample suffers this effect, and given the redshift uncertainties, we cannot exclude these galaxies from further analysis. In conclusion, we consider that the $J0515$ EW and line fluxes correspond to the sum of the [OIII]4959 and [OIII]5007 emission lines, while the r line flux and EW correspond to $H\alpha$. For the 18 galaxies with photometric redshift z_{phot} < 0.017, the $H\alpha$ line falls within the $J0660$ filter wavelength range, and the precision of the $H\alpha$ measurement is much higher, with an offset of $\sim$ 0.002 dex and a scatter of $\sim$ 0.02 dex.

3.2. Selection of the EELG sample

Our original selection of candidates was performed using a very rough estimation of the stellar continuum below the [OIII] line (simply the r band flux), and this results in an extended [OIII] EW distribution (see Fig. 7). We decided to perform an additional cut to obtain a clearly defined EELG sample, with a strict limit in EW. We considered (only for galaxies with SNR> 3 in [OIII] EW) how the logarithm of the number of galaxies in each EW bin varies with EW, given the expected decreasing exponential relation (see Fig. 10). This exponential relation is clear (with R^2=0.998), for example, when selecting galaxies with EW([OIII])>10 Å and SNR>3 in the SDSS spectroscopic
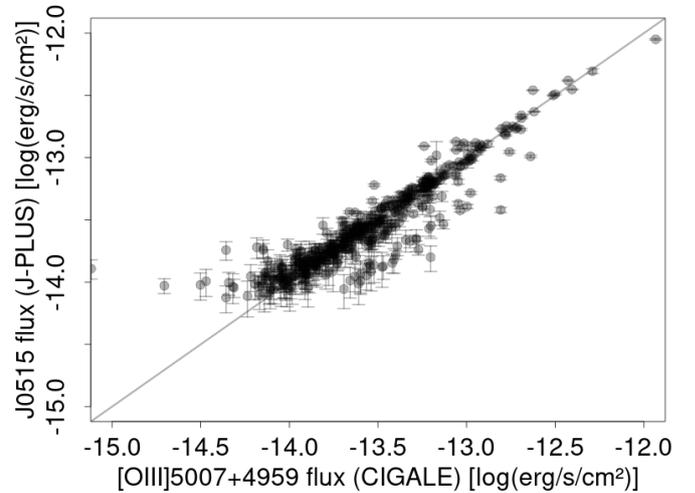
Fig. 8. Comparison between the emission line flux measured in the $J0515$ filter and the [OIII]5007+4959 flux estimated by the `CIGALE` SED fit for the EELG sample. The black line represents the one-to-one relation.
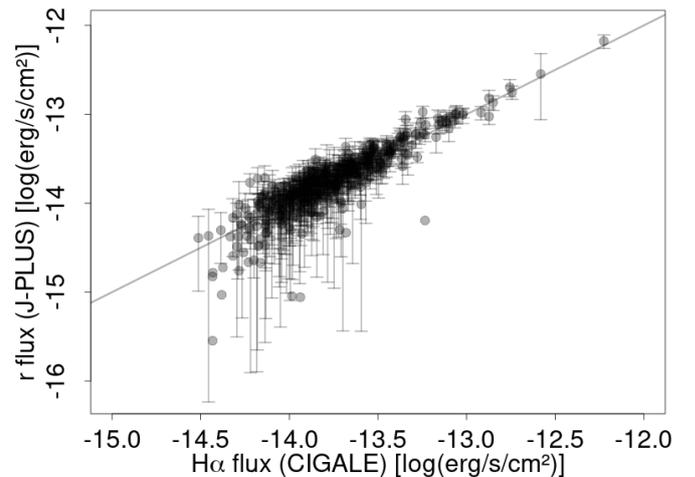
Fig. 9. Comparison between the emission line flux measured in the r filter and the $H\alpha$ flux estimated by the `CIGALE` SED fit for the EELG sample. The black line represents the one-to-one relation.

database (the GalSpecLine table from DR8, Aihara et al. 2011). In our case, for a given lower limit of EW, the number of galaxies will not increase as fast as an exponential: that would be our completeness limit. In order to compute it accurately, we perform linear fits with EW bins of 0.1 dex in log(Å), limiting the analysis in the high EW range to the bins showing more than 10 galaxies, which translates into log(EW) < 3.05 log(Å). For the low EW limit, we perform three separate linear fits, covering the bins from 2.45, 2.55, and 2.65 log(EW), respectively. For log(EW) values lower than 2.5 (EW $\sim$ 300 Å), the measured value drops below that of a linear fit to the bins above it (Fig.

10), with the fits limited to 2.55 and 2.65 showing comparable parameters. Therefore we place our threshold at EW = 300 Å in [OIII] to keep a complete sample, obtaining a total of 466 EELGs. In the higher EW limit, completeness is limited by the $r<20$ threshold defined in Sect. 2, given that fainter (and lower mass) galaxies tend to show higher EWs (see Sect. 4.2.1). This magnitude cut is nevertheless necessary in order to provide with a clearly defined sample, and not to include significant contamination of high-z AGNs (as seen in Fig. 3).

The value we selected is higher than some limits found in the literature for starburst galaxies or EELGs (100 Å in Amorín et al. 2015, 80 Å in Hinojosa-Goñi et al. 2016), but lower than most of the galaxies in other surveys (Cardamone et al. 2009; Yang et al. 2017), especially those at high redshift (van der Wel et al. 2011; Atek et al. 2011). We consider that the intermediate value of 300 Å is appropriate, given the width of the filter used to select the sample. It provides a large sample, showing undoubtedly strong emitters, allowing both a statistical analysis and low contamination. A higher threshold would have removed interesting objects and restricted the sample, while a lower one could provide a high rate of contaminants unless we removed the lower brightness objects, which would be large fraction or the sample. Comparing with the SDSS spectroscopic database (using again the GalSpecLine table from DR8), we see that with a 300 Å threshold we keep only the top 2% of the EW([OIII]) distribution (considering only galaxies with EW([OIII])>10Å), which further confirms that the emission properties of our sample are extreme.

We consider as well the effect of the initial selection (Eq. 1) in the Hα flux distribution of the final sample. Since the Hα lines falls within the wavelength range of the r filter and Eq. 1 selects objects with high $J0515$ to r ratio, we could be biasing the sample, removing objects with relatively high [OIII] EW yet low [OIII]5007/Hα ratio. To test if this effect was present, we artificially increased the `CIGALE`-estimated Hα fluxes in the selected galaxies (those with EW([OIII])>300 Å) and check if they still fulfilled Eq. 1. A 25% increase in the Hα flux was used, since that produced a significant offset in the [OIII]5007/Hα ratio as a function of EW([OIII]), unobserved in the spectroscopic data. Using that 25% increase, only 16 objects (3.2%) were rejected, therefore we are confident that no significant bias is added to our sample against galaxies with low [OIII]5007/Hα ratio due to Eq. 1.

This sample selection results in 466 EELGs. In order to determine how novel the J-PLUS database and the present work are, we investigate the amount of objects in our sample that have not been previously identified as EELGs. To do so, we first query the NASA Extragalactic Database (NED) to select catalogued objects near the positions of the galaxies in our sample. Only 53 have a reported spectroscopic redshift, and we consider that those galaxies have been already classified as EELGs. In addition, 14 galaxies without spectroscopic redshift in NED have been referenced in at least one publication, with 2 of them described as EELGs or a similar category. Finally, we check that all objects with spectroscopic confirmation in this EELG sample are indeed low redshift galaxies, which translates into a theoretical purity of 100%. In conclusion, the present work identifies as EELGs 411 galaxies (88% of the sample) that were previously unknown to belong to this class. This highlights how the J-PLUS survey is well suited to find new extreme emitters, reaching fainter magnitudes than spectroscopic surveys covering wide areas of the sky.
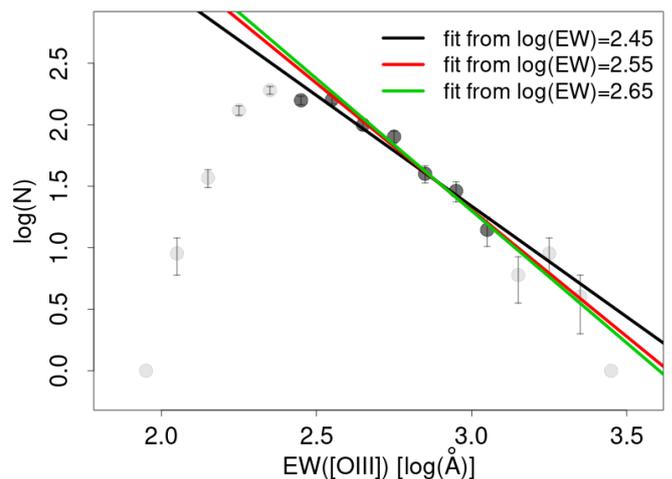
Fig. 10. Diagram used to asses the completeness of the EELG sample. We plot the number of galaxies per log(EW[OIII]) bin as a function of log(EW[OIII]). We overplot the linear fits to the data considering three different limits in the lower end 2.45, 2.55, and 2.65 log(Å).
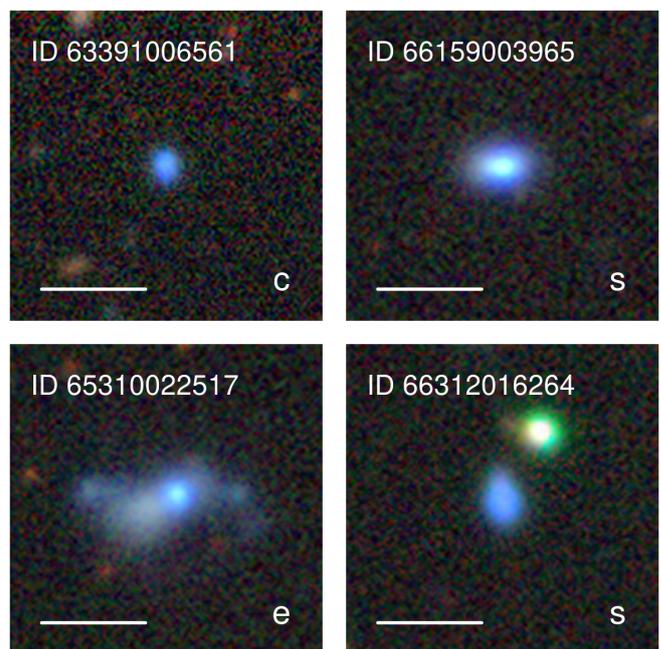
Fig. 11. RGB composites (made using images taken with z, r, and g filters) of four galaxies in the EELG sample. In the bottom right of each panel a letter indicates the type of morphology: compact (c), semi-compact (s), or extended (e). All images have been obtained using the Legacy Survey viewer, with data from the BASS and MzLS surveys (Zou et al. 2019; Dey et al. 2019).

