The Milky Way Project: New Galaxies in the Zone of Avoidance[*]

R. J. Simpson[1]†, K. Clements[2], A. Gazzard[3], B. Simmons[1], C. J. Lintott[1]

[1] *Oxford Astrophysics, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK*
[2] *Wheatley School, Oxford, UK*
[2] *A. N. Other School, Oxford, UK*

Draft form 15 February 2013.

ABSTRACT

We report the discovery of 35 previously unknown galaxies in the Zone of Avoidance. We also present a method for identifying galaxies through data gathered from the citizen science website 'The Milky Way Project'. 43 candidate galaxies are identified through correlation of markings from volunteers, inspecting Spitzer GLIMPSE and MIPSGAL survey data on 'The Milky Way Project' website. The characteristics of these candidate, such as their locations, spectral energy distributions and redshift velocities are used to confirm whether they are galaxies. The magnitudes and integrated flux of the candidates are calculated in GLIMPSE bands 3.6μm, 4.5μm, 5.8μm and 8.0μm, and MIPSGAL band 24μm. The velocities of these 35 newly uncovered galaxies appear to correlate with the velocities of over-densities in the large-scale structure, at very low Galactic latitudes either side of the Zone of Avoidance. Our findings therefore suggest that the large-scale structure continues throughout the Zone of Avoidance, and highlight the potential for many more galaxies to be found in far-infrared survey data.

Key words: H II regions – infrared: ISM – ISM: dust – stars: formation

1 INTRODUCTION

The Zone of Avoidance (ZoA) remains a literal gap in our knowledge of the large-scale structure of galaxies in the universe, and prevents us from determining the dynamics of our local Universe. Redshift surveys have steadily improved our knowledge of the large-scale structure above and below the ZoA (e.g., CfA Redshift Survey, Geller & Huchra 1989; 2dF Redshift Survey, Colless 1999; Las Campanas Redshift Survey, Shectman et al. 1996). However, the precise way the coherent structures of the two halves connect remains largely unknown, and currently needs to be simulated – a problem highlighted by the case of the 'Great Attractor' at Galactic longitude l 320° and latitude b 0° (Kolatt et al. 1995).

The Spizter GLIMPSE and MIPSGAL surveys (Benjamin et. al. 2003, Carey et al. 2008), surveyed the Galactic midplane in a multitude of infrared wavelengths. These landmark surveys provided the opportunity to occaisonally see through the obscuring material of our own Galaxy. One of the first extragalactic results from GLIMPSE/MIPSGAL was the Jarrett et al. (2007) discovery of two galaxies obvserved through our Galaxy, toward of the Great Attractor.

Following this Marleau et al. (2008) demonstrated further the discovery potential of Spitzer in the ZoA by revealing 25 further highly-obscured galaxies.

2 THE MILKY WAY PROJECT

The Milky Way Project (MWP, Simpson et al. 2012) is the ninth online citizen science project created using the Zooniverse[1] platform (Smith et al. 2011). The Zooniverse platform supports the activities of all Zooniverse citizen science projects. Built originally for Galaxy Zoo 2 this software, and successive versions of it, have now been used by more than 20 different projects across a range of research disciplines. The Zooniverse toolset is designed primarily as a way to of serving a large collection of 'assets' (for example, images or video) to a user interface, and collecting back user-generated interactions with these assets.

Galaxy Zoo (Lintott et al. 2008, 2011) and the larger suite of Zooniverse projects have successfully built a large community of volunteers[2] eager to participate in scientific

---

† Email: robert.simpson@astro.ox.ac.uk

[1] http://www.zooniverse.org
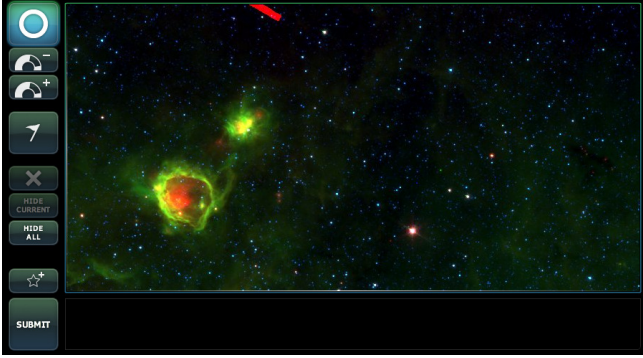[2] Over 750,000 registered volunteers at time of writing.

Figure 1. Screenshot of the Milky Way Project user interface. Colour figure available online.

activities. The Zooniverse has shown that enlisting these 'citizen scientists' via the internet is a powerful way to analyze large amounts of data. Human brains excel at pattern recognition tasks, and most people will reach a level of accuracy as high as any expert after a brief introduction. By enlisting citizen scientists, researchers can extend visual classification to large samples of images, having each image examined by a large number of independent classifiers. This allows researchers to tap into the 'wisdom of the crowd' effect where the consensus of a group of non-experts is often more accurate than the testimony of a single expert.

Another advantage of enlisting human classifiers is their ability to recognize unusual objects (e.g. galaxies, star clusters, dark clouds in the MWP) which computer search algorithms may be unable to spot. This has been shown by the serendipitous discovery of Hanny's Voorwerp (Lintott et al. 2009) and the case of the Galaxy Zoo 'Green Peas' (Cardamone et al. 2009). Spitzer GLIMPSE data is ideally suited to classification by citizen scientists as the amount of data is large and the images contain complex, overlapping structures that are extremely difficult to disentangle using automated algorithms.

The assets in the MWP are multiband, false-colour JPEG images, created by gridding the *Spitzer* GLIMPSE and MIPSGAL mosaics into smaller images at three different zoom levels. The highest zoom level provides users with tiles of $0.3° \times 0.15°$, and at a resolution of 800×400 pixels these tiles nearly reproduce the $1.2''$ pixel scale of the GLIMPSE survey images. Larger tile sizes of $0.75° \times 0.375°$ and $1.5° \times 0.75°$ were also generated. The tiles were plotted in an overlapping grid to allow all parts of the inner Galactic plane ($|l| \leq 65°$, $|b| \leq 1°$) to be viewed by the MWP users, at all zoom levels. To provide an optimal representation of the dynamic range within each tile, each of 3 single-band images was independently scaled to a square-root stretch function (with the faintest 5% of image pixels clipped to black and the brightest 0.2% clipped to white), assigned to a colour channel (red=24 μm, green=8.0 μm, blue=4.5 μm), and finally composited into a 3–colour image. The MIPSGAL 24 μm mosaics frequently saturate in regions of bright nebulosity, and saturated 24 μm pixels were set to maximum red to preserve the visual appeal of the images and to avoid presenting MWP users with saturation artefacts. The resulting composite images allow visual identification of both bright and faint features within a given image tile.

The MWP user interface (see Figure 1) was built using Flash, based upon the pre-existing Moon Zoo interface (Joy et al. 2011). Volunteers are primarily encouraged to draw ellipses onto the image to mark the locations of bubbles. A short, online tutorial shows how to use the tool, and examples of prominent bubbles are given[3]. As a secondary task, users can also mark rectangular areas of interest, which can be labelled as small bubbles, green knots, dark nebulae, star clusters, galaxies, fuzzy red objects or 'other'. Examples of these are also given in a tutorial on the website, and it is the galaxies that will be dealt with in this study. Users can add as many annotations as they wish before submitting the image, at which point they are given another image for annotation. Each image's annotations are stored in a database as a classification, and users can see the images they have classified in a part of the site called 'My Galaxy'. Users can only classify a given image once.

Volunteers are primarily asked to draw bubbles on the images, by placing elliptical annuli that match the structures they see in the data. They are also encouraged to mark other regions interest on an image, marking these are either star clusters, green knots, etc. Do do so, volunteers simply draw rectangles. Since these objects are secondary to the main bubble-finding task, the site was designed so that they should be simple and quick to mark. Simple rectangles allow us to record the positions and approximate sizes of any these other interesting objects.

3   CATALOGUE CONSTRUCTION

The galaxy candidates presented here are derived from 30,529 galaxy annotations by volunteers on the MWP – literally drawn as boxes on the sky at locations where they believed there *might* be a galaxy. These boxes were gridded together on the basis of position and size, with a grid size equal to the smallest MWP annotation size of 20 pixels. The resulting list of 1,003 unique objects each have a calculated averaged position, height and width as well as a ratio of number-of-times-viewed to number-of-times-drawn by MWP volunteers (known as the 'hit rate'). Only those objects drawn by at least three independent volunteers, and with a hit-rate of at least 10%, were retained. This quality control step ensures that most spurious and unintentional drawings do not contaninate the reduced set and a similar measures is taken in Simpson et al. (2012). The resultant 87 objects were visually inspected. Any obvious outliers or misclassifications were removed (e.g. stars, diffraction spikes and small bubbles) and the final set of objects forms the list of 43 candidates shown in Table 3.

3.1   Aperture photometry

Data from GLIMPSE (Benjamin et. al 2003) and MIPSGAL (Carey et. al. 2008) were downloaded from the NASA/IPAC Infrared Science Archive[4] for each of the 43 candidate galaxies. These were in the form of cutouts in each of the four IRAC (Fazio et al. 2004) bands (3.6, 4.5, 5.8 and 8μm) and

---

[3] http://www.milkywayproject.org/tutorial
[4] http://irsa.ipac.caltech.edu

Table 1. Table of galaxy parameters, including galactic coordinates, corrected magnitudes for Spitzer IRAC bands 1, 2, 3 and 4, and the MIPS 24μm flux.

