Mon. Not. R. Astron. Soc. **369,** 969–975 (2006) doi:10.1111/j.1365-2966.2006.10353.x

The build-up of the red sequence in the galaxy cluster MS1054−0321 at $z = 0.831$

S. Andreon★
INAF – Osservatorio Astronomico di Brera, Milano, Italy

Accepted 2006 March 20. Received 2006 March 17; in original form 2005 December 27

ABSTRACT

Using one of the deepest data sets available, we determine that the red sequence of the massive cluster MS1054−0321 at $z = 0.831$ is well populated at *all* studied magnitudes, showing no deficit of faint (down to $M^* + 3.5$) red galaxies: the faint end of the colour–magnitude relation is neither empty nor underpopulated. The effect is quantified by the computation of the luminosity function (LF) of red galaxies. We found a flat slope, showing that the abundance of red galaxies is similar at faint and at intermediate magnitudes. Comparison with present-day and $z \sim 0.4$ LFs suggests that the slope of the LF is not changed, within the errors, between $z = 0.831$ and 0. Therefore, the analysis of the LF shows no evidence for a decreasing (with magnitude or redshift) number of faint red galaxies. The presence of faint red galaxies in high-redshift clusters disfavours scenarios where the evolution of red galaxies is mass-dependent, because the mass dependency should differentially depopulate the red sequence, while the MS1054−0321 colour–magnitude relation is populated as in nearby clusters and as in $z \sim 0.4$ clusters. The presence of abundant faint red galaxies in the high-redshift cluster MS1054−0321 restricts the room for allocating descendants of Butcher–Oemler galaxies, because they should change the faint end slope of the LF of red galaxies, while instead the same faint end slopes are observed in MS1054−0321, at $z \sim 0$ and at $z \sim 0.4$. In the rich MS1054−0321 cluster, the colour–magnitude relation seems to be fully in place at $z = 0.831$ and therefore red galaxies of all magnitudes were wholly assembled at higher redshift.

Key words: methods: statistical – galaxies: clusters: individual: MS1054–0321 – galaxies: evolution – galaxies: formation – galaxies: fundamental parameters – galaxies: luminosity function, mass function.

1 INTRODUCTION

Intensive studies of galaxy properties, such as star formation rates and morphology, have significantly improved our understanding of galaxies in the Universe. However, it is still unclear how galaxies evolve over Hubble time, largely because of the heterogeneous nature of galaxy properties: galaxies differ in size, mass, star formation histories, etc. Galaxies evolve with time at least as a natural consequence of stellar evolution.

Recently conducted large surveys, such as the 2dF Galaxy Redshift Survey (Colless et al. 2001) and the Sloan Digital Sky Survey (SDSS; York et al. 2000), revealed that galaxy properties show strong bimodality in their distribution (e.g. Strateva et al. 2001). That is, there are two distinct populations: red early-type galaxies and blue late-type galaxies. This bimodality is found to be a strong function of the mass of galaxies in the sense that massive galaxies tend to be red early-type galaxies, while less massive galaxies tend to be blue late-type galaxies.

The faint end of the red population has been suggested to be a newcomer population, because the colour–magnitude diagrams of ESO Distant Clusters Survey (EDis CS; De Lucia et al. 2004) at $z \sim 0.8$ 'show a deficiency of red, relatively faint galaxies and suggest that such a deficit may be a universal phenomenon in clusters at these redshifts'. A similar deficit has been hinted also at higher redshift ($z \sim 1.2$, Kajisawa et al. 2000; Nakata et al. 2001), although the authors cautioned about the possibility that their deficit is spurious. Similarly, Kodama et al. (2004) notice an apparent absence of galaxies on the red colour-magnitude sequence at $M^* + 2$. Goto et al. (2005) suggest a similar deficit for the cluster MS1054−0321 at $z = 0.831$. These faint red galaxies are instead observed in present-day clusters (e.g. Secker, Harris & Plummer 1997). The absence of faint red galaxies at high redshift implies that a large fraction of present-day passive faint galaxies must have moved on to the colour–magnitude relation at redshift below $z \sim 0.8$ and therefore their star formation activity must have ended at larger redshift (De Lucia et al. 2004).

In this paper, making use of the uniform and very deep photometry collected by the Faint Infrared Extragalactic (FIRES) team (Labbé et al. 2003; Förster Schreiber et al. 2005), we test whether

★E-mail: andreon@brera.mi.astro.it

Table 1. The cluster and control field sample.