3.3. Morphology of the EELG sample

We have performed a simple visual classification of the morphology of the extreme [OIII] emitters. The EELGs in the sample have been classified into three categories: compact (c), semi-compact (s), and extended (e). Compact those galaxies show a circular shape, with no further structure discernible. We classify them in the semi-compact category when their shape is dominated by a circular and bright clump, but some fainter structure is revealed: a tail, a halo, etc. We consider extended galaxies those showing a complex morphology, those with no bright clump, or
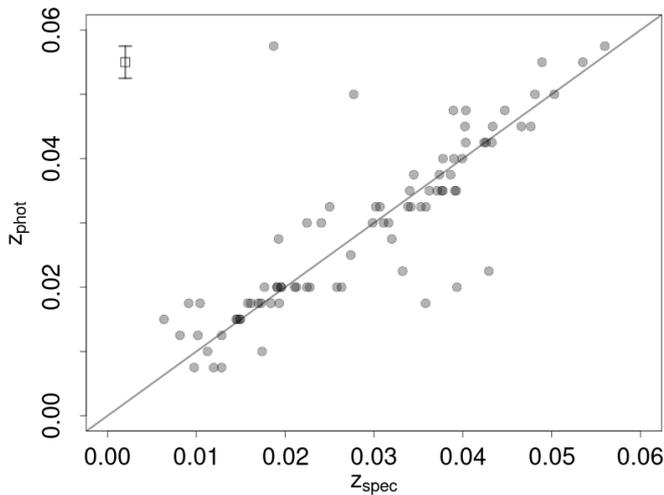
Fig. 12. `CIGALE` photometric redshift as a function of SDSS-derived spectroscopic redshift, for the galaxies in our sample of candidates that have available spectra (excluding the three misclassified objects). The line represent the one to one relationship. On the top left of the figure we show as error bars the minimum step in photometric redshift considered with `CIGALE` (0.0025).
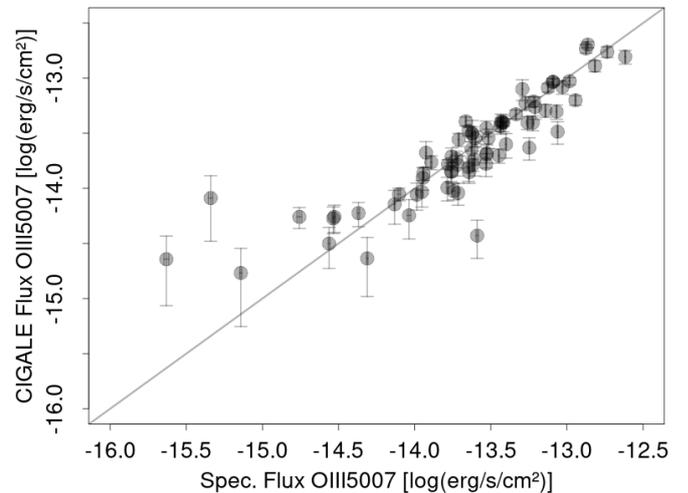

Fig. 13. [OIII]5007 flux estimated using `CIGALE` SED fitting on J-PLUS data as a function of [OIII] flux measured in SDSS spectra. The black line represents the one-to-one relation.

those where the clump or clumps does not dominate the light profile in the galaxy. As an example, we show the RGB post stamps of four galaxies in Fig. 11.

For this analysis we took the images from the Legacy Survey database[2] that had been obtained in the framework of the BASS and MzLS surveys (Zou et al. 2019; Dey et al. 2019). They provide images in the g, r, and z bands, with median depths of 23.65, 23.08, and 22.60 mag respectively (considering 5σ detections of point sources), and typical FWHM values of their PSFs of 1.61", 1.47", and 1.01" (Dey et al. 2019). These images reach low surface brightness detection limits in r of 27.9 mag arcsec^2 for 3σ detection of a 100 arcsec^2 feature, fainter than the SDSS and Panstarrs PS1 surveys, and comparable to SDSS Stripe-82 (Hood et al. 2018). Therefore, even if fainter features than those visually noticeable in the RGB composite images may exist (and change the morphological classification), we consider the present analysis provides with results accurate enough for this work.

The compact class is the most common with 43% of the EELG sample, followed by the semi-compact one (38%), the rest being classified as extended. This is consistent with previous studies in this topic (e.g. Izotov et al. 2011; Yang et al. 2017), focused in the compact or semi-compact galaxies. Nevertheless, we show that a significant amount (19%) of the EELGs in our sample are more properly classified as extended systems, suggesting possible extensions of previous EELG searches.

3.4. Comparison with spectra

We assessed the accuracy of the properties derived using J-PLUS photometry and the `CIGALE` SED analysis by disucssing the results obtained for the 82 sources in the candidate sample with SDSS spectra. For this analysis, we removed the misclassified objects (one star and two high-redshift galaxy). First, we measured the accuracy of the photometric redshift determined with `CIGALE`, which are compared in Fig. 12 with the spectroscopic redshifts, showing very good agreement. We computed a param-

---

[2] https://www.legacysurvey.org/viewer

eter to quantify the quality of the agreement, the σNMAD, defined as

$$\sigma_{\text{NMAD}} = 1.48 \times \text{median}\left|\Delta z - \text{median}(\Delta z)\right|, \quad (5)$$

where $\Delta z = (z_{\text{best}} - z_{\text{spec}})/(1+z_{\text{spec}})$. We obtained a value of σNMAD ~ 0.003, similar to the mini-JPAS results, with 56 narrow band filters (Hernán-Caballero et al. 2021). This high accuracy is likely due in part to the limited redshift range covered, to the extreme intensity of the emission lines, but most importantly, to the fact that those lines ([OII]3727, [OIII], and Hα) lie inside or near the wavelength range of narrow or medium-band filters, which produces very clear photometric features. In fact, for more than half of this subsample (56%), the difference between photometric and spectroscopic redshift is smaller than the minimum step used in the photometric redshift determination (0.0025). This high accuracy diminishes the possible uncertainties in our determinations of absolute parameters, such as luminosities and stellar masses. We used the best fitting `CIGALE` models instead of the Bayesian estimates, in part because the agreement with spectroscopic redshift values drops when using the latter. In consequence, the values discussed for the SED fitting (ages, masses, etc.) correspond to the best-fit models.

We also compare the fluxes and EWs estimated in the J-PLUS photometry with those measured using the SDSS spectra. We use our own code to measure the emission line fluxes and EWs in the spectra, a revised version of the procedure used in Lumbreras-Calle et al. (2019a). Briefly, we compute the continuum underneath each emission line by masking the line and performing a linear fit to the remaining spectrum in a 200 Å aperture around the line. We estimate the values and uncertainties of the linear fits by performing Bootstrap simulations. We compute the flux of each emission line simply adding the measured flux over the spectral window considered and subtracting the continuum value (and then dividing by the continuum value to get the EW).

It is important to remind the reader that the photometry used in this work has been the PSFCOR J-PLUS photometry, rescaled to the AUTO aperture value in r. In order to perform an accurate comparison, we need to use the 3ARCSEC aperture in the J-PLUS catalog, which should match more closely the SDSS spectroscopic measurement (which is performed with a 3 arcsecond wide fiber in most of the sample). Nevertheless,
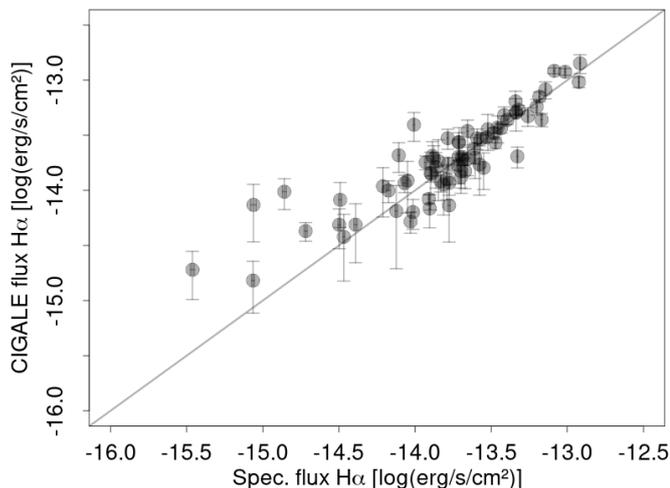
Fig. 14. Hα flux estimated using `CIGALE` SED fitting on J-PLUS data as a function of Hα flux measured in SDSS spectra.The black line represents the one-to-one relation.
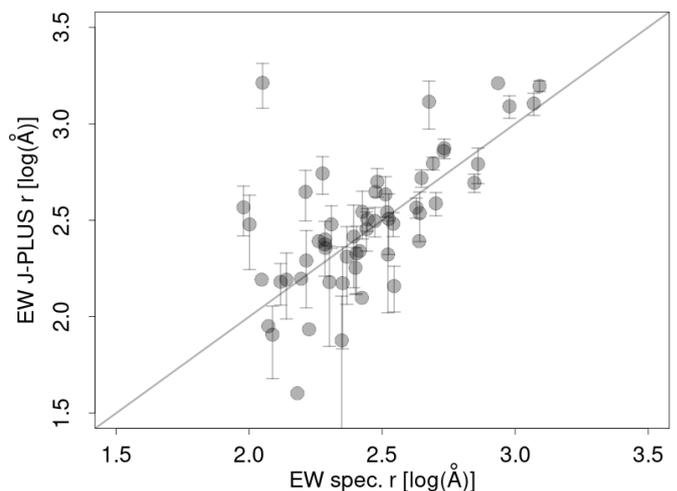

Fig. 15. J0515 EW directly measured on J-PLUS data as a function of J0515 EW estimated convolving SDSS spectra with the filter transmission.

Fig. 16. r EW directly measured on J-PLUS data as a function of r EW estimated convolving SDSS spectra with the filter transmission.

some sources of discrepancy still exist: different seeing conditions, small differences between the sky position of the objects, or a possible offset between absolute calibrations. We compared the 3ARCSEC r fluxes in J-PLUS with the synthetic r fluxes computed from the spectra, and found very good agreement (a $1\sigma=0.12$ dex scatter) with a small but noticeable offset (0.08 dex). Therefore, we re-scaled the 3ARCSEC photometry to the SDSS flux using this r band offset.

For a precise comparison, line by line, we need to use the `CIGALE` model output, which provides us with fluxes for several emission lines ([NII]6584, Hα, Hβ, [OIII]5007, [OIII]4959, and [OII]3727). In Figures 13 and 14 we show the comparison between spectra and SED models for the [OIII]5007 and Hα lines respectively. For [OIII]5007 the agreement between the measurements is very good, with almost no bias (a median difference of -0.001 dex) and very low scatter, a 1σ value of ~ 0.18. We can also compare our result with others in the literature. In the Census of the Local Univers (CLU) preliminary fields, Cook et al. (2019) show the comparison between photometric and spectroscopic Hα fluxes in their figure 10. Considering only their 5σ detections, we estimate a 1σ scatter of ~ 0.25. This value is higher than ours, which is striking considering that their filters are notably narrower than ours (from 76 to 92 Å compared to 200 Å). For a more similar comparison, using J-PLUS data, we look at the results in Logroño-García et al. (2019). Considering Hα fluxes measured in the J0660 filter, these authors reach more accurate values than ours, with a 1σ scatter of ~ 0.11. This result is expected, given that the J0660 is narrower than J0515 (~ 140 vs. ~ 200 Å), and that they took extreme care in matching the apertures of the photometric and spectroscopic measurements, which included integral field unit data.