ID	GLON (deg)	GLAT (deg)	3.6μm (mag)	4.5μm (mag)	5.8μm (mag)	8μm (mag)	24μm (mag)
1	004.548483	−0.781395	13.80	13.53	12.35	10.76	7.09
2	011.818667	−0.025800	11.88	11.80	9.84	8.41	4.69
3	012.741267	−0.000400	10.10	10.03	10.15	9.46	5.19
4	017.225237	−0.584033	11.78	11.86	11.73	9.98	7.05
5	017.784417	+0.034222	11.30	11.77	10.39	8.70	5.88
6	028.141688	−0.192917	11.57	11.72	11.26	9.82	
7	028.729475	−0.539626	12.85	12.72	11.50	9.53	5.32
8	036.399831	−0.466475	12.84	13.06	13.46	11.12	5.27
9	041.581565	−0.884437	13.86	13.70	12.88	11.85	8.86
10	042.908694	−0.888909	14.23	13.88	13.45	11.27	8.13
11	047.824267	−0.390300	13.78	13.98	12.04	10.55	8.16
12	048.707895	+0.961547	13.25	13.38	12.60	11.42	7.93
13	054.574569	−0.511079	13.40	13.47	12.01	10.56	7.83
14	054.657243	−0.256119	12.94	12.83	11.52	9.89	6.62
15	055.307833	+0.810834	13.34	13.57	12.01	10.51	7.56
16	057.474400	−0.387300	13.44	13.28	12.11	10.52	7.14
17	058.441232	−0.470409	12.86	12.98	12.10	10.87	6.21
18	059.217833	+0.396467	14.39	14.78	15.41	12.04	8.35
19	059.545904	+0.776002	13.58	13.49	12.66	10.62	7.81
20	061.350500	−0.108500	13.98	14.02	13.44	11.62	8.66
21	063.349203	−0.768282	14.00	14.02	13.07	11.60	8.58
22	299.362072	+0.731691	14.22	14.80	15.95	12.44	9.98
23	299.398377	+0.888066	12.84	12.90	12.45	10.89	7.49
24	302.037426	−0.528374	13.12	12.96	11.62	9.65	6.68
25	302.373292	−0.023500	13.91	12.68	11.73	11.37	6.72
26	306.681500	+0.015125	10.86	10.80	10.47	9.82	6.56
27	314.143892	−0.682605	14.69	14.14	13.28	12.36	8.65
28	316.873061	−0.599138	12.03	12.04	10.99	9.64	6.47
29	317.039057	−0.497820	12.77	12.75	11.34	9.75	5.64
30	318.527084	+0.504229	13.02	12.89	12.20	10.60	7.08
31	318.729619	−0.097762	12.44	12.17	11.07	9.42	5.28
32	321.402694	+0.652191	11.52	11.27	9.84	8.33	5.12
33	322.552422	+0.815422	11.48	11.58	11.88	11.68	5.24
34	324.323250	+0.317000	13.48	13.39	12.26	10.96	7.90
35	325.245517	−0.799997	12.95	13.20	11.87	10.38	7.16
36	333.445875	+0.356375	12.10	11.86	10.60	9.06	7.23
37	335.104583	−0.066792	11.30	11.09	8.95	7.50	5.68
38	335.517000	−0.378042	11.02	11.05	10.38	9.12	5.24
39	341.857467	−0.842779	13.52	13.29	12.12	10.40	7.67
40	350.457146	−0.041708	13.45	13.46	14.21	13.08	7.47
41	350.101407	+0.389093	12.81	12.42	11.34	9.87	5.39
42	357.148741	−0.572130	13.61	13.47	12.53	10.89	6.83
43	357.737833	−0.092033	12.96	12.69	10.50	9.13	5.71

in the 24μm MIPS (Rieke et al. 2004) band. Three-colour IRAC composites for each candidate are shown in Figures 3 and 3. The original MWP images seen by users, are included in the appendix, alongside a selection of other data for each candidate.

The spectral energy disturbtion for each candidate was derived using photomtetric measurements from these data. The aperture size and position for each candidate were created using the visual controur of the object in the 8μm band in order to enclose as much flux as possible from the object, without encompassing nearby stars. The background contributing to the total flux density within a similarly-sized sky aperture was removed.

$$correction\,factor = \frac{true\,flux}{flux} = [Ae^{-radius^B}] + C \qquad (1)$$

The IRAC flux densities were multiplied by the aperture correction factor according to the formulation documented by the IRAC team of the Spitzer Science Centre website[5] (see Equation 1). The zero magnitude flux densities for IRAC were taken from Reach et al. (2005) ans for MIPS from Engelbracht et al. (2007). They are 280.9 (IRAC 3.6μm), 179.7 (IRAC 4.5μm), 115.0 (IRAC 5.8μm), and 64.9 Jy (IRAC 8μm) and 7.17 Jy (MIPS 24μm). Finally, the extinction corrections at these wavelengths were obtained us-

---

[5] http://ssc.spitzer.caltech.edu/irac/calib/extcal/

ing the transformation given in Indebetouw et al. (2005), which is shown (without uncertainties) as Equation 2.

$$\log\left[\frac{A\lambda}{AK}\right] = 0.61 - 2.22\log(\lambda) + 1.21[\log(\lambda)]^2 \qquad (2)$$

All extinction-corrected photometric measurements are given in Table 3. Uncertainties in the flux density measurements varied with passband, and we assumed flux density uncertainties of the order of 20%.

4   RESULTS

4.1   Spectral energy distributions

Figures AA1–AA23 show the SED distributions of the candidate galaxies alongside models of normal star-forming galaxies from Dale et al. (2001) and Dale & Helou (2002). These models cover cover a wide wavelength range, from 3 to 1100 m, but are described by only the parameter, fν (60 μm)/fν (100 μm). This range is based on IRAS and ISO measurements and is useful when considering our galaxy candidates, for which the flux-densities derived at longer wavelengths are more reliable, given the extreme reddening at shorter wavelengths.

The Dale & Helou (2002) model SEDs assume a power-law distribution of dust mass in a given galaxy. The dust mass, dM_d heated by a radiation field at intensity U is given by,

$$dM_d(U) \propto U^{-\alpha}dU, \qquad (3)$$

where the exponent α represents the relative contributions of the different local SEDs. The range of α that describes normal galaxy SEDs is approximately $1.0 < \alpha < 2.5$. The models were obtained from Dales Web site[6], which provided model SEDs for $0.0625 \leq \alpha \leq 4.0$ (with a step size of 0.0625).

The model SEDs are shown in Figures AA1–AA23 with grey lines. The models showing the two extreme α values are displayed as solid lines, along with a dashed-line showing a mid-range value. This direct comparison between the model and data is used to decide whether or not each candiate is a likely galaxy. Figure 4 shows examples of candidates SEDs that display a good match to the models. Conversely, Figure 5 shows candidates which are unlikely to be galaxies based on their comparsion to the models. Table 4.3 shows the list of candidates that are likely to be galaxies, based on inspection of their SEDs.

4.2   Comparison with large-scale structure from 2MASS

We estimate the distance (and thus redshift) to each candidate galaxy using the Ks-band magnitudes of each, and the distance luminosity function from Kochanek et al. (2001). We derive the Ks-band magnitude using the IRAC channel 1 and 2 flux density measurements. These are less affected by uncertainties, due to extinction and the presence of polycyclic aromatic hydrocarbon (PAH) emission features,
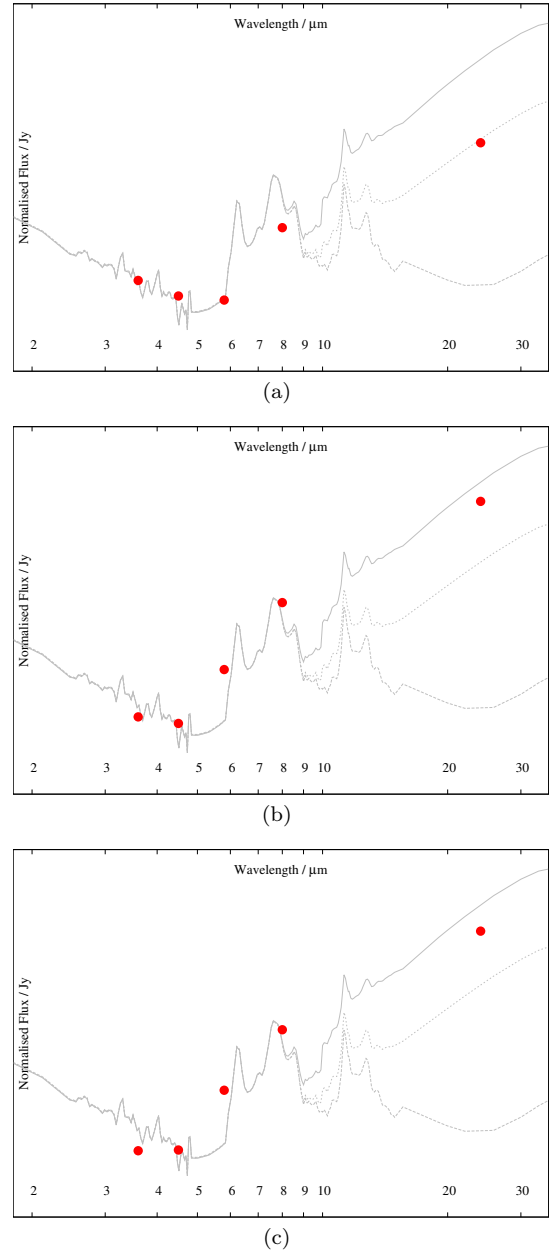
Figure 4. Data with model SEDs for galaxy candidates 4, 14 and 32. These candiates fit well to the model SEDs.

than the IRAC 3 and 4 channels. The resultant, fitted redshift values are shown in Table 4.3. Using this method, the galaxies have estimated redshifts ranging between 0.011 and 0.155. This places them at distances between 47 and 730 Mpc, assuming a cosmology with $H_0 = 71$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.27$ and $\Omega_\Lambda = 0.7$.