	MS1054	HDF-S[a]
$T_{\rm exp}$ V (ks)	6.5	97
$T_{\rm exp}$ I (ks)	6.5	112
$T_{\rm exp}$ K (ks)	21-28	128
Completeness K (mag)	21.5[b]	22.5
Area (arcmin^2)	20.7	5.8

[a]*Hubble Deep Field*–South.
[b]Estimated.

there is a deficit of faint red galaxies in a high-redshift cluster: MS1054−0321 (MS1054 for short). MS1054 has been detected in the Extended Medium Sensitivity Survey (Gioia et al. 1990) and has a redshift of $z = 0.831$. It is quite bright in X-ray, having $L_{\rm X}$(2–10 keV) $= 2.2 \times 10^{45}\ h_{50}^{-1}$ erg s^{-1} (Donahue et al. 1998). MS1054 has an Abell richness class of 3 and a line-of-sight velocity dispersion of $\sigma_v = 1170 \pm 150$ km s^{-1} (Tran et al. 1999). The weak-lensing mass is estimated to be $9.9 \pm 0.4 \times 10^{14}$ M$_\odot$ (Jee et al. 2005).

Throughout this paper, we assume $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$ km s^{-1} Mpc^{-1}. All magnitudes are in the Vega system.

2 THE DATA

The cluster and control field images used in this analysis have been taken in the K_s band with the Infrared Spectrometer And Array Camera (ISAAC) on the European Southern Observatory (ESO) Very Large Telescope (VLT) and in the V_{606} and I_{814} bands with the Wide Field Planetary Camera 2 on the *Hubble Space Telescope*. From now on, MS1054 and control field photometry in V_{606}, I_{814} and K_s is denoted with V, I and K for the sake of clarity. Images on MS1054 cover a 4.3-Mpc^2 area.

We adopt here the photometry provided by the FIRES team and described in full detail in Labbé et al. (2003) and Förster Schreiber et al. (2005). Key quantities are listed in Table 1. Specifically, we use their 'total' fluxes and 'optimal' colours.

Completeness limits (for extended sources) have been estimated according to the prescription of Garilli, Maccagni & Andreon (1999) and Andreon et al. (2000), by considering the magnitude of the brightest galaxies displaying the lowest detected central surface brightness values. Only galaxies brighter than this completeness limit are considered in this paper.

3 RESULTS

3.1 Qualitative evidence for no deficit of red galaxies

Fig. 1 shows the $V - I$ versus K diagram of galaxies in the cluster line of sight (top panel) and in the reference line of sight (bottom panel). At the cluster redshift, $V - I$ approximately samples the rest-frame $U - B$ colour, which straddles the 4000 Balmer break and is therefore very sensitive to any recent or ongoing star formation. Instead, the K band samples the rest-frame J band, which is far less sensitive to short episodes of star formation than the optical U and B bands. The cluster line of sight is incredibly rich with galaxies having $V - I \sim 2$ mag, which are almost absent in the control field line of sight. A red sequence is clearly visible and it extends down to the completeness magnitude (marked with a vertical dashed line). The large circle shows the colour and luminosity of a 4×10^{11} M$_\odot$ single stellar population model, based on Bruzual & Charlot (2003).

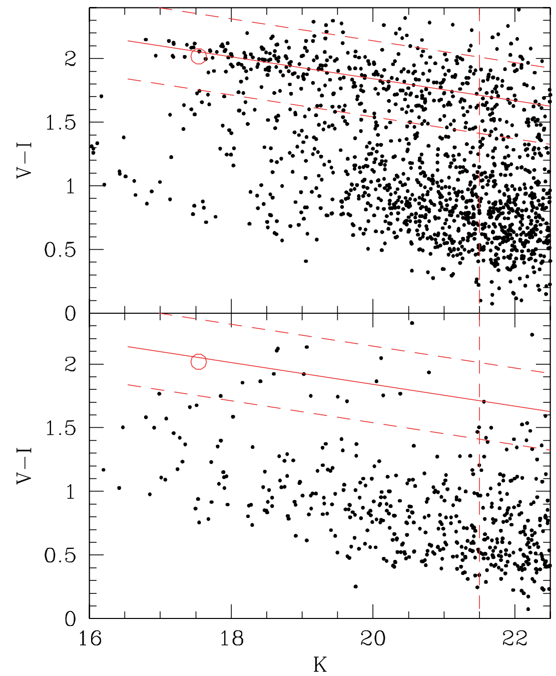

Figure 1. Colour–magnitude relation in the cluster (top panel) and control field (bottom panel) line-of-sight directions. The solid line marks the colour–magnitude relation, whereas the dotted lines delimit the portion of plane that qualifies galaxies to be called red. The vertical line marks the sample completeness.