The comparison for the Hα line is also very successful, with a similar offset ($\sim$ dex) and scatter (~ 0.18 dec.) even if the line flux is estimated using the broadband r filter. This value is likely low thanks to the data in the J0660 filter, that very accurately traces the continuum near the Hα line for galaxies at z>0.017. For the few galaxies in our sample with available spectra and z<0.017, `CIGALE` uses data from the J0660 filter to estimate the Hα flux, obtaining an even better agreement (with an offset of $\sim$-0.01 and a scatter of ~ 0.12). Other emission lines show a worse agreement between spectra and `CIGALE` models, some simply
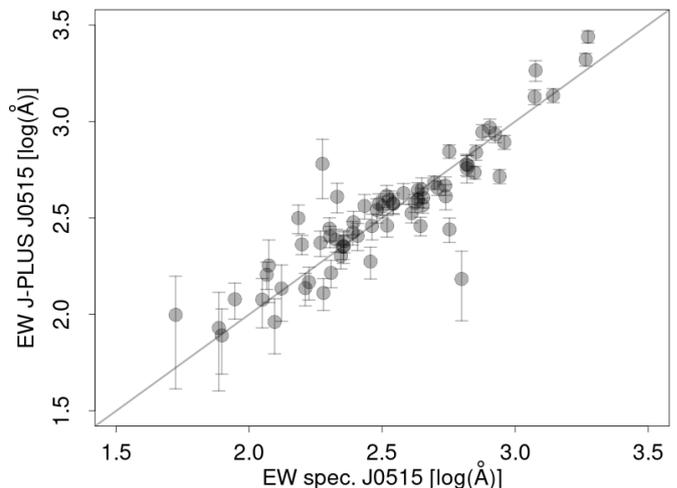
because they are less intense (for example, the [NII]6584 line, which in addition is very close to the much more intense Hα line). Some others are not only faint, but also located in regions were the spectroscopic and photometric analysis is more complex. For example, the [OII]3727 line lies outside the wavelength range of most of the spectra (and on the edge of some), and it can fall in the gap between two blue narrowband J-PLUS filters (J0378 and J0395) which are also harder to calibrate than the redder ones. Nevertheless, it is worth noting that making direct measurements on the J-PLUS data, the [OIII]/[OII] ratio of the sample of EELGs reaches extremely high values compared to typical star-forming galaxies. This indicates very hard ionizing radiation, and is similar to very high redshift galaxies. We will explore this result in depth in an upcoming paper analyzing the spectra of some of the galaxies in the current sample.

The previous comparison between line fluxes, even if very successful for [OIII]5007 and Hα, relies heavily on the SED models. For a less model-dependant version, we look into the comparison between the EW measured directly in the J-PLUS photometry and the synthetic EW values for each filter, obtained convolving the emission lines in the SDSS spectra with the filter

transmission curves. We should note that there is still some modeling implied, since we used the CIGALE fits in order to estimate the continuum in each filter. We show the results for $J0515$ and r in Figures 15 and 16, respectively. In this case, the agreement is even better, with similarly small offsets (0.02 for $J0515$ and 0.04 for r) but lower scatter values (0.12 and 0.17, respectively). This result strengthens our confidence in the accuracy of J-PLUS photometry and our SED and EW analysis.

4. Discussion

In this section we compare our sample with different works in the literature, regarding their broadband colors, number density values, and the ratio between [OIII] and the Hα and Hβ emission lines. In addition, we discuss the physical properties derived for our EELG sample, considering EW, star formation rate (SFR), mass, and other properties in the context of the literature at different redshifts.

Given the good agreement between the photometric and spectroscopic redshifts, as well as between the photometric and spectroscopic line fluxes and EWs, we are able to discuss the results of this work without performing follow-up spectroscopic observations on the whole sample of galaxies. That step has been however necessary in previous works that used only broadband selection (Yang et al. 2017; Senchyna & Stark 2019; Kojima et al. 2020) or narrow band selection like Hα dots (Kellar et al. 2012; Salzer et al. 2020) and the CLU survey (Cook et al. 2019). Even if the physical information we can obtain is limited compared to what can be derived from spectroscopic observations, it is still significant, considering the data spans 2 176 sq. deg. down to $r = 20$. In addition, this photometric analysis allows for a more efficient spectroscopic follow-up, targeting galaxies with specific physical properties. This characteristics will improve significantly in the upcoming J-PAS survey, with deeper observations and higher spectral resolution.

4.1. Testing the sample selection

4.1.1. Comparison with broadband color-color selection

Several works over the past years have used broadband colors from large photometric surveys (most notably SDSS) in order to select galaxies with strong emission lines and/or extremely metal poor gas. (Cardamone et al. 2009; Yang et al. 2017; Senchyna & Stark 2019; Kojima et al. 2020). For the selection of extreme [OIII] emitters with broadband data, it is necessary to define regions of extreme $g - r$ color (or $r - i$ at higher redshift) in order to avoid selecting typical galaxies. Line ratios and redshift affect these colors, preventing some galaxies with very high [OIII] emission from showing extreme colors. This is presented in Figure 17, where many extreme [OIII] emitters in our sample are indistinguishable from the main SDSS galaxy population using only the g, r and i bands. Our sample of EELGs covers as expected the region of the Yang et al. (2017) blueberry galaxies (since they present extreme [OIII] emission at our redshift range), but it also covers the same color space as extremely metal-poor galaxies both from observations and models (Kojima et al. 2020). This effect is mainly due to the different Hα/[OIII] line ratios, given how dominated the broadband fluxes are by the line emission in this kind of galaxies. If this ratio is high, the g-r color will not be extreme, while if the ratio is low, the g-r color will be very negative. This ratio, even if high, does not prevent us from identifying these extreme emitters, since we look at the
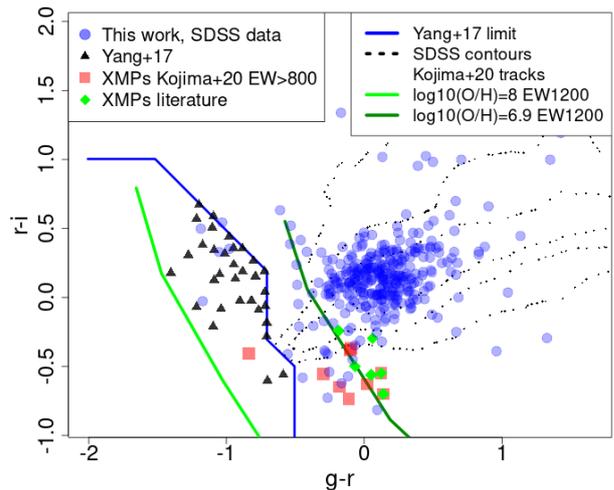
Fig. 17. Broadband color-color diagram, created using SDSS data. We plot our EELG sample in blue dots, alongside the blueberry galaxies in Yang et al. (2017) and extremely metal poor galaxies (XMP) from Kojima et al. (2020) and the literature. The blue line represent the Yang et al. (2017) sample selection limits, while the green lines follow the evolutionary tracks of models in Kojima et al. (2020). The small black dots represent the density contours of typical SDSS galaxies (Kojima et al. 2020).

flux in the $J0515$ mediumband filter, proving again the added value of the mediumband filters in the J-PLUS survey.

4.1.2. Number density of EELGs

The detection of EELGs presented in this work follows clear and reproducible procedures, without pre-selection of targets (other than a magnitude limit). We can therefore use our sample to estimate the number density of this class of galaxies in the redshift range probed, and compare it with other works in the literature. In order to roughly estimate uncertainties, we consider two sources of errors: Poisson noise (associated with the discreet nature of galaxies) and cosmic variance. We approximate the Poisson noise by $\sqrt{N}$, where N is the number of galaxies, and estimate the error associated with cosmic variance using the code presented in Driver & Robotham (2010). We add in quadrature both errors to obtain an uncertainty estimation. The results of this analysis are shown in Table 5.

First, the number density of EELGs in our J-PLUS sample is computed. We limit the analysis of our sample to the redshift range where both [OIII] lines lie within the $J0515$ filter, $0.016 < z < 0.048$, in order to ensure completeness. We find 394 EELGs in this range, which translates into a number density of $(2.45 \pm 0.27) \cdot 10^{-4}$ Mpc^{-3}. This is about 1 EELG every 4000 Mpc^3.

In order to compare the number density of EELGs we measure with other works, we impose limitations in magnitude (or mass), [OIII] EW, and redshift range. This is done both for the literature samples and ours, in order to match them appropriately. We start the comparisons with the GAMA survey, which provides with precise spectroscopic data, while being still relatively deep and complete. They obtained fiber spectra of essentially all sources brighter than a certain magnitude threshold in the targeted fields. In order to perform a density comparison we limit ourselves to three of the four fields with available data (ex-

Table 5. Number density of EELGs in this work, and comparison with several literature samples.

Sample	Selection method	N	Density [Mpc^{-3}]	Redshift range	EW limit [Å][a]	Mag. limit[b]	Mass limit log $(M_\star/M_\odot)$
J-PLUS	Mediumband	394	$(2.45 \pm 0.27) \cdot 10^{-4}$	0.016 - 0.048	300	20.0	-
GAMA	Spectra	13	$(8.77 \pm 3.40) \cdot 10^{-5}$	0.016 - 0.048	300	19.0	-
J-PLUS comparison		113	$(6.96 \pm 0.96) \cdot 10^{-5}$				
SDSS	Spectra	90	$(1.37 \pm 0.23) \cdot 10^{-5}$	0.016 - 0.048	300	17.7	-
J-PLUS comparison		16	$(0.99 \pm 0.27) \cdot 10^{-5}$				
Yang et al. (2017)	Broadband	7	$(0.56 \pm 0.22) \cdot 10^{-6}$	0.016 - 0.048	1600	20.0	-
J-PLUS comparison		14 [c]	$(8.38 \pm 2.39) \cdot 10^{-6}$				
Cardamone et al. (2009)	Broadband	27	$(0.020 \pm 0.004) \cdot 10^{-6}$	0.18 - 0.32	400	-	9.52
J-PLUS comparison		2 [c]	$(1.240 \pm 0.890) \cdot 10^{-6}$	0.016 - 0.048			
Hα dots	Narrowband	6	$(1.60 \pm 1.50) \cdot 10^{-3}$	0.001 - 0.024	300	20.0	-
J-PLUS comparison		104 [c]	$(0.70 \pm 0.10) \cdot 10^{-3}$	0.016 - 0.024			

[a] This limit refers to the [OIII]4959+5007 EW.