To compare the positions of our newly discovered galaxies with the large-scale structure of galaxies in the nearby universe we use the 2MASS Redshift Survey (2MRS, Huchra et. al 2011)[7]. Figures 6–9 show sections of the 2MRS in the region covered by GLIMPSE, with candidate galaxies shown in red.
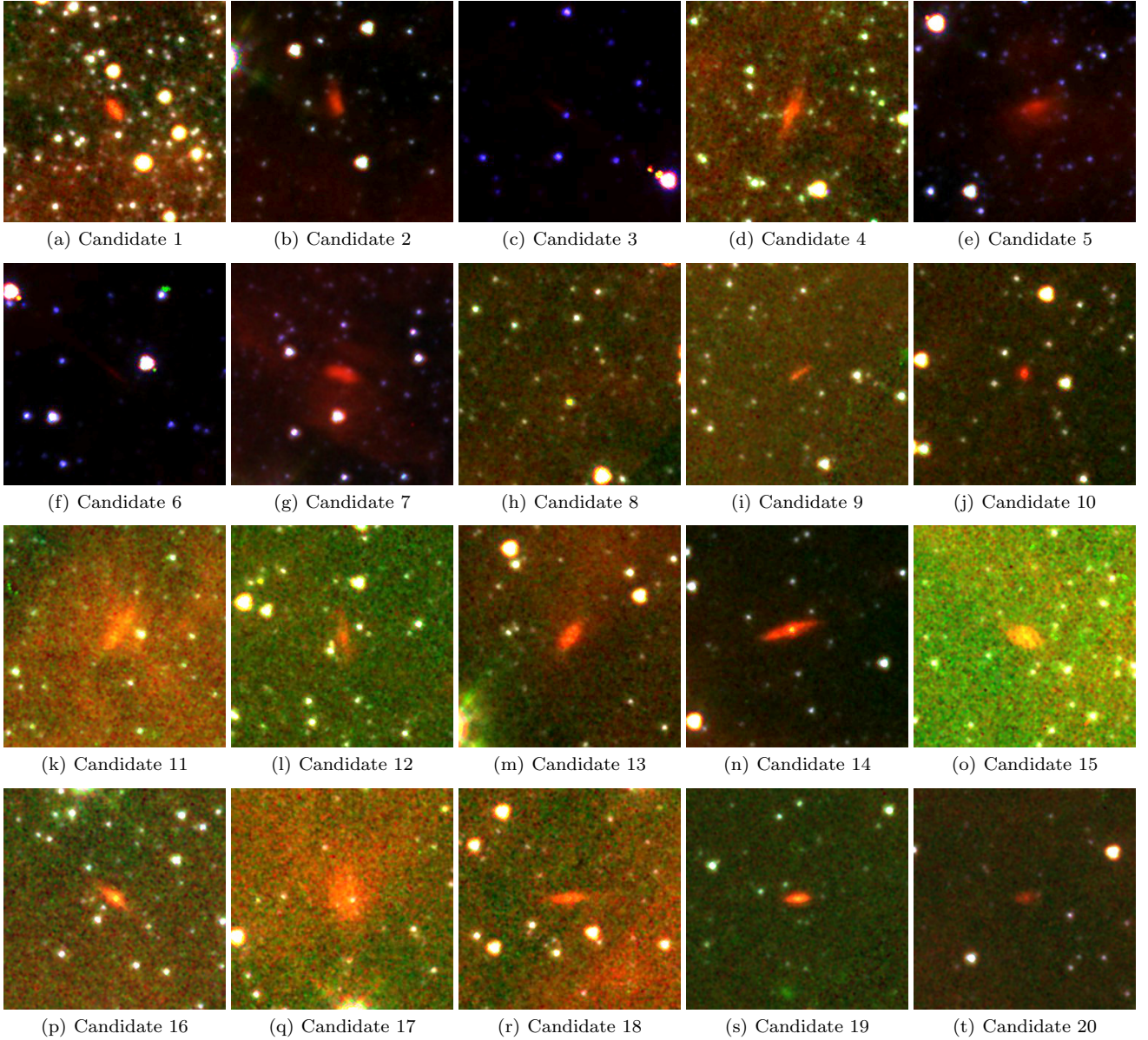
(a) Candidate 1   (b) Candidate 2   (c) Candidate 3   (d) Candidate 4   (e) Candidate 5

(f) Candidate 6   (g) Candidate 7   (h) Candidate 8   (i) Candidate 9   (j) Candidate 10

(k) Candidate 11   (l) Candidate 12   (m) Candidate 13   (n) Candidate 14   (o) Candidate 15

(p) Candidate 16   (q) Candidate 17   (r) Candidate 18   (s) Candidate 19   (t) Candidate 20

Figure 2. Three-colour composite images of candidate galaxies 1–20, with RGB channels using Spitzer IRAC bands 8μm, 4.5μm and 3.6μm respectively.

4.3 Galaxy catalogue

Our list of 35 previously-unknown galaxies in the ZoA is shown in Table 4.3. These objects are morphologically consistent with reddened, extended sources, when seen in the original MWP images and in the optimised IRAC mosaics. They were selected from a large sample created by multiple, independent observations by online volunteers. Measurements of their flux densities from Spitzer at 3.6, 4.5, 5.8, 8 and 24μm show SEDs consistent with model galaxy SEDs from Dale (2001).

5 DISCUSSION

6 SUMMARY AND FUTURE WORK

(a) Candidate 21   (b) Candidate 22   (c) Candidate 23   (d) Candidate 24   (e) Candidate 25

(f) Candidate 26   (g) Candidate 27   (h) Candidate 28   (i) Candidate 29   (j) Candidate 30

(k) Candidate 31   (l) Candidate 32   (m) Candidate 33   (n) Candidate 34   (o) Candidate 35

(p) Candidate 36   (q) Candidate 37   (r) Candidate 38   (s) Candidate 39   (t) Candidate 40

(u) Candidate 41   (v) Candidate 42   (w) Candidate 43
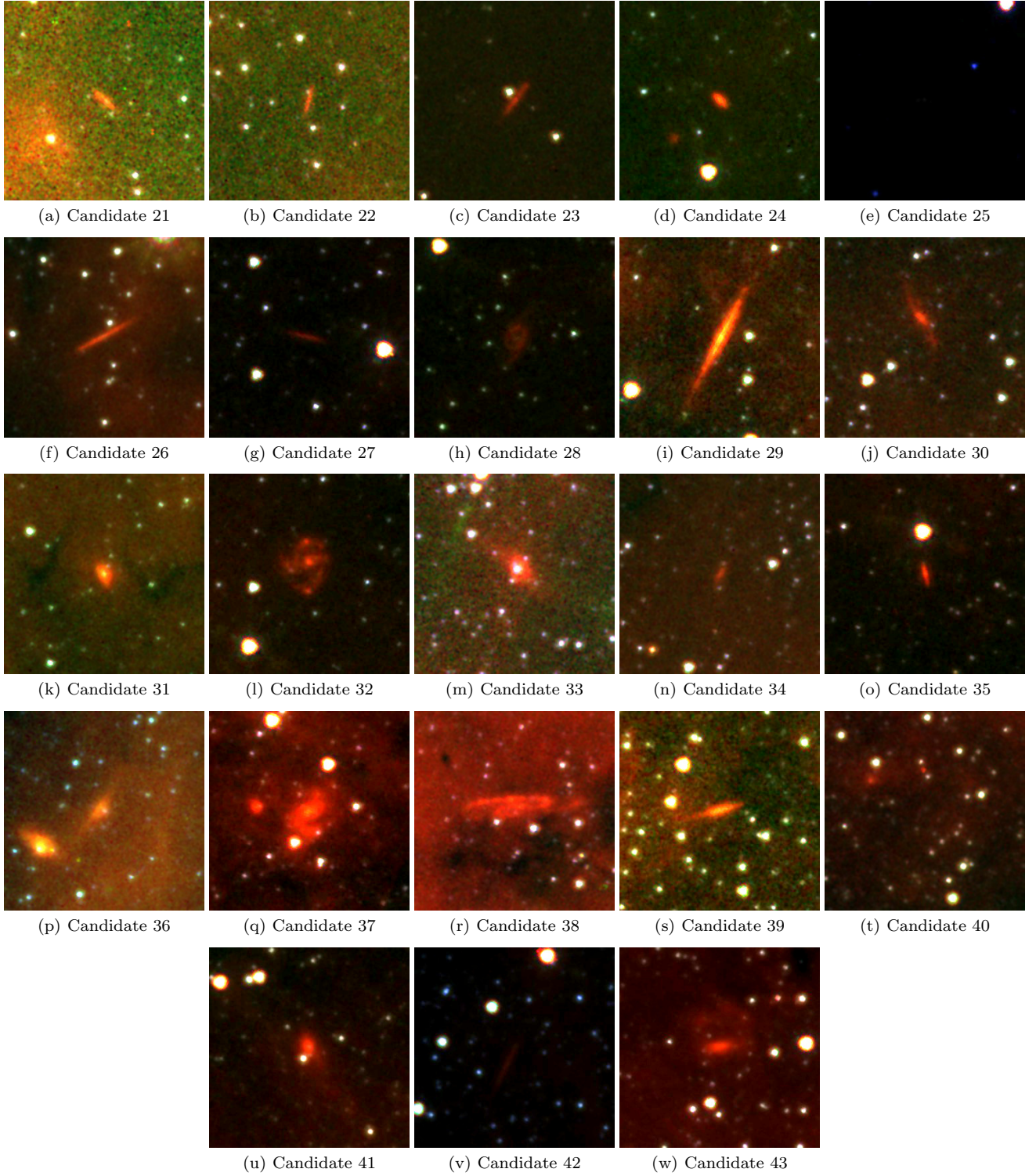
Figure 3. Three-colour composite images of candidate galaxies 21–43, with RGB channels using Spitzer IRAC bands 8μm, 4.5μm and 3.6μm respectively.