A Salpeter initial mass function is used, with the lower/upper limit to the mass range fixed at 0.1/100 M$_\odot$. We assume a formation redshift $z_{\rm f} = 10$, solar metallicity and Padua 1994 tracks. The slope, intercept and scatter of the colour–magnitude relation are computed in Appendix A, following laws of probabilities. Fig. 2 shows the colour–magnitude diagram in the cluster line of sight, using I as abscissa, disregarding near-infrared photometry.

Perhaps the most interesting result of our analysis is that the red sequence in MS1054 is well populated at *all* magnitudes, showing no deficit of faint ($K > 20$ or $I > 23.5$) galaxies, a result which is quantified in the following sections based on colour-dependent luminosity and stellar mass functions. Qualitatively, our Figs 1 and 2 should be contrasted with similar plots in Kodama et al. (2004) and De Lucia et al. (2004), which both show an almost empty (or under-populated) region of the colour–magnitude diagram at the magnitudes and colour of faint red galaxies, i.e. a large deficit of faint red

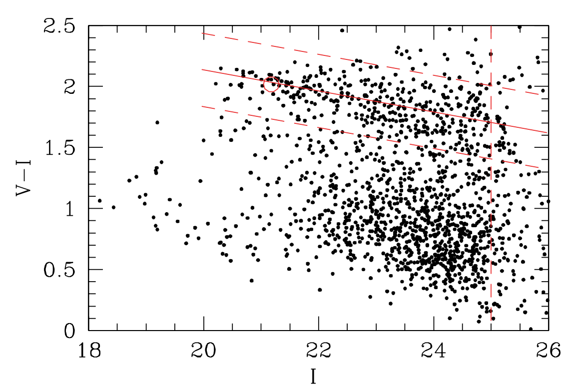

Figure 2. Colour–magnitude relation in the cluster line-of-sight direction. The meaning of lines is as in Fig. 1.

galaxies in high-redshift clusters. We anticipate that, in the cluster rest frame, our data are deeper than theirs.

3.2 Quantitative evidence for no deficit

In order to quantify the abundance of faint red galaxies, we compute the luminosity function (LF) of red galaxies. We define galaxies as red if their colour is within 0.3 mag from the colour–magnitude relation (within 3 times the measured dispersion around the colour–magnitude relation in the K band). We compute the LF by fitting the unbinned colour-selected galaxy counts in the cluster and control field direction. Here, we follow the rigorous method set forth in Andreon, Punzi & Grado (2005, APG hereafter), which is an extension of the Sandage, Tammann & Yahil (1979, STY) method to the case where a background is present, and that adopts the extended likelihood instead of the conditional likelihood used by STY. For display purposes only, the LF is also computed as the difference between the observed (binned) counts in the cluster direction and the best fit to the counts in the control field direction. Errors on data points, again for display purpose only, are simply given by the square root of the variance of the observed counts in the cluster line of sight. These binned data and these errors are not used in the mentioned rigorous method, which instead computes 68 per cent confidence intervals using the Likelihood Ratio theorem (Wilks 1938, 1963).

As a model for the cluster LF, we use a Schechter (1976) function:

$$\phi(m) = \phi^* 10^{0.4(\alpha+1)(m^*-m)} \exp\left(-10^{0.4(m^*-m)}\right), \qquad (1)$$

where m^*, α and ϕ_i^* are the characteristic magnitude, slope and normalization, respectively.

There are 313 red galaxies in the cluster line of sight, ≈80 per cent of which are cluster members. As is also obvious from the bottom panel of Fig. 1, the contamination of galaxy counts by the background at the considered colour is low and therefore the LF parameters are almost unaffected by the uncertainty in the background counts.

We found $K^* = 17.99 \pm 0.27$ mag (i.e. $M_J^* = -24.0$ mag) and $\alpha = -0.95 \pm 0.15$. The resulting LF is shown in Fig. 3. Fitting I-band counts and without any use of the near-infrared photometry, we found $I^* = 21.53 \pm 0.21$ mag (i.e. $M_B^* = -21.0$ mag) and $\alpha = -0.80 \pm 0.12$. The resulting LF is shown in Fig. 4. In I, the data turn out to be complete down to $I = 25.0$ mag.