[b] The magnitude limits are placed considering *r* magnitude in all samples except for the Hα dots, where *R* magnitude was used instead. In the GAMA and SDSS samples and comparisons, petrosian magnitude was used. For the Yang et al. (2017) sample (where they used `cmodelmag` magnitudes) and for the Hα dots sample (circular apertures), we used AUTO apertures in the J-PLUS comparisons.

[c] In the Yang et al. (2017), Cardamone et al. (2009), and Hα dots comparisons, we restrict the J-PLUS sample to compact or semi-compact objects, to reproduce their selection criteria.

cluding G02 due to the use of a different input catalog). To further homogenize the selection, we limit the comparison to galaxies brighter than petrosian magnitude 19 in the *r* band, where both GAMA and our sample are complete. Using the data in the GaussFitSimple table within the SpecLineSFR data management unit, we select galaxies with EW([OIII]5007)>225 Å in GAMA, which would correspond to ~ 300 Å in our J-PLUS analysis ([OIII]4959+5007). This yields 13 GAMA galaxies over the 0.016 < z < 0.048 redshift range. A similar cut in petrosian magnitude and EW in our sample yields 113 objects. When the relative areas and the volume of the Universe between those redshifts are taken into account, the density of GAMA sources is $(8.77 \pm 3.40) \cdot 10^{-5}$ Mpc^{-3} and $(6.96 \pm 0.96) \cdot 10^{-5}$ Mpc^{-3} for J-PLUS.

We also compare to the SDSS database, limiting our analysis in this Section to the main legacy survey and the MPA-JHU catalog (Kauffmann et al. 2003; Brinchmann et al. 2004; Tremonti et al. 2004; Salim et al. 2007), presented in the GalSpecLine table from DR8 (Aihara et al. 2011). The clear selection of this sample (essentially petrosian magnitude brighter than 17.7 in *r*, Strauss et al. 2002) allows for an accurate comparison. We choose galaxies in the same redshift and EW ranges as with the GAMA survey. Their density in this case is $(1.37 \pm 0.23) \cdot 10^{-5}$ Mpc^{-3}, compatible with the $(0.99 \pm 0.27) \cdot 10^{-5}$ Mpc^{-3} density in our J-PLUS data for the same cuts.

The results from the previous two comparisons show that the density of EELGs computed in this work using J-PLUS is compatible with the ones derived from magnitude-limited spectroscopic surveys, which can be considered as "ground truth". This result shows the strength of our contribution, which is able to select essentially all EELGs at this redshift range down to a certain magnitude threshold. Moreover, our work has some advantages over spectroscopic analyses: we cover much wider areas than pencil beam surveys such as GAMA, and we are complete down to deeper magnitudes than a wide-field spectroscopic survey like SDSS.

We can also compare our density values with broadband-selected samples, like the Blueberry Galaxies (Yang et al. 2017) and Green Peas (Cardamone et al. 2009). In this case, we restrict our sample to only the compact and semi-compact categories, to more closely match their sample selection. In both works the authors select only galaxies with very high EW, but without a clear threshold value. Therefore we limit our comparison to the galaxies with the highest EW values of these samples, in order to approach completeness. For Yang et al. (2017), we have placed the EW limit roughly where the number of galaxies per EW bin decreases with increasing EW of the [OIII]5007 emission line (around 1200 Å, corresponding to 1600 Å in [OIII]5007+4959). In addition, we cut the comparison sample at *r* < 20 to simulate our limit in brightness, given the similar redshift range covered in both works. Estimating these limits for Cardamone et al. (2009) is more complex, given their EW distribution and higher redshift range (which results in larger masses). As an exercise, we place the limits in EW>300 Å and log $(M_\star/M_\odot) > 9.5$, obtaining only two galaxies in our sample, against 27 in their case. These thresholds result in a density of $(0.56 \pm 0.22) \cdot 10^{-6}$ Mpc^{-3} for Yang et al. (2017), and more than ten times higher in J-PLUS, $(8.38 \pm 2.39) \cdot 10^{-6}$ Mpc^{-3}. The comparison between the Cardamone et al. (2009) sample and our work is even more striking, with values of $(0.020 \pm 0.004) \cdot 10^{-6}$ Mpc^{-3} vs. $(1.24 \pm 0.890) \cdot 10^{-6}$ Mpc^{-3}, but more uncertain.

Other surveys use narrowband imaging to identify emission line galaxies. In principle, they should be more sensitive to low EW emission lines than J-PLUS, since their filters are narrower than our *J*0515 filter, and therefore the contrast in brightness between narrow and broadband should be higher. Nevertheless, given that we limit our analysis only to the most extreme events of star formation, the strong contrast is enough to avoid missing a significant amount of EELGs. The sky area and redshift range that narrowband surveys cover is smaller than ours, which limits their ability to identify extreme, rare objects. To our knowledge, there has not been any narrowband survey targeting the [OIII] emission line in the local Universe, so we cannot perform

any direct comparison. A similar example are the Hα dots identified in the ALFALFA Hα survey (Kellar et al. 2012; Salzer et al. 2020). While they select galaxies based on Hα emission at our redshift range, they provide spectroscopic follow-up and thus [OIII] EW. We compute the density of Hα dots showing [OIII] EW higher than 300 Å and brighter than $r = 20$ located at $z < 0.024$ (beyond that the transmission of their reddest Hα filter drops). The value we obtain for their number density $(1.60 \pm 1.50 \cdot 10^{-3} \text{ Mpc}^{-3})$ is higher than the corresponding value for our sample $(0.70 \pm 0.10 \cdot 10^{-3} \text{ Mpc}^{-3})$, considering again only compact or semi-compact galaxies in our J-PLUS comparison. Nevertheless, given the high uncertainty in the density value derived for the Hα dots survey ($\sim 80\%$, driven mostly by cosmic variance), the density values are compatible within one standard deviation.

4.1.3. [OIII]5007/Hα and [OIII]5007/Hβ ratios

The present work intends to provide a sample of EELGs selected purely on magnitude and EW([OIII]), with no direct bias on line ratios, since it is only based on the detection of flux excess in a mediumband filter. In contrast with this, other works based on broadband photometry detection (such as Yang et al. 2017 and Cardamone et al. 2009) are biased against low [OIII]5007/Hα systems (and therefore also against low [OIII]5007/Hβ ones). This is because they demand that the flux of the filter where the [OIII] doublet falls (g in Yang et al. 2017, r in Cardamone et al. 2009) must be significantly stronger than the filter where the Hα line flux contribution lies (r in Yang et al. 2017, i or z in Cardamone et al. 2009). According to stellar population synthesis and nebular photoionization models (such as Inoue 2011), the [OIII]5007/Hα ratio reaches a maximum value for gas metallicities around $12 + \log(O/H) \sim 8.0$, which is further confirmed by the typical and metallicities of Yang et al. (2017) and Cardamone et al. (2009). Therefore, selecting galaxies with high [OIII]5007/Hα ratios biases the samples against very low metallicities (Senchyna & Stark 2019; Kojima et al. 2020).

Since we only select extreme [OIII] emitters, we are certainly missing systems that have [OIII] EW below our threshold, but above it we should recover all galaxies, regardless of their [OIII]5007/Hα ratio. In Fig. 18 we compare the distribution of the [OIII]5007/Hα and [OIII]5007/Hβ ratios for several samples of ELGs. We consider more reliable the [OIII]5007/Hβ ones, since the emission lines are much closer in wavelength, minimizing the effect of dust extinction correction and/or flux calibration issues. We see that our work shows lower median values for these ratios than broadband selections such as Yang et al. (2017) or Cardamone et al. (2009), and we reach much lower values. Our selection shows slightly higher values than a pure spectroscopic selection (Amorín et al. 2015), specially in the [OIII]5007/Hα ratio. A stronger difference is present when comparing with Hα selected samples (the Hα dots, Kellar et al. 2012; Salzer et al. 2020) and extremely metal poor galaxies (Hsyu et al. 2018; Guseva et al. 2017), specially if we include systems with low Hβ EW (a proxy for low sSFR). Nevertheless, we can see that the range of values that our survey covers overlaps with most of the range of the extremely metal poor and strongly star forming samples, in contrast with Yang et al. (2017) and Cardamone et al. (2009). In addition, we plot as well the typical values for other types of systems, some with strong ionizing spectra: HeII emitters (Pérez-Montero et al. 2020) and Lyman α emitters (Matthee et al. 2021; Izotov et al. 2021). In these cases, their ratios are more similar to ours.
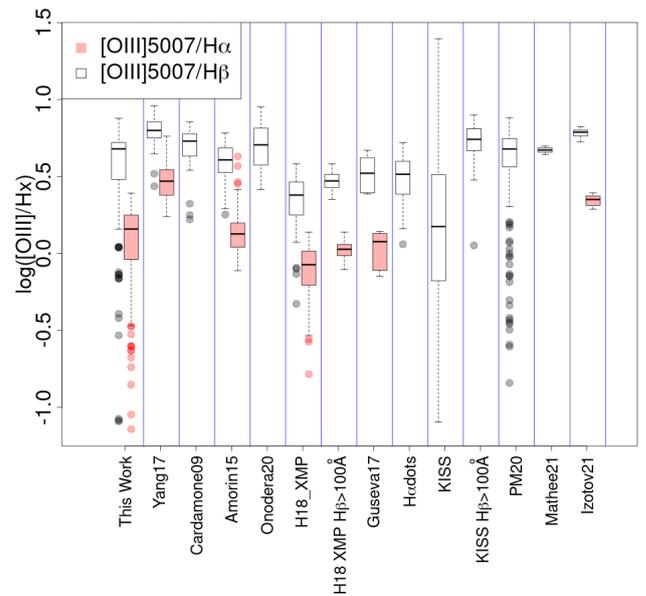
Fig. 18. Distribution of the ratios [OIII]5007/Hα (red-filled rectangles and dots) and [OIII]5007/Hβ (white-filled rectangles and grey dots). The thick black line represent the median value, while the box ranges from the first quartile to the third one. The error bars represent the maximum and minimum value without outliers, which are plotted as dots (and defined as the values that lie beyond 1.5 times the inter quartilic range). We plot in this figure results from the present work and several samples from the literature (Yang et al. 2017; Cardamone et al. 2009; Amorín et al. 2015; Onodera et al. 2020; Hsyu et al. 2018; Guseva et al. 2017; Pérez-Montero et al. 2020; Matthee et al. 2021; Izotov et al. 2021), as well as the KISS survey (Wegner et al. 2003; Gronwall et al. 2004; Jangren et al. 2005; Salzer et al. 2005) and the Hα dots class (Kellar et al. 2012; Salzer et al. 2020). We have selected the subsample with extremely low metallicity in Hsyu et al. (2018) (H18_XMP), as well as those galaxies with EW(Hβ)>100 Å in both Hsyu et al. (2018) XMPs (H18_XMP Hβ>100 Å) and the KISS survey (KISS Hβ>100 Å). Different samples are separated by vertical blue lines.