REFERENCES

Cardamone, C., et al. 2009, MNRAS, 399, 1191
Jarrett, T. H., et al. 2007, AJ, 133, 979
Joy, K., et al. 2011, Astronomy and Geophysics, 52, 020000
Lintott, C., et al. 2011, MNRAS, 410, 166
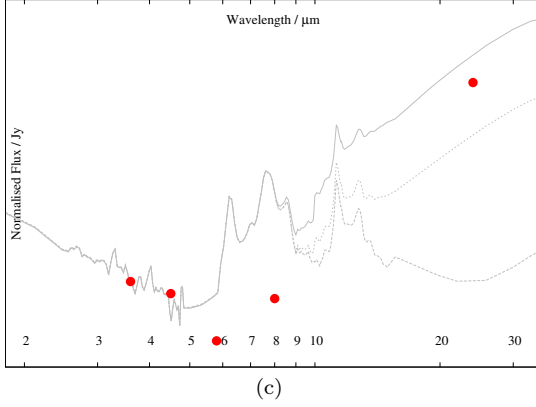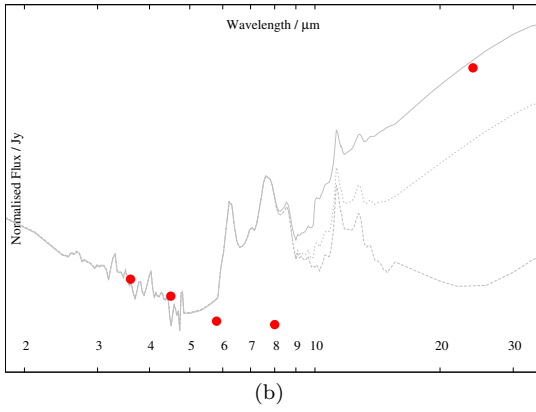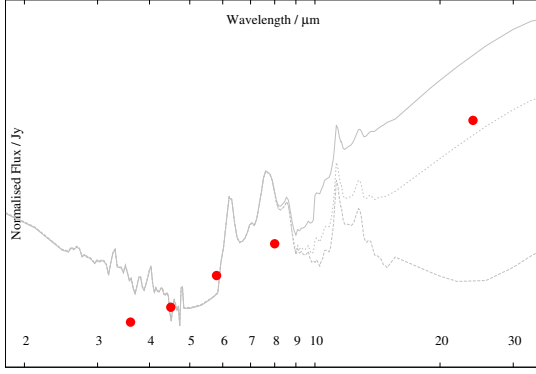
Figure 5. Data with model SEDs for galaxy candidates 27, 33 and 40. These candiates do not fit well to the model SEDs.

Figure 7. Positions of galaxies relative to our own.
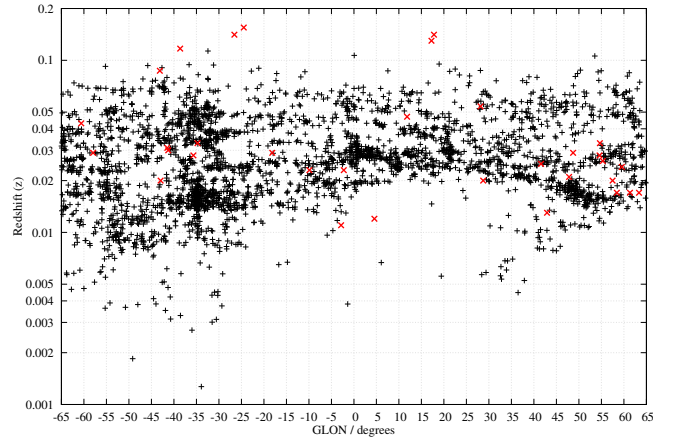
Figure 8. Positions of galaxies relative to our own.

Lintott, C. J., et al. 2008, MNRAS, 389, 1179
—. 2009, MNRAS, 399, 129
Marleau, F. R., et al. 2008, AJ, 136, 662
Simpson, R. J., et al. 2012, MNRAS, 424, 2442
Smith, A. M., et al. 2011, MNRAS, 412, 1309

Figure 6. Positions of galaxies relative to our own.
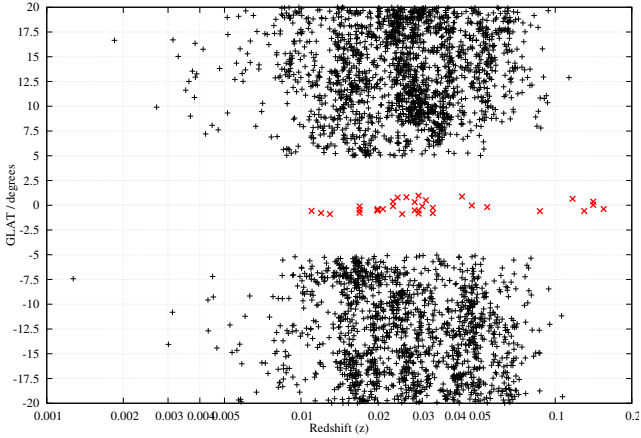
Figure 9. Positions of galaxies relative to our own.

Table 2. Table of galaxy parameters, including galactic coordinates, corrected magnitudes for Spitzer IRAC bands 1, 2, 3 and 4, and the MIPS 24μm flux. In cases where a single IRAC channel does not fit the model well, this is noted in the *outliers* column.

ID	GLON (deg)	GLAT (deg)	Fitted Redshift	Outliers (channel)
1	004.548483	−0.781395	0.012	
2	011.818667	−0.025800	0.047	
4	017.225237	−0.584033	0.130	
5	017.784417	+0.034222	0.141	
6	028.141688	−0.192917	0.054	
7	028.729475	−0.539626	0.020	
9	041.581565	−0.884437	0.025	
10	042.908694	−0.888909	0.013	
11	047.824267	−0.390300	0.021	
12	048.707895	+0.961547	0.029	
13	054.574569	−0.511079	0.028	
14	054.657243	−0.256119	0.033	
15	055.307833	+0.810834	0.026	
16	057.474400	−0.387300	0.020	
17	058.441232	−0.470409	0.017	
18	059.217833	+0.396467	0.004	3
19	059.545904	+0.776002	0.024	
20	061.350500	−0.108500	0.017	
21	063.349203	−0.768282	0.017	
23	299.398377	+0.888066	0.043	
24	302.037426	−0.528374	0.029	
26	306.681500	+0.015125	0.310	4
28	316.873061	−0.599138	0.087	
29	317.039057	−0.497820	0.020	
30	318.527084	+0.504229	0.031	
31	318.729619	−0.097762	0.030	
32	321.402694	+0.652191	0.117	
34	324.323250	+0.317000	0.028	
35	325.245517	−0.799997	0.033	
36	333.445875	+0.356375	0.141	
38	335.517000	−0.378042	0.155	
39	341.857467	−0.842779	0.029	
41	350.101407	+0.389093	0.023	
42	357.148741	−0.572130	0.011	
43	357.737833	−0.092033	0.023	

Table 3. Table of candidates with SEDs that do not look like galaxies.