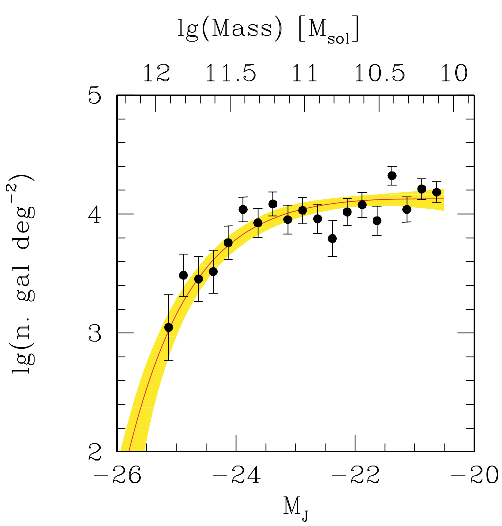

Figure 3. J-band luminosity and mass function of red galaxies in MS1054. The curve shows the rigorous LF determination, whereas the dots and error bars mark approximate values derived by standard (e.g. Zwicky 1957; Oemler 1974) background subtraction (see text for details). The shaded area marks the model uncertainty.

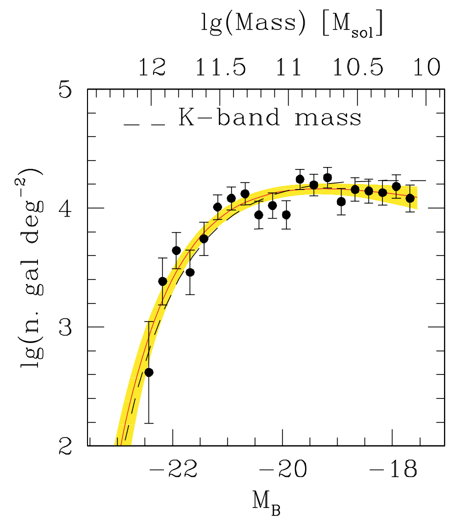

Figure 4. B-band luminosity and mass function of red galaxies in MS1054. Solid curve, dots, error bars and shading as in Fig. 3. The dashed line is the near-infrared band derived mass function shown in Fig. 3.

The slope α of the LF is flat in both the B and J bands (see also Figs 3 and 4). There is no evidence for a decrease in the number of red galaxies, down to the magnitude of completeness. Therefore, there is no deficit of faint red galaxies: the number of red galaxies is similar at faint and intermediate magnitudes, down to the magnitude of completeness. We have here quantified our qualitative statement about the abundance of faint red galaxies in MS1054.

Our determination of the slope α is robust to minor changes of the ‘red’ definition: first of all, we found that blue galaxies represent a minority population in the studied region of MS1054, therefore the precise location of the blue side of the red region can be moved without affecting our results. Furthermore, the colour distribution of red galaxies is bell-shaped, which implies that moving the colour boundaries by, say, 0.1 mag redward or blueward removes from the sample a minority population (2 per cent, for a Gaussian with $\sigma = 0.1$ mag). The robustness of our definition of ‘red’ simplifies the comparison with similar literature determinations, when the adopted definitions of ‘red’ are designed to encompass the whole red population, as our definition does. Finally, by adopting the colour boundary far from the central colour of the red population (3σ in colour), we strongly limit Malmquist-like biases that would otherwise systematically deplete the LF at fainter magnitudes.

We now look for an evolution of the number of faint red galaxies between high and low redshift. If faint red galaxies are newcomers, then the slope of the LF should change between low and high redshift and, specifically, the LF should be flatter (less negative) at higher redshift than at low redshift. Tanaka et al. (2005) find $\alpha = -0.9 \pm 0.1$ (as revised in Tanaka et al. 2006) in the V band for a composite sample of clusters at $z = 0$, in good agreement with our values at $z = 0.831$, $\alpha = -0.95 \pm 0.15$ in the J band and $\alpha = -0.80 \pm 0.12$ in the B band. Barrientos & Lilly (2003) define red the galaxies having $\approx U - V$ within 0.25 mag from the red sequence and found $\alpha = -0.80 \pm 0.03$ in the $\approx V$-band rest frame for a composite sample of eight clusters at $z \sim 0.4$. As mentioned, the slope does not seem to be band-dependent between B and J, therefore our comparison with the value observed by Tanaka et al. (2005) and Barrientos &

Lilly (2003) in the V band should be appropriate. Such a comparison shows that, within the error, the slope has not changed from $z = 0$ and 0.831, in good agreement with the Tanaka et al. (2005) comparison of their $z = 0$ clusters to another $z \sim 0.83$ cluster, blessed however by larger uncertainties.