Finally, it is also important to keep in mind that several properties in the samples affect the [OIII]5007/Hα ratios, apart from selection biases and metallicity. The SFR is also positively correlated with the ratio (as seen for example when restricting the Hsyu et al. (2018) sample only to objects with high Hβ EW). Stellar mass also plays a role (with higher mass galaxies showing lower ratios), but it is less relevant than the other factors: the Cardamone et al. (2009) sample has a typical mass ~ 100 times higher than those of Hsyu et al. (2018) or Guseva et al. (2017), yet their ratios are higher. While these factors play a role in the results shown in this section, our selection method is clearly open to the selection of low [OIII]/Hα and Hβ galaxies, in contrast with others.

4.2. Physical properties of the EELG sample

4.2.1. [OIII] EW and stellar mass

A clear negative correlation between emission line EW and stellar mass has been found in the literature (e.g. Fumagalli et al. 2012; Sobral et al. 2014; Khostovan et al. 2016; Reddy et al. 2018; Lumbreras-Calle et al. 2022 *in prep.*). This indicates that galaxies with lower masses tend to have stronger recent star-formation events relative to their mass (higher sSFR). The relationship between the two variables has been fitted with a linear

model (in logarithmic units), and depends on the emission line studied (both intercept and slope) and the redshift of the galaxies (mostly the intercept), although discrepancies are still present (i.e. between Khostovan et al. 2016 and Reddy et al. 2018, but see Khostovan et al. 2021). It is beyond the scope of the present paper to shed light into the values of that linear relation, given that our selection process focuses only in a region of the EW - mass diagram, making a linear fit to our data meaningless in this context. Nevertheless, we can compare the values obtained to those available in the literature for the [OIII] line (see Fig. 19).

The most clear result is that, even if our galaxies reside in the very low redshift Universe, they populate almost exclusively the regions of the diagram defined by the linear fits to galaxies at intermediate redshift (z > 0.84), with EW values similar to those typically seen at z ∼ 1.4 (Khostovan et al. 2016). We have as well in our sample some galaxies with EW similar to those at medium-high redshift: z ∼ 2.2 − 3.4 for Khostovan et al. (2016), z ∼ 3.4 for Reddy et al. (2018), and even z ∼ 8 (De Barros et al. 2019). The very high EW values, low [OII]/[OIII] ratios, and typically compact morphologies are similar to those found in very high redshift galaxies (e.g. Onodera et al. 2020). According to detailed spectroscopic studies in the low redshift universe (Izotov et al. 2021), this type of galaxies share physical properties with those forming stars at very high redshift, likely leaking Lyman continuum radiation, and playing an important role in the reionization of the Universe. In consequence, we consider the present sample may be useful in improving the understanding of the very early Universe.

We split the sample in the three morphological categories described in Sect. 3.3, and show them in different colors in Fig. 19. Compact and semi-compact galaxies have slightly lower masses, with typical values of $8.06^{+0.65}_{-0.54}$ and $8.15^{+0.58}_{-0.59}$ respectively, compared to $8.38^{+0.56}_{-0.67}$ for extended galaxies. Driven mostly by this difference and the EW - mass anticorrelation, compact galaxies show lower typical EW values. Nevertheless, for the lower mass range, compact galaxies show clearly higher EW values than extended ones even at similar masses. This suggests that there is an additional physical effect that favors higher EW values in compact galaxies over extended ones.

The linear relation shown in Khostovan et al. (2015) and Reddy et al. (2018) between EW([OIII]) and stellar mass appears to have a limit at EW([OIII]) ∼ 3 000 Å (Reddy et al. 2018). This limit, reached for galaxies dominated by the star-forming burst, is likely to be physical in nature rather than purely observational, given that for galaxies with a given mass, there is no bias against selecting higher EW systems. Understanding the reason behind this limit could provide insights into the star formation cycle in the most extreme environment, in galaxies at the highest redshifts and with the lowest masses and metallicities. Our sample provides several previously unknown nearby candidates around that limit, which can be targeted with follow-up spectroscopic observations to extract relevant conclusions on this topic.

4.2.2. SFR and Main Sequence

One of the most important properties of star-forming galaxies is their SFR. We will use one of the most prominent methods to estimate this parameter: the flux in the Hα emission line. In particular, we will use the flux estimated by the CIGALE SED fit, with the prescription in Kennicutt (1998) to transform Hα flux into SFR. We correct for extinction using the E(B-V) value from the SED fit and the Calzetti et al. (2000) extinction law with R_V=4.05 for consistency with the CIGALE implementation (us-
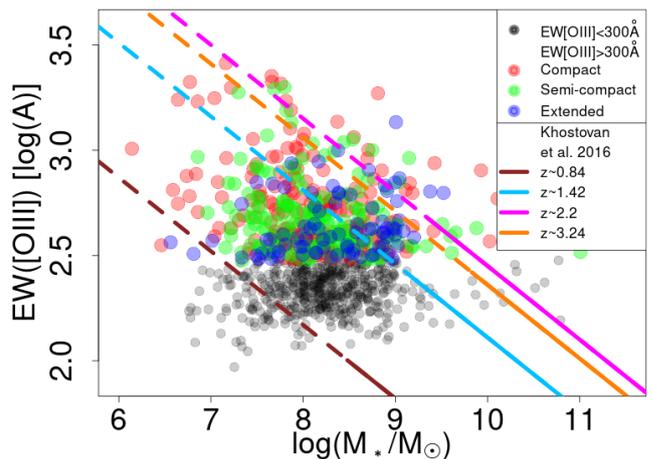
Fig. 19. [OIII] EW of the EELG (color dots) and candidate (gray dots) samples as a function of stellar mass. We plot in different colors the galaxies in the EELG sample according to their morphology: red corresponds to compact objects, green to semi-compact, and blue to extended. In addition, we overplot linear relations from Khostovan et al. (2016).
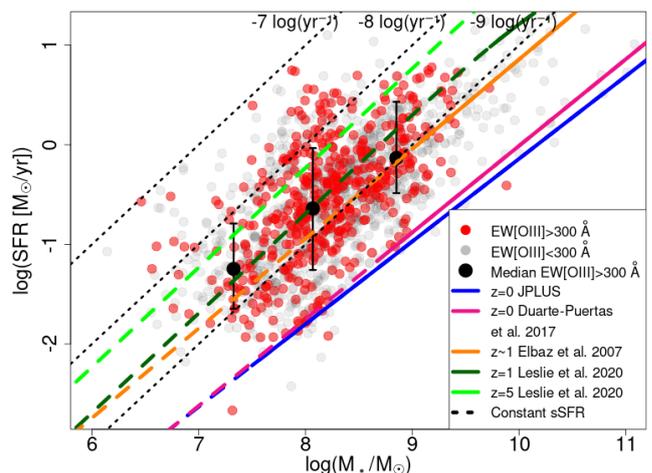
Fig. 20. SFR as a function of stellar mass for the galaxies in the EELG sample (red dots) and the rest of the candidate sample (grey dots). Large black dots represent the median values for the EELG sample grouped in mass bins. Colored lines represent main sequence fits from the literature, with black dotted lines represent constant sSFR.

ing R_V=3.1 would only create an offset with a median value of ∼ -0.04 dex in SFR). We use the Hα flux since it is the most direct SFR tracer that we can access with the J-PLUS data, and using other sources would entail other corrections beyond the scope of this paper. The Hα flux value, even if calculated using CIGALE (and thus not entirely independant from the stellar mass value we use), is in very good agreement with both photometric and spectroscopic estimations (see Figs. 9 and 14). Taking directly the SFR estimation by CIGALE would be much less independent and make it harder to compare with other works in the literature. In any case, we only use the SFR values to illustrate the extreme nature of our sample, not to provide any definitive physical conclusions.

A tight correlation has been proven to exist between the stellar mass of galaxies and their SFR often called "star formation

main sequence" (e.g. Brinchmann et al. 2004; Elbaz et al. 2007; Noeske et al. 2007; Speagle et al. 2014). This relation has been used to define both starburst galaxies (above the relation) and quiescent galaxies (below it). In the present work, our selection method is intentionally oriented to identify galaxies with very high SFR, and thus our data cannot be used to compute the main sequence parameters. In Fig. 20, we plot the SFR of the galaxies in our sample as a function of their stellar mass, as well as the constant specific SFR (sSFR) lines. We overplot several main sequence relations for different redshifts (Leslie et al. 2020; Elbaz et al. 2007), with the fits derived in Duarte Puertas et al. (2017) and Vilella-Rojo et al. (2021) as local comparisons. It is clear that our sample of EELGs lies well above the local main sequence relation, with sSFR values more comparable to those of typical high redshift star forming galaxies. We compute the difference between the SFR values measured in our sample and those expected if the galaxies followed exactly the Vilella-Rojo et al. (2021) relation. The typical difference is very high, with our EELG sample showing SFR values $1.31^{+0.55}_{-0.46}$ dex higher. In particular, the differences are higher at lower stellar masses, with $M_\star < 10^{7.5}\ M_\odot$ galaxies having $1.71^{+0.54}_{-0.48}$ dex difference and those with $M_\star > 10^{8.5}\ M_\odot$ presenting $1.08^{+0.34}_{-0.32}$ offset. This is partially due to selection effects, given that the detection of galaxies with low stellar masses and low sSFR is less likely. Nevertheless, the detection of high sSFR low mass galaxies and the lack of high sSFR high mass galaxies are not due to biases. This result is consistent with many works that have discovered how low mass galaxies undergo bursts of star formation more intense and more often than their higher mass counterparts, and therefore are more likely to be observed in a starburst phase (Sparre et al. 2017; Guo et al. 2016; Khostovan et al. 2021). This effect is actually already present (albeit softer) in typical star-forming galaxies, hence the slope of the main sequence being lower than one.

This differences put the vast majority of our sample above the threshold for galaxies undergoing starbursts, which is usually set around 0.48-0.6 dex above the main sequence (Rodighiero et al. 2011; Elbaz et al. 2018; Bluck et al. 2020). They are in fact more similar to main sequence galaxies at very high redshift, with some reaching values that are typical at z ~ 5, considering the extrapolation to low mass galaxies of the main sequence at that redshift.

4.2.3. Comparison with typical low redshift star-forming galaxies

Some properties derived from the SED fitting process can be complex to evaluate, given the uncertainties in the models, the assumptions taken (most notably the SFH), and the relatively sparse wavelength coverage of our data. In order to provide a fair comparison with the literature, we performed the same SED fitting analysis, adopting the same assumptions on a different sample of galaxies. The absolute values obtained for ages, metallicities, etc. may differ from others due to our choices in the SED fitting process and not due to physical differences in the galaxies, but the relative difference in those magnitudes when applying the same method provides more accurate insights.