ID	GLON (deg)	GLAT (deg)	Candidate Status
3	012.741267	−0.000400	
8	036.399831	−0.466475	
22	299.362072	+0.731691	
25	302.373292	−0.023500	
27	314.143892	−0.682605	
33	322.552422	+0.815422	
37	335.104583	−0.066792	
40	350.457146	−0.041708	

Galaxy Images and SEDs

APPENDIX A: CANDIDATE IMAGES AND SEDS

This appendix shows images and SEDs for each of the galaxy candidates in this study. In each case subfigures show (a) MWP image cutout; (b) IRAC 8μm data; (c) Three-colour composite with RGB channels using Spitzer IRAC bands 8μm, 4.5μm and 3.6μm respectively; and (d) measured SED for this object.
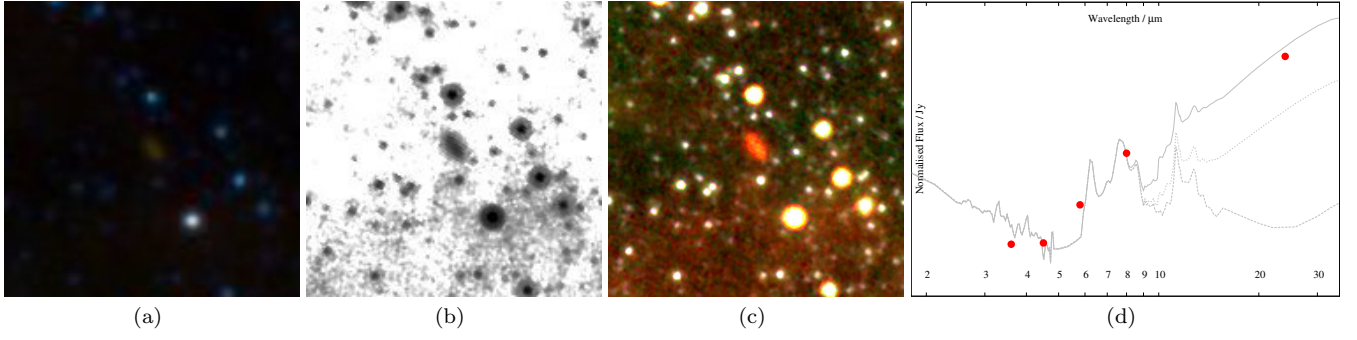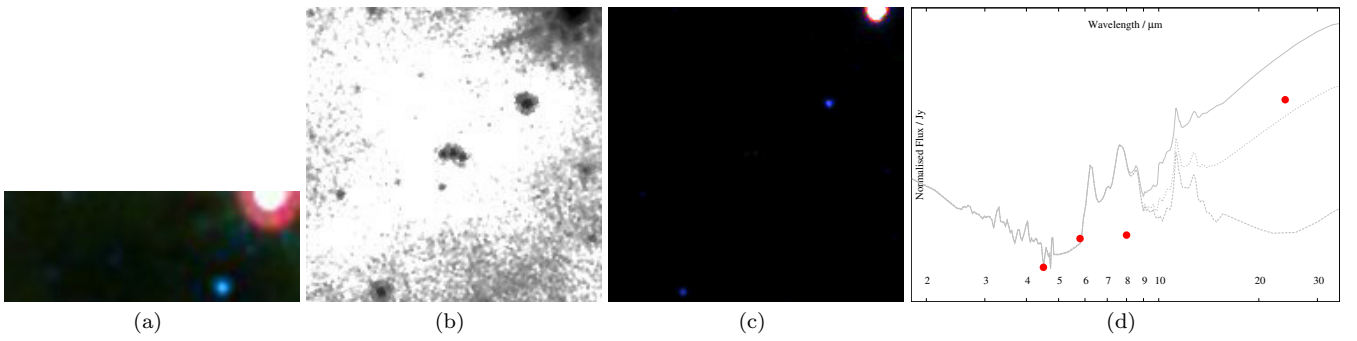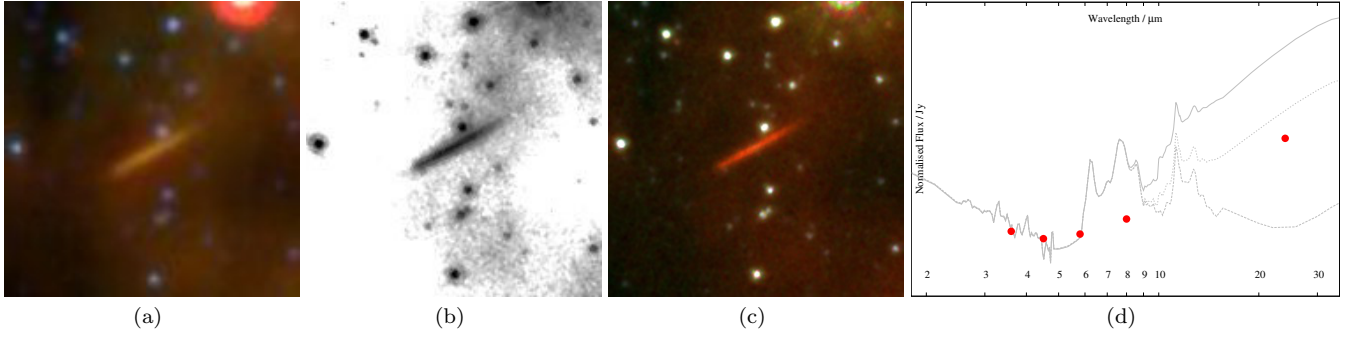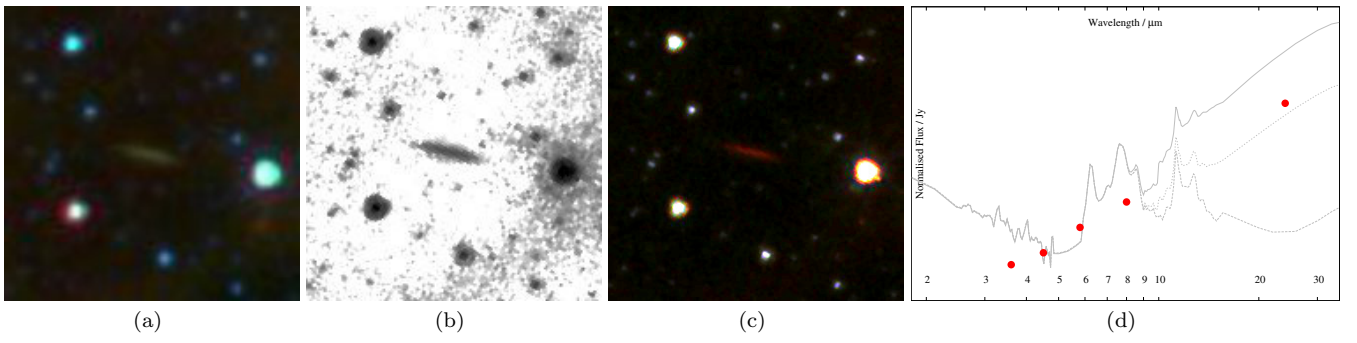
Figure A1. Data for galaxy candidate 1.