In order to do this, we choose the sample from Vilella-Rojo et al. (2021) as a comparison. In that work, the authors used the J-PLUS first data release (DR1; Cenarro et al. 2019) to identify 805 local (z < 0.017) star-forming galaxies selected as Hα emitters. The comparison is therefore straightforward, given that the filters used are the same. `CIGALE` is run over the Vilella-Rojo

et al. (2021) sample of Hα emitters, with the same parameters used for our analysis (Sect. 3.1.1). Even if their galaxies are similar to ours, they selected brighter (r < 18 mag) objects, at a slightly lower redshift. Therefore, their typical masses are higher than ours, with median and 1σ limits of $\log(M_\star/M_\odot) = 8.64^{+1.12}_{-0.88}$ compared to $\log(M_\star/M_\odot) = 8.15^{+0.55}_{-0.59}$. In order to correct for this difference, we limit our analysis to their galaxies showing stellar masses below $10^{8.9} M_\odot$, which lowers their typical mass values to $\log(M_\star/M_\odot) = 8.21^{+0.44}_{-0.80}$, in agreement with our sample (see Panel d) in Fig. 21).

The results of the comparison between the other SED parameters is shown in the rest of the panels in Fig. 21 (we only plot the most relevant histograms for this analysis). The most striking difference appears in the age of the young stellar populations, with extremely low values for our extreme [OIII] emitters sample ($3.0^{+2.7}_{-1.0}$ Myr) compared to more typical values for the Vilella-Rojo et al. (2021) sample (7^{+2}_{-1} Myr). This result was expected, considering the difference in the sample selection in their work (with a minimum EW in Hα of around 12 Å) and ours (with a minimum EW in [OIII] of 300 Å), resulting in 95% of our galaxies showing Hα EWs higher than 56 Å, and 90% above 135 Å. It is clear in all models of star formation that the EW of the emission lines decreases rapidly after the initial burst, in a few Myrs. Therefore, galaxies with higher EW values will tend to have younger bursts of star formation. Other parameters (fraction of young population, metallicity, ionization parameters) will also play a role, but the very high EW in our work necessarily implies very young ages.

In line with the EW differences stated, we see that the mass fraction of the young population is slightly higher in our sample than in Vilella-Rojo et al. (2021) with $\log(M_{young}/M_{old}) = -2.0^{+0.7}_{-0.3}$ and $\log(M_{young}/M_{old}) = -2.3^{+1.0}_{-0.3}$, respectively. The metallicities (not shown in the figure) are indistinguishable between the samples, with a p-value of 0.27 in the Kolmogorov-Smirnoff test. The ionization parameter shows a difference going from $\log U = -3.0^{+1.0}_{-0.5}$ in their sample to $\log U = -2.5^{+1.0}_{-0.5}$ in ours. However, it must be kept in mind that mass fraction and ionization parameter are subject to degeneracies and are hard to accurately measure just with the photometric data we have.

While the results presented in this work are insufficient to explain why the EELGs are undergoing such extreme events, some inferences can be made. The very young ages of the bursts indicate that it is likely that these galaxies only show these properties for briefs periods of time, which explains in part why they are so uncommon. Many more galaxies may have undergone similar starburst phases in other moments of their SFH (see e.g. Sánchez Almeida et al. 2008, 2018). The trigger of the extreme star formation events remains an open question, yet some hypothesis have been presented in the literature. For decades, mergers have been considered likely to be causing a fraction of the events (Barnes & Hernquist 1991), and more recently their importance has also been shown in dwarf galaxy starbursts both from simulations (Bekki 2008) and observations (Stierwalt et al. 2015). Disrupted morphologies have also been observed with increased frequency in EELGs (e.g. Calabrò et al. 2017). While in the present work most galaxies are compact, there is a significant fraction that show extended, often complex and clumpy morphologies. In addition, while some may appear undisturbed, traces of a recent merger may be hidden due to their low surface brightness (Martínez-Delgado et al. 2012). Another key driver for starbursts is the infall of cold gas into the galaxy (Ceverino et al. 2010; Sánchez Almeida et al. 2015), which is inline with the typically low metallicity of these systems, especially if the
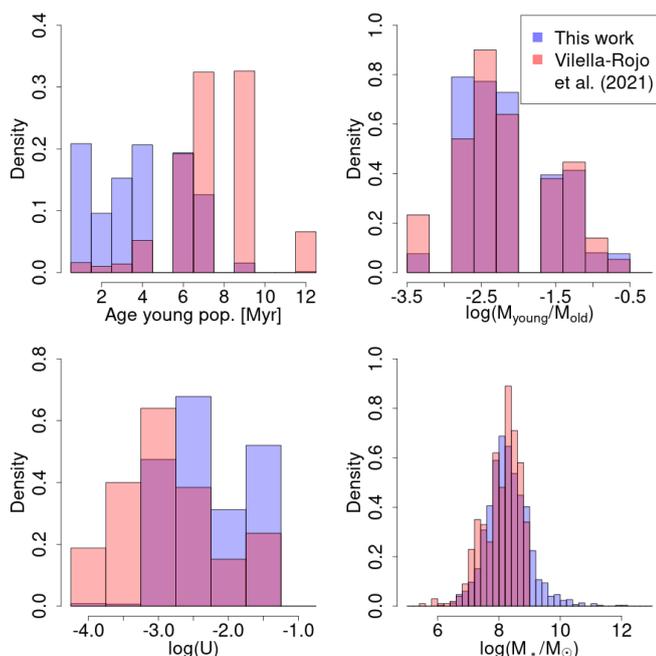
Fig. 21. Histograms comparing the properties of our EELG sample (blue) with a subsample of typical star forming galaxies from Vilella-Rojo et al. (2021) (red). We show the best fitting values of several magnitudes, obtained fitting J-PLUS photometry to SED models using CIGALE. From left to right and top to bottom, they are the age of the young stellar population, the mass ratio between the young and the old populations, the ionization parameter, and the stellar masses.

gas is near pristine and falling from the cosmic web. This process also explains why the EELGs are more common at higher redshift, where cold gas was much more abundant. Another physical process that could trigger the starbursts is the star formation feedback Sparre et al. (2017). It could be due to the removal of gas from the galaxy, that later "rains down" on it, causing he starburst, or by compressing the gas inside the galaxy, triggering further star formation (Tenorio-Tagle et al. 2005). While disentangling the causes of the extreme events of star formation is beyond the scope of this paper, the sample presented here it can be useful in that regard in future analyses.

5. Summary and conclusions

We have used the J-PLUS DR2 to select a sample of 466 EELGs in the local Universe over 2 176 square degrees, with EW[OIII] above 300 Å. Out of theme, 411 (89%) were previously unclassified as EELGs.

To build the sample, we have identified objects with excess of emission in the unique ~ 200Å J-PLUS $J0515$ filter compared to the r band. To remove contaminants (stars and high redshift systems), we rely in color-color and color-magnitude diagrams using infrared WISE observations, reaching 96% purity and 92% completeness considering the available SDSS spectra.

We perform SED fitting on the J-PLUS data using the CIGALE software, finding the following key properties of the sample of galaxies:

- Low stellar masses, with typical values of $\log(M_\star/M_\odot) = 8.13^{+0.61}_{-0.58}$.
- Very young ages, $3^{+2.7}_{-2}$ Myr, the main driver of the very high EW.

- Moderately low metallicities, Z=$0.008^{+0.012}_{-0.004}$, and high ionization parameters, $\log U = -2.50^{+1.00}_{-0.35}$, contributing as well to the exceptionally high EW.
- Low extinction, $E(B-V) = 0.1^{+0.2}_{-0.1}$, typical of low mass galaxies with high star formation activity.

We have compared our sample with several results in the literature. We find a very good agreement for those galaxies that have spectroscopic data, comparing both line fluxes and EW. We measure negligible offsets in the comparisons between spectroscopic and our photometric J-PLUS measurements in the Hα and [OIII]5007 lines, with very little scatter in fluxes (~ 0.18 dex) and EWs (~ 0.1 dex).

We analyze as well the efficiency of our selection process compared to other recent results. We find that the number density of EELGs we calculate is in agreement with magnitude-limited spectroscopic surveys such as the main SDSS legacy survey and GAMA, while reaching fainter galaxies than both and much wider areas than GAMA. We measure comparable densities as narrow-band surveys, with caveats due to selection criteria. We are, in contrast, much more efficient than searches made using broadband surveys, such as the blueberry galaxies (Yang et al. 2017) or the green peas (Cardamone et al. 2009). We find 20 - 50 times more EELGs per unit of volume, when controlling for EW and magnitude or stellar mass. It is also likely that we are able to access lower metallicity systems than broadband surveys, since we are not directly biased against high Hα/[OIII]5007 systems, which is a property of extremely metal poor galaxies (Senchyna & Stark 2019). The [OIII]5007 EW and SFR as a function of mass diagrams place our sample of EELGs in the typical values for the high redshift Universe (z $\sim 2 - 5$) while being located at z < 0.06.

In conclusion, we have presented a sample of mostly previously unclassified EELGs, which despite residing in the local Universe, share characteristics (mass, sSFR, EW, metallicity) that make them similar to those at high redshift. Therefore, spectroscopic follow-up of this sample may shed light into the properties of the galaxies forming in the very early Universe.

References

Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, ApJS, 249, 3
Aihara, H., Allende Prieto, C., An, D., et al. 2011, ApJS, 193, 29
Amorín, R., Aguerri, J. A. L., Muñoz-Tuñón, C., & Cairós, L. M. 2009, A&A, 501, 75
Amorín, R., Pérez-Montero, E., Contini, T., et al. 2015, A&A, 578, A105
Amorín, R. O., Muñoz-Tuñón, C., Aguerri, J. A. L., Cairós, L. M., & Caon, N. 2007, A&A, 467, 541
Arrabal Haro, P., Rodríguez Espinosa, J. M., Muñoz-Tuñón, C., et al. 2020, MNRAS, 495, 1807
Atek, H., Siana, B., Scarlata, C., et al. 2011, ApJ, 743, 121
Bacon, R., Conseil, S., Mary, D., et al. 2017, A&A, 608, A1
Barnes, J. E. & Hernquist, L. E. 1991, ApJ, 370, L65
Bekki, K. 2008, MNRAS, 388, L10
Benítez, N., Dupke, R., Moles, M., et al. 2014, arXiv e-prints [arXiv:1403.5237]
Benítez, N., Gaztañaga, E., Miquel, R., et al. 2009, ApJ, 691, 241
Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393
Bluck, A. F. L., Maiolino, R., Piotrowska, J. M., et al. 2020, MNRAS, 499, 230
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Boquien, M., Kennicutt, R., Calzetti, D., et al. 2016, A&A, 591, A6
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000
Calabrò, A., Amorín, R., Fontana, A., et al. 2017, A&A, 601, A95
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, MNRAS, 399, 1191
Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, ApJS, 189, 270
Catalán-Torrecilla, C., Gil de Paz, A., Castillo-Morales, A., et al. 2015, A&A, 584, A87
Cava, A., Pérez-González, P. G., Eliche-Moral, M. C., et al. 2015, ApJ, 812, 155
Cenarro, A. J., Moles, M., Cristóbal-Hornillos, D., et al. 2019, A&A, 622, A176
Cenarro, A. J., Moles, M., Marín-Franch, A., et al. 2014, in Proc. SPIE, Vol. 9149, Observatory Operations: Strategies, Processes, and Systems V, 91491I
Ceverino, D., Dekel, A., & Bournaud, F. 2010, MNRAS, 404, 2151
Chabrier, G. 2003, PASP, 115, 763
Cook, D. O., Kasliwal, M. M., Van Sistine, A., et al. 2019, ApJ, 880, 7
Cristóbal-Hornillos, D., Gruel, N., Varela, J., et al. 2012, in SPIE CS, Vol. 8451
De Barros, S., Oesch, P. A., Labbé, I., et al. 2019, MNRAS, 489, 2355
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Driver, S. P. & Robotham, A. S. G. 2010, MNRAS, 407, 2131
Duarte Puertas, S., Vilchez, J. M., Iglesias-Páramo, J., et al. 2017, A&A, 599, A71
Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33
Elbaz, D., Leiton, R., Nagar, N., et al. 2018, A&A, 616, A110
Fumagalli, M., Patel, S. G., Franx, M., et al. 2012, ApJ, 757, L22
Garn, T. & Best, P. N. 2010, MNRAS, 409, 421
Geach, J. E., Smail, I., Best, P. N., et al. 2008, MNRAS, 388, 1473
González Delgado, R. M., Díaz-García, L. A., de Amorim, A., et al. 2021, A&A, 649, A79
Gronwall, C., Ciardullo, R., Hickey, T., et al. 2007, ApJ, 667, 79
Gronwall, C., Jangren, A., Salzer, J. J., Werk, J. K., & Ciardullo, R. 2004, AJ, 128, 644
Guo, Y., Koo, D. C., Lu, Y., et al. 2016, ApJ, 822, 103
Guseva, N. G., Izotov, Y. I., Fricke, K. J., & Henkel, C. 2017, A&A, 599, A65
Haro, G. 1956, Boletin de los Observatorios Tonantzintla y Tacubaya, 2, 8
Hernán-Caballero, A., Hatziminaoglou, E., Alonso-Herrero, A., & Mateos, S. 2016, MNRAS, 463, 2064
Hernán-Caballero, A., Varela, J., López-Sanjuan, C., et al. 2021, arXiv e-prints, arXiv:2108.03271
Hinojosa-Goñi, R., Muñoz-Tuñón, C., & Méndez-Abreu, J. 2016, A&A, 592, A122
Hippelein, H., Maier, C., Meisenheimer, K., et al. 2003, A&A, 402, 65
Hood, C. E., Kannappan, S. J., Stark, D. V., et al. 2018, ApJ, 857, 144
Hsyu, T., Cooke, R. J., Prochaska, J. X., & Bolte, M. 2018, ApJ, 863, 134
Inoue, A. K. 2011, MNRAS, 415, 2920
Izotov, Y. I., Guseva, N. G., Fricke, K. J., et al. 2021, A&A, 646, A138
Izotov, Y. I., Guseva, N. G., & Thuan, T. X. 2011, ApJ, 728, 161
Jangren, A., Wegner, G., Salzer, J. J., Werk, J. K., & Gronwall, C. 2005, AJ, 130, 496
Jiang, T., Malhotra, S., Rhoads, J. E., & Yang, H. 2019, ApJ, 872, 145
Kakazu, Y., Cowie, L. L., & Hu, E. M. 2007, ApJ, 668, 853
Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, MNRAS, 341, 33
Kellar, J. A., Salzer, J. J., Wegner, G., Gronwall, C., & Williams, A. 2012, AJ, 143, 145
Kennicutt, Robert C., J. 1998, ApJ, 498, 541
Khostovan, A. A., Malhotra, S., Rhoads, J. E., et al. 2021, MNRAS, 503, 5115
Khostovan, A. A., Sobral, D., Mobasher, B., et al. 2015, MNRAS, 452, 3948
Khostovan, A. A., Sobral, D., Mobasher, B., et al. 2016, MNRAS, 463, 2363
Kojima, T., Ouchi, M., Rauch, M., et al. 2020, ApJ, 898, 142
Kron, R. G. 1980, ApJS, 43, 305
Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3
Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, ApJ, 899, 58
Logroño-García, R., Vilella-Rojo, G., López-Sanjuan, C., et al. 2019, A&A, 622, A180
López-Sanjuan, C., Tempel, E., Benítez, N., et al. 2017, A&A, 599, A62
López-Sanjuan, C., Varela, J., Cristóbal-Hornillos, D., et al. 2019, A&A, 631, A119
López-Sanjuan, C., Yuan, H., Vázquez Ramió, H., et al. 2021, A&A, in press [2101.12407]
Lumbreras-Calle, A., Muñoz-Tuñón, C., Méndez-Abreu, J., et al. 2019a, A&A, 621, A52
Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415
Maier, C., Meisenheimer, K., Thommes, E., et al. 2003, A&A, 402, 79
Marín-Franch, A., Taylor, K., Cenarro, J., Cristobal-Hornillos, D., & Moles, M. 2015, in IAU General Assembly, Vol. 29, 2257381
Markarian, B. E. 1967, Astrofizika, 3, 24
Martínez-Delgado, D., Romanowsky, A. J., Gabany, R. J., et al. 2012, ApJ, 748, L24
Maseda, M. V., van der Wel, A., Rix, H.-W., et al. 2018, ApJ, 854, 29
Matthee, J., Sobral, D., Hayes, M., et al. 2021, MNRAS, 505, 1382
Meisenheimer, K., Beckwith, S., Fockenbrock, H., et al. 1998, in Astronomical Society of the Pacific Conference Series, Vol. 146, The Young Universe: Galaxy Formation and Evolution at Intermediate and High Redshift, ed. S. D'Odorico, A. Fontana, & E. Giallongo, 134
Melnick, J., Terlevich, R., & Eggleton, P. P. 1985, MNRAS, 216, 255
Mendes de Oliveira, C., Ribeiro, T., Schoenell, W., et al. 2019, MNRAS, 489, 241
Moles, M., Benítez, N., Aguerri, J. A. L., et al. 2008, AJ, 136, 1325
Molino, A., Costa-Duarte, M. V., Mendes de Oliveira, C., et al. 2019, A&A, 622, A178
Nilsson, K. K., Östlin, G., Møller, P., et al. 2011, A&A, 529, A9
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
Onodera, M., Shimakawa, R., Suzuki, T. L., et al. 2020, ApJ, 904, 180
Ouchi, M., Mobasher, B., Shimasaku, K., et al. 2009, ApJ, 706, 1136
Pérez-González, P. G., Cava, A., Barro, G., et al. 2013, ApJ, 762, 46
Pérez-Montero, E., Kehrig, C., Vílchez, J. M., et al. 2020, A&A, 643, A80
Pirzkal, N., Rothberg, B., Ly, C., et al. 2013, ApJ, 772, 48
Reddy, N. A., Shapley, A. E., Sanders, R. L., et al. 2018, ApJ, 869, 92
Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJ, 739, L40
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Salpeter, E. E. 1955, ApJ, 121, 161
Salzer, J. J., Feddersen, J. R., Derloshon, K., et al. 2020, AJ, 160, 242
Salzer, J. J., Jangren, A., Gronwall, C., et al. 2005, AJ, 130, 2584
Sánchez Almeida, J., Elmegreen, B. G., Muñoz-Tuñón, C., et al. 2015, ApJ, 810, L15
Sánchez Almeida, J., Muñoz-Tuñón, C., Amorín, R., et al. 2008, ApJ, 685, 194
Sánchez Almeida, J., Olmo-García, A., Elmegreen, B. G., et al. 2018, ApJ, 869, 40
Sánchez Almeida, J., Pérez-Montero, E., Morales-Luis, A. B., et al. 2016, ApJ, 819, 110
Sargent, W. L. W. & Searle, L. 1970, ApJ, 162, L155
Schlafly, E. F., Meisner, A. M., & Green, G. M. 2019, ApJS, 240, 30
Searle, L. & Sargent, W. L. 1972, ApJ, 173, 25
Senchyna, P. & Stark, D. P. 2019, MNRAS, 484, 1270
Sobral, D., Best, P. N., Smail, I., et al. 2014, MNRAS, 437, 3516
Sobral, D., Matthee, J., Best, P. N., et al. 2015, MNRAS, 451, 2303
Sobral, D., Santos, S., Matthee, J., et al. 2018, MNRAS, 476, 4725
Sobral, D., Smail, I., Best, P. N., et al. 2013, MNRAS, 428, 1128
Sparre, M., Hayward, C. C., Feldmann, R., et al. 2017, MNRAS, 466, 88
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Spinoso, D., Orsi, A., López-Sanjuan, C., et al. 2020, A&A, 643, A149
Stierwalt, S., Besla, G., Patton, D., et al. 2015, ApJ, 805, 2
Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810
Stroe, A., Sobral, D., Matthee, J., Calhau, J., & Oteo, I. 2017a, MNRAS, 471, 2558
Stroe, A., Sobral, D., Matthee, J., Calhau, J., & Oteo, I. 2017b, MNRAS, 471, 2575
Tang, M., Stark, D. P., Chevallard, J., et al. 2021, MNRAS, 503, 4105
Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., et al. 2015, PASJ, 67, 104
Tenorio-Tagle, G., Silich, S., Rodríguez-González, A., & Muñoz-Tuñón, C. 2005, ApJ, 628, L13
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
van der Wel, A., Straughn, A. N., Rix, H. W., et al. 2011, ApJ, 742, 111
Vilella-Rojo, G., Logroño-García, R., López-Sanjuan, C., et al. 2021, A&A, 650, A68
Wegner, G., Salzer, J. J., Jangren, A., Gronwall, C., & Melbourne, J. 2003, AJ, 125, 2373
Wolf, C., Meisenheimer, K., Rix, H.-W., et al. 2003, A&A, 401, 73
Wolf, C., Meisenheimer, K., & Röser, H.-J. 2001, A&A, 365, 660
Yang, H., Malhotra, S., Rhoads, J. E., & Wang, J. 2017, ApJ, 847, 38
Zou, H., Zhou, X., Fan, X., et al. 2019, ApJS, 245, 4
Zwicky, F. 1966, ApJ, 143, 192

Appendix A: ADQL query

In this appendix we reproduce the ADQL query used in Section 2 to obtain the parent sample for this work. In addition to the main condition (significant emission in *J*0515 compared to *r*), described in Sect. 2, we include some other restrictions:

- Objects must have been detected with at least 3σ significance in *g*, *r*, and *J*0515.
- The flux in the *J*0515 and *g* filters cannot be compatible, considering the uncertainties. This is to ensure that there is significant emission in *J*0515, not just a very steep slope towards *r*.
- The *r* fluxes in both PSFCOR and AUTO photometry must be higher than certain threshold. This is only to ensure appropriate detections and avoid numeric errors, given that the thresholds are much lower than the final one imposed (*r*< 20 mag).
- Objects with FLAGS other than 0, 1, or 2 are rejected, only keeping those with either no issues (0), close neighbors (1), deblended (2), or a combination of those.
- Objects with any MASK_FLAG are rejected.

```
SELECT flu.*, z.*, s.*, e.*
FROM jplus.FLambdaDualObj flu,
jplus.PhotoZLephare z,
jplus.StarGalClass s, jplus.MWExtinction e
WHERE  flu.TILE_ID = z.TILE_ID
AND z.TILE_ID = s.TILE_ID
AND s.TILE_ID = e.TILE_ID
AND flu.NUMBER = z.NUMBER
AND z.NUMBER=s.NUMBER
AND s.NUMBER=e.NUMBER
AND flu.FLUX_AUTO[jplus::rSDSS]>44.8
AND flu.FLUX_PSFCOR[jplus::rSDSS]>17.8
AND flu.FLUX_RELERR_PSFCOR[jplus::rSDSS]
<0.333
AND flu.FLUX_RELERR_PSFCOR[jplus::gSDSS]
<0.333
AND flu.FLUX_RELERR_PSFCOR[jplus::J0515]
<0.333
AND (flu.FLUX_PSFCOR[jplus::J0515] -
flu.FLUX_PSFCOR[jplus::rSDSS])/
flu.FLUX_PSFCOR[jplus::rSDSS] > 1
AND (flu.FLUX_PSFCOR[jplus::J0515]
- flu.FLUX_RELERR_PSFCOR[jplus::J0515]
*flu.FLUX_PSFCOR[jplus::J0515])
>(flu.FLUX_PSFCOR[jplus::gSDSS]+
flu.FLUX_RELERR_PSFCOR[jplus::gSDSS]
*flu.FLUX_PSFCOR[jplus::gSDSS])
AND flu.FLAGS[jplus::rSDSS] <4
AND flu.MASK_FLAGS[jplus::rSDSS] = 0
AND flu.FLAGS[jplus::J0515] < 4
AND flu.MASK_FLAGS[jplus::J0515] = 0
```

Appendix B: Re-calculation of J-PLUS photometry for extended sources with blending flag

As mentioned in Section 2.2.6, some of the objects in our sample present photometric flags indicating that they have been deblended from other source. This, in some cases, just means that there was a different object nearby (such a star or a different galaxy), and the photometric data we are using only belongs to
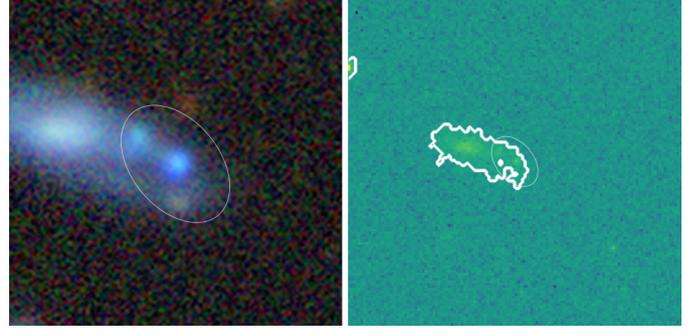
Fig. B.1. Example of a improperly deblended source, and the SExtractor run that covers the whole galaxy

the galaxy we are interested in. Nevertheless, in some cases, the nearby object is in fact the main galaxy body, and the object we are detecting is only a specific region of the galaxy. Keeping these regions as "galaxies" would bias our analysis towards lower masses and higher EWs. In order to avoid this, we run SExtractor over the J-PLUS images.

Appendix B.1: Visual classification

We inspect the 371 objects with deblending flag in the main sample, selecting those where our object of interest is in fact part of a larger galaxy. While can be ultimately a "judgment call", we follow some rules to use our own SExtractor photometry of an object instead of the original J-PLUS one:

- If in the Legacy survey image (or in the PANSTARRS, if outside the Legacy footprint) there is a clear connection of blue emission between our object and an extended galaxy that is not covered by the J-PLUS aperture.
- If there is a large galaxy nearby, with several HII regions showing similar color as our object, at similar distances.

Some cases were not considered as improper deblending:

- If the color of the object near our target is very red and no clear structure is found between them, we consider that they do not belong to the same physical object and keep the J-PLUS original photometry.
- If the object and our target have a similar size and color, and there is no emission between them, we consider them as satellites and keep the original J-PLUS photometry.

In some cases, particularly involving mergers, the decision between considering a system one galaxy or several based upon the photometry is, to a certain degree, arbitrary. We have chosen to keep the original photometry if the system is very disturbed and our object is physically separated from the larger galaxy or galaxies. We have as well preferred to keep the original photometry in cases were the SExtractor apertures were contaminated significantly by other nearby objects, that were however masked out in the J-PLUS photometry. After this visual inspection, we identified 175 galaxies where the original J-PLUS photometry was appropriate, and 179 where additional SExtractor runs where necessary to capture all the flux from the galaxy. We identified in addition 20 spurious objects, that were removed from the sample. Most of them were spikes from bright stars, measured over large apertures.

An example of a galaxy with deblending flag where we decided to keep the SExtractor photometry is shown in Fig. B.1.

Table B.1. Parameters for the SExtractor runs in blended objects

	Set 1	Set 2
DETECT_MINAREA	20	20
DETECT_THRESH	1.1	0.6
ANALYSIS_THRESH	3	3
DEBLEND_NTHRESH	1	1
DEBLEND_MINCONT	0.05	0.005

Appendix B.2: SExtractor runs

We run SExtractor over the sample of 179 galaxies that were not accurately represented by the original J-PLUS photometric aperture, using the default SExtractor parameter set except for a few parameters shown in Table B.1 (Set 1). After the first run using Set 1 we inspected visually the result, and considered that for 36 of these galaxies neither the SExtractor run nor the original photometry were appropriate. In most of those cases the our target was a region of a larger galaxy, but both the original photometry and our first SExtractor run failed to identify them as the same object. We chose to run SExtractor with a second set of parameters (Set 2 in Table B.1). This was enough to recover accurate photometry for 27 out of the 36 galaxies. The remaining 9 were removed from the sample, since in any case they are so extended that they would not make the EW cut to join the EELG sample.

Appendix C: CIGALE SED fits and parameters

In this Appendix we will review in more detail the SED fitting process and describe in full the CIGALE input parameters we used. We reproduce here the full set of input parameter in the main SED fit presented in this work.

```
  [[sfh2exp]]

e-folding time of the main stellar
population model in Myr.
tau_main = 50

e-folding 1 of the late starburst
population model in Myr.
tau_burst = 1

Mass fraction of the late burst
population.
f_burst = 0.0005, 0.0025,
0.005, 0.01, 0.03, 0.05, 0.075

Age of the main stellar population in the
galaxy in Myr. The precision is 1 Myr.
age = 200, 500, 1000, 2000, 5000

Age of the late burst in Myr.
The precision is 1 Myr.
burst_age = 1, 2, 3, 4, 5, 7, 12

Value of SFR at t = 0 in M_sun/yr.
sfr_0 = 1.0

Normalise the SFH to produce one
solar mass.
normalise = True
```

```
  [[bc03]]

Initial mass function: 0 (Salpeter) or
1 (Chabrier).
imf = 0

Metalicity. Possible values are:
0.0001, 0.0004, 0.004, 0.008, 0.02, 0.05.
metallicity = 0.0001, 0.0004, 0.004,
0.008, 0.02


  [[nebular]]

Ionisation parameter
logU = -4.0, -3.5, -3.0, -2.5, -2.0, -1.5

Fraction of Lyman continuum photons
escaping the galaxy
f_esc = 0.0, 0.2

Fraction of Lyman continuum photons
     absorbed by dust
f_dust = 0.0

Line width in km/s
lines_width = 300.0

Include nebular emission.
emission = True


  [[dustatt_calzleit]]

E(B-V)*, the colour excess of the stellar
continuum light for the young population.

E_BVs_young = 0, 0.1, 0.2, 0.3, 0.5

Reduction factor for the E(B-V)* of the old
population compared to the young one (<1).
E_BVs_old_factor = 0.44, 1

Central wavelength of the UV bump in nm.
uv_bump_wavelength = 217.5

Width (FWHM) of the UV bump in nm.
uv_bump_width = 35.0

Amplitude of the UV bump.
For the Milky Way: 3.
uv_bump_amplitude = 0.0

Slope delta of the power law modifying
the attenuation curve.
powerlaw_slope = 0.0
```

It is important to note that they main focus of the SED fits is to compute accurate values for the continuum and the emission line fluxes, while also providing a stellar mass estimate. The specific results for the stellar population parameters are subject to our choices in the free parameters and several physical degeneracies, thus extracting high level scientific conclusions from

them is very challenging. We perform several CIGALE runs on fractions of the main sample, with variations on the input parameters, to test the effect that different choices may have on the results.

The SFH we choose, with two single stellar populations, represents two almost instantaneous bursts of star formation. Other possibilities may as well provide good fits to the data, for instance, including options for a more extended SFH in the old stellar populations (τ_{old}=50, 500, 1000 Myr). This would provide accurate fits as well, while making almost no impact in the main results (negligible in log(U) or E(B-V), and just 0.1 dex in stellar mass, 0.02-0.04 dex in emission line fluxes, and 0.6 Myr in age of the young population). Even comparing the age of the old stellar population we only see a small scatter (0.15 dex) and no offset.

In our main run we only considered two values (0 and 0.2) for the fraction of Lyman continuum photons escaping the galaxy (f_esc), and only one value (0) for the fraction of Lyman photons absorbed by dust (f_dust). We included some variation to provide more flexible fits, but we make no attempt at extracting physical conclusions from these parameters. Increasing the set of possible values to f_esc = 0, 0.2, 0.5 and f_dust=0, 0.5 yields very similar results to the original SED fits: the median values of the difference between parameters remain virtually the same, with only small 1σ scatters: 0.08 dex in stellar mass, 0.02-0.04 dex in emission line fluxes, 1.5 Myr in age of the young population, 0.05 mag in E(B-V), and negligible in log(U).

Finally, we changed our assumption of a Salpeter (1955) IMF to a Chabrier (2003) one. In this case, we see the expected 0.23 dex offset in the stellar mass, along with a 0.16 dex scatter. The offsets in the rest of the parameters are negligible, with low scatter, as in the previous tests: 0.04 - 0.07 dex in emission line fluxes, 1.5 Myr in the age of the young population, 0.1 mag in E(B-V), and negligible in log(U).