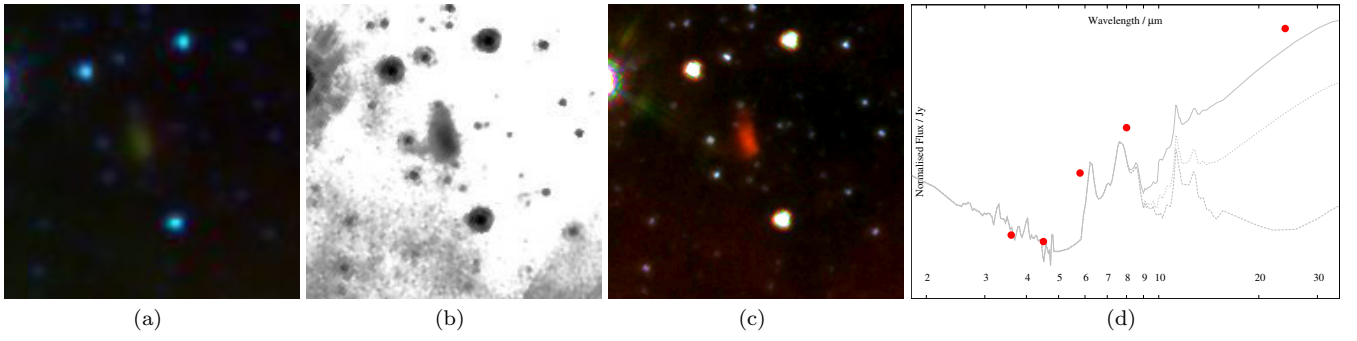
Figure A2. Data for galaxy candidate 2.
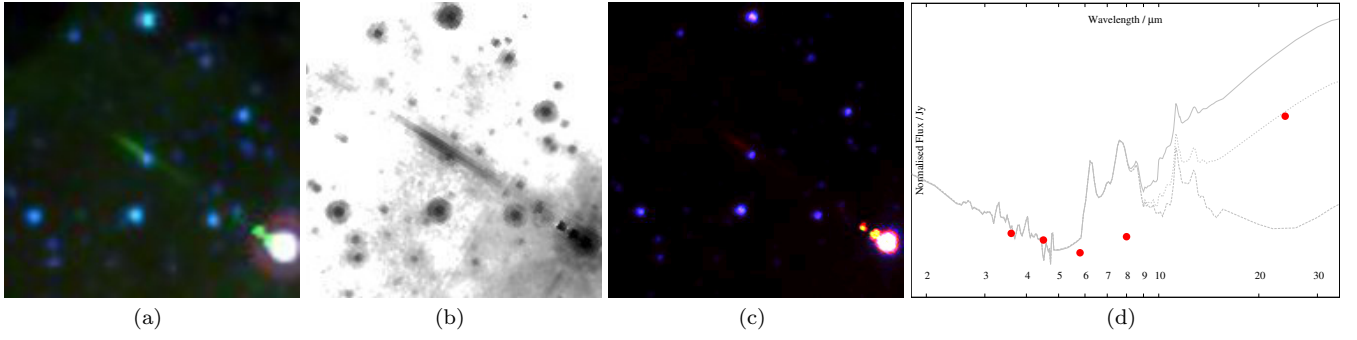
Figure A3. Data for galaxy candidate 3.
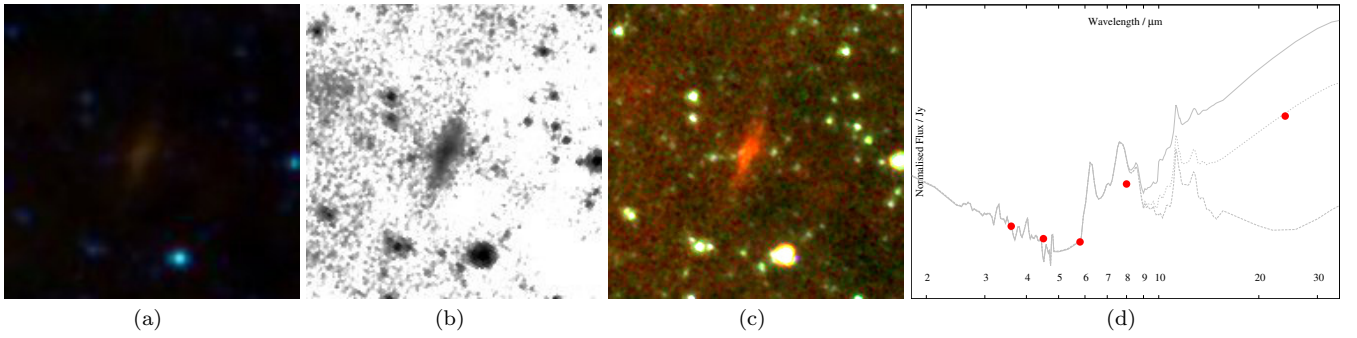
Figure A4. Data for galaxy candidate 4.
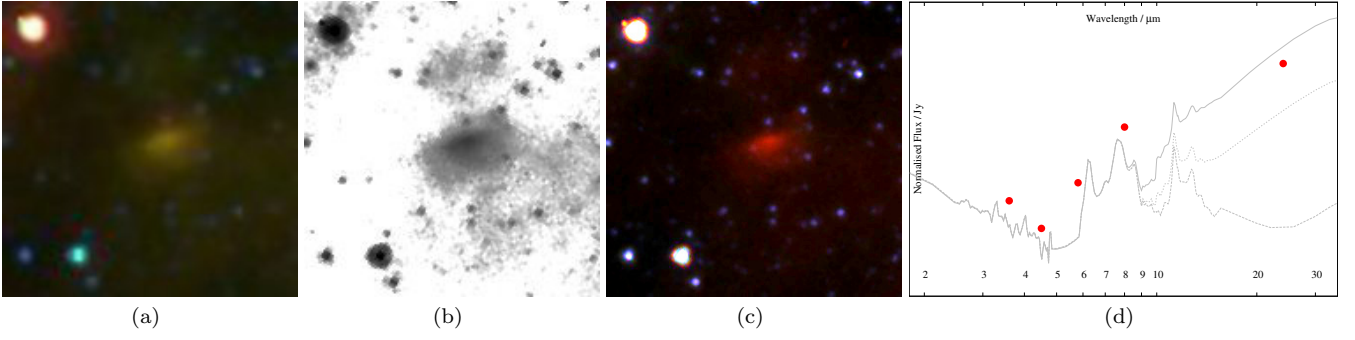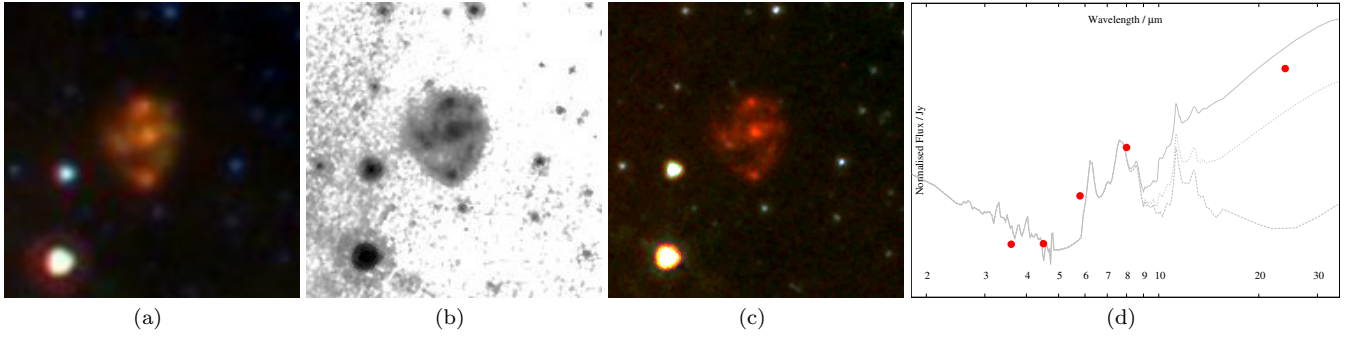
Figure A5. Data for galaxy candidate 5.
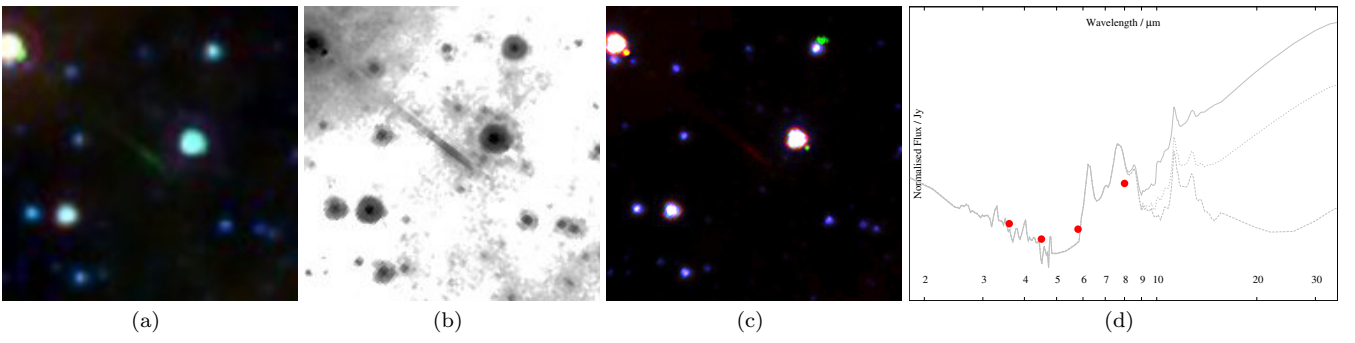
Figure A6. Data for galaxy candidate 6.
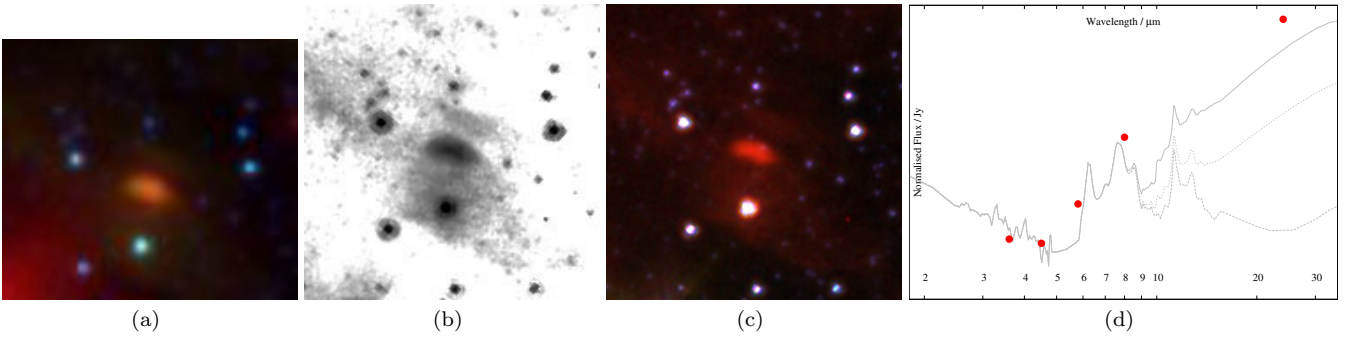
Figure A7. Data for galaxy candidate 7.
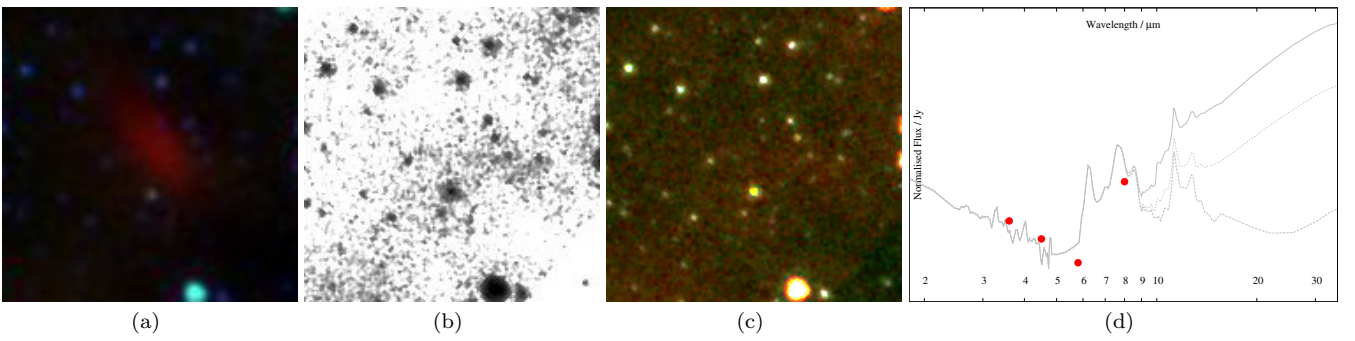
Figure A8. Data for galaxy candidate 8.
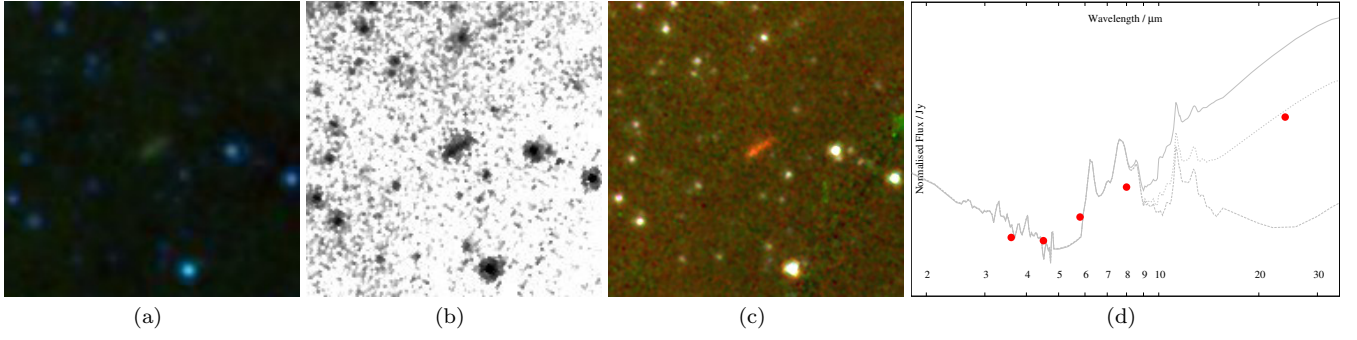
Figure A9. Data for galaxy candidate 9.
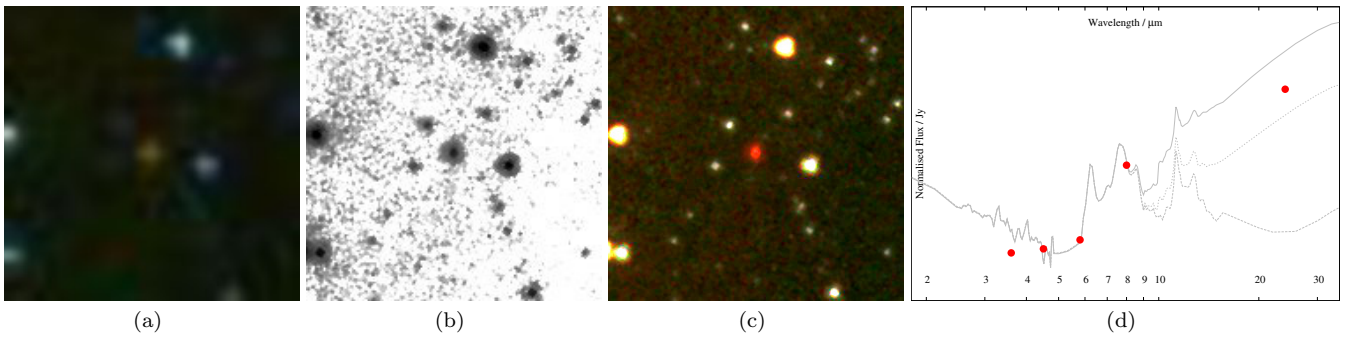
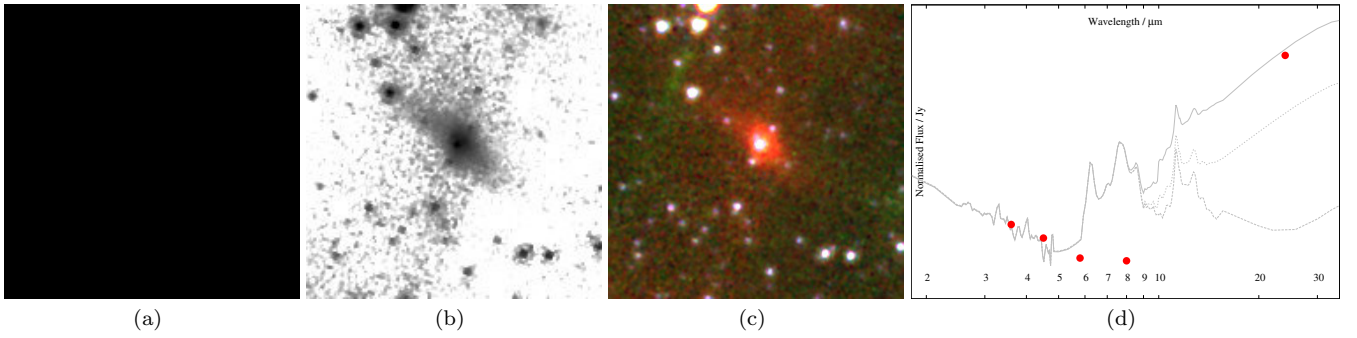

Figure A10. Data for galaxy candidate 10.
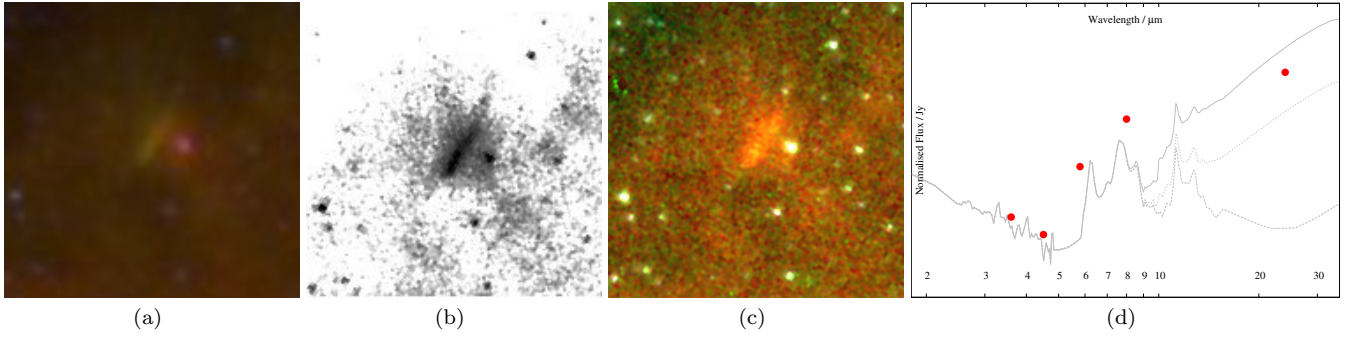
Figure A11. Data for galaxy candidate 11.
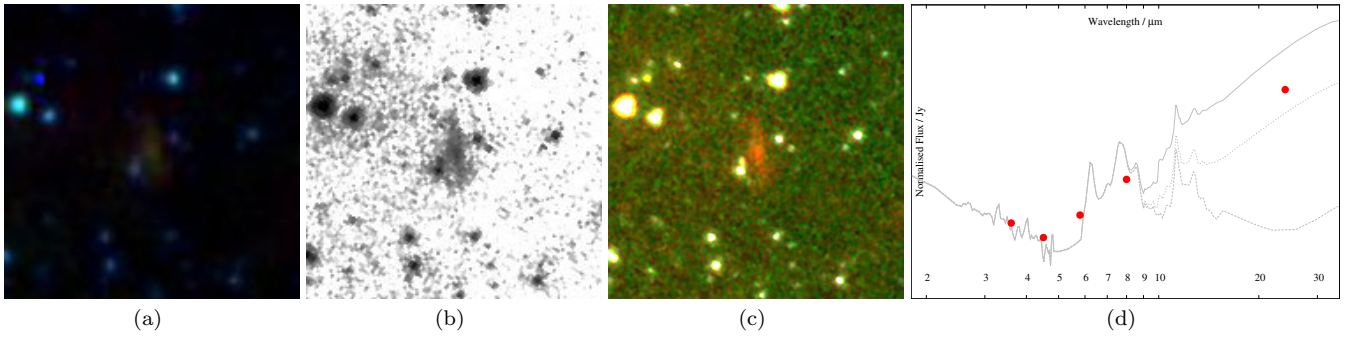
Figure A12. Data for galaxy candidate 12.
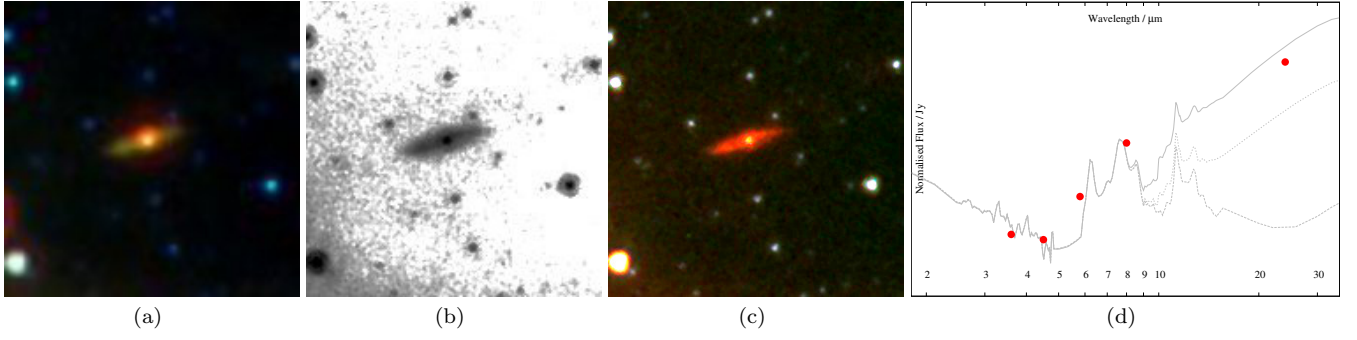
Figure A13. Data for galaxy candidate 14.
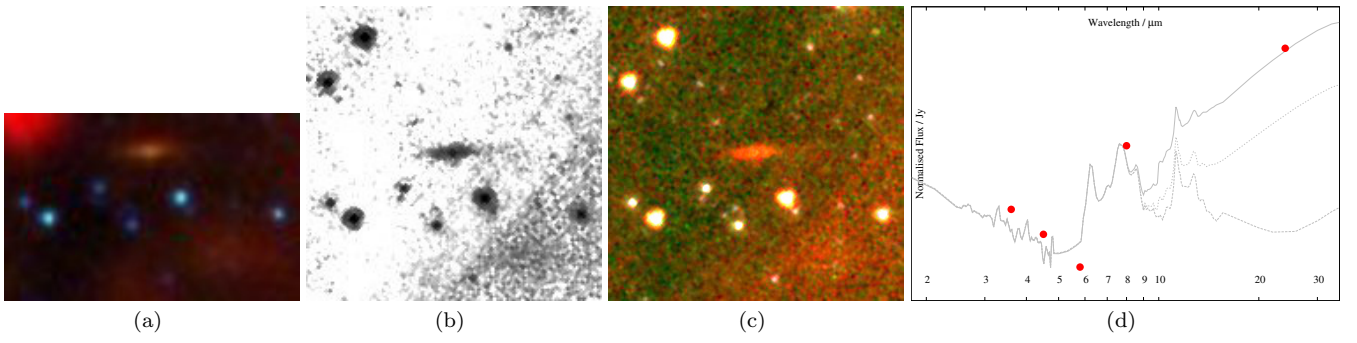
Figure A14. Data for galaxy candidate 18.
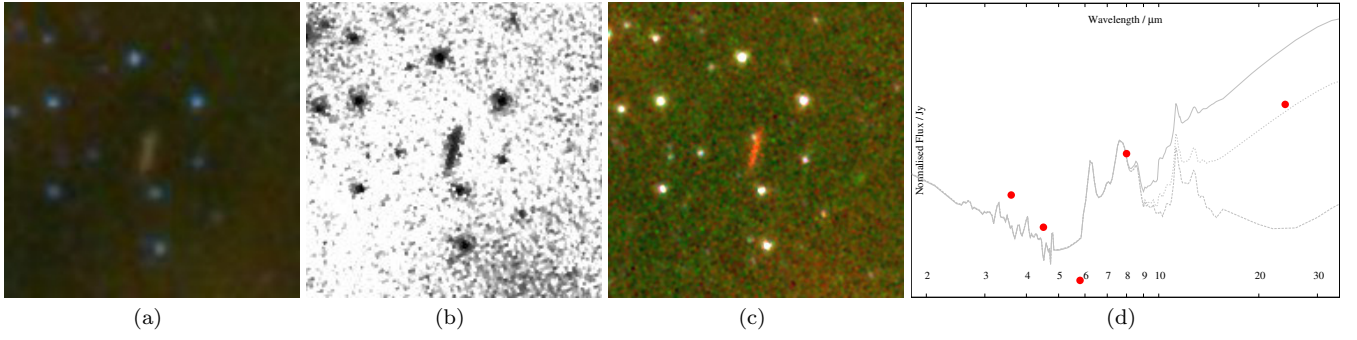
Figure A15. Data for galaxy candidate 22.

Figure A16. Data for galaxy candidate 25.

Figure A17. Data for galaxy candidate 26.

Figure A18. Data for galaxy candidate 27.

Figure A19. Data for galaxy candidate 32.
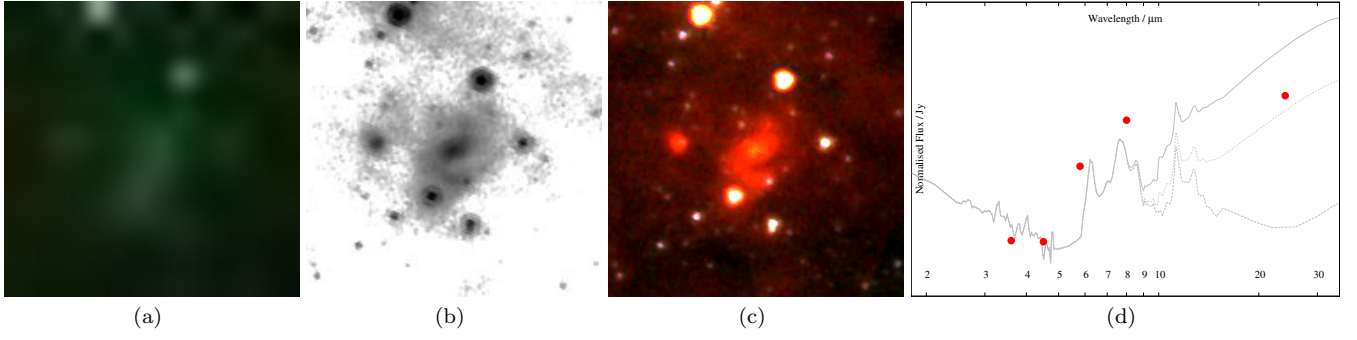

Figure A20. Data for galaxy candidate 33.

Figure A21. Data for galaxy candidate 37.
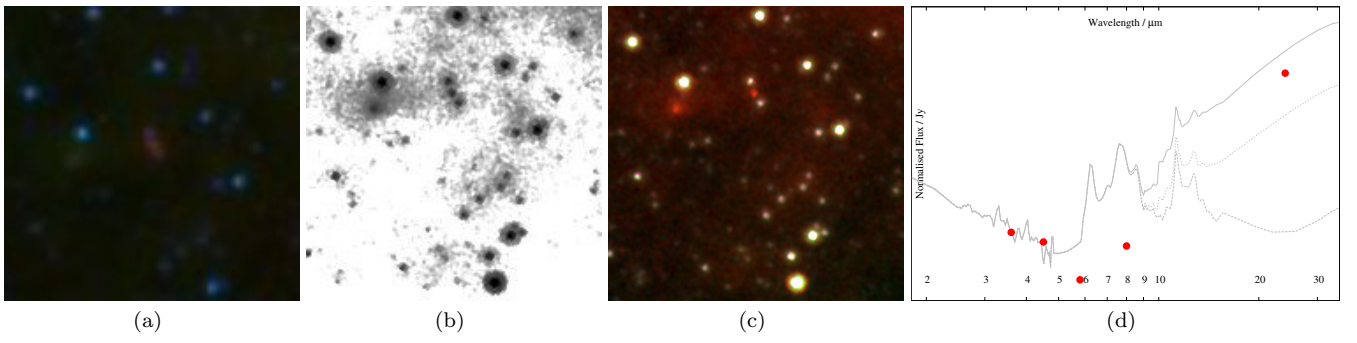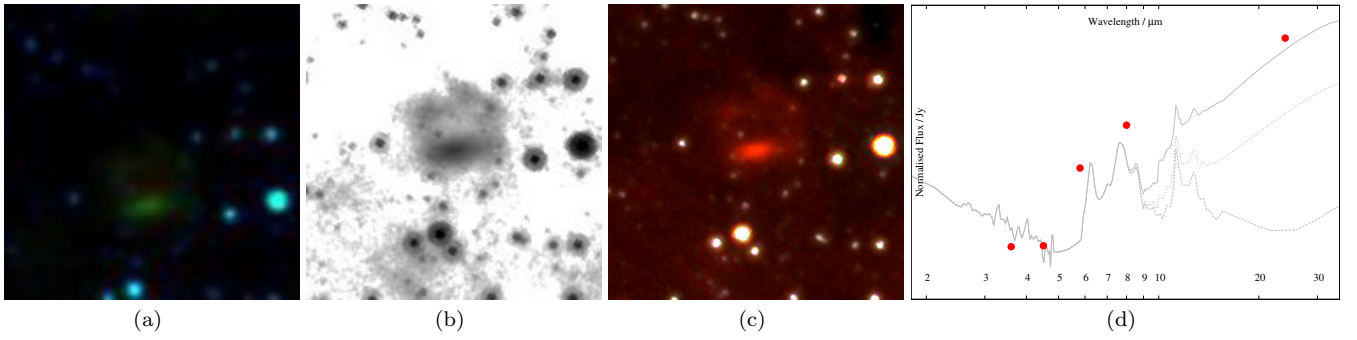

Figure A22. Data for galaxy candidate 40.

Figure A23. Data for galaxy candidate 43.