To summarize, the analysis of the LF shows no evidence for a decreasing (with magnitude or redshift) number of faint red galaxies, i.e. we do not find in MS1054 the suggested deficit of faint red galaxies at high redshift observed in other clusters.

3.3 Stellar mass function of red galaxies

Absolute magnitudes can be converted into a stellar mass scale, using the $\mathcal{M}/L$ ratio of our passive evolving model. The mass scale is shown as the upper abscissa in Figs 3 and 4. We derive a characteristic mass of 2.9×10^{11} $M_{\odot}$ from both the J- and B-band (rest-frame) photometry. The statistical accuracy is 30 per cent, derived from the m^* uncertainty only. The absolute value of the characteristic mass, $\mathcal{M}^*$, depends on several key model parameters (e.g. the lower mass limit of the initial mass function, see Bell & de Jong 2001) and caution should be exercised in the comparison of masses derived using different models.

In Fig. 4, the dashed line reproduces the mass function derived from near-infrared photometry shown in Fig. 3. The agreement is remarkable, especially considering that there are no free parameters in such a comparison, and clarifies that red galaxies are sufficiently simple objects that their mass can be derived equally well from B or J photometry, in spite of the fact that the near-infrared emission is often claimed to be a better tracer of stellar mass than the optical emission. At the same time, the agreement is unsurprising, because the model is chosen to agree with the observed colours.

4 DISCUSSION

The most significant result of this paper is the lack of a deficit: the faint end of the colour–magnitude relation of MS1054 is neither empty nor underpopulated. Furthermore, the slope of the LF of red galaxies in our cluster at $z = 0.831$ is equal to the value observed in $z \sim 0$ and 0.4 clusters. This claim is at odds with some recent works, claiming a deficit of faint red galaxies. Evidence for a 'truncation' of the red sequence has been noticed in a cluster at $z \sim 1.2$ by Kajisawa et al. (2000) and Nakata et al. (2001), at $z \sim 1.1$ by Kodama et al. (2004) and at $z \sim 0.75$ by De Lucia et al. (2004). Our colour–magnitude determination goes at least as deep as those, reaching $M^* + 3.5$ versus $M^* + 3.1$ (Kodama et al. 2004) and $M^* + 2.4$ (De Lucia et al. 2004). Therefore, our data are deep enough to detect the effect, if present.

If true, the lack of faint red galaxies has a profound implication for our understanding of galaxy evolution: the deficit of faint red galaxies at high redshift and the presence of these galaxies in nearby clusters has been interpreted (e.g. De Lucia et al. 2004; Kodama et al. 2004) as evidence in favour of an end of star formation activity at $z \sim 0.8$ of many cluster galaxies. Equivalently, a large fraction of present-day passive galaxies must have moved on to the colour magnitude relation at redshifts lower than $z \sim 0.8$, which also provide a counterpart for Butcher–Oemler galaxies (Butcher & Oemler 1984, BO), i.e. galaxies blue at the time of observation (i.e. at high redshift) and almost absent in the present-day Universe. Furthermore, the lack of faint red galaxies rejects a formation scenario in which all red galaxies in clusters evolved passively after a synchronous monolithic collapse at high redshift and suggests a different evolutionary path for present-day passive galaxies. Finally, such a path should depend on luminosity, because faint red galaxies are missing at high redshift, while instead massive galaxies are already abundant in distant clusters.

The presence of faint red galaxies in MS1054 questions and reverses the above interpretation: it disfavours scenarios where the evolution of red galaxies is mass-dependent and restricts the room for allocating descendants of Butcher–Oemler galaxies. In the rich MS1054 cluster, the colour–magnitude relation seems to be fully in place at $z = 0.831$ and therefore red galaxies of all magnitudes wholly assembled at higher redshift.

In view of the centrality of our detection of many faint and red galaxies, previous claims of a deficit should be commented. Kodama et al. (2004) discuss why determinations of a deficit of red galaxies predating their paper should be taken with caution.

De Lucia et al. (2004) use photometric redshifts to discard likely interlopers and found a deficit, whereas we statistically account for interlopers using a control field and we do not found any deficit. While the photometric redshift approach seems superior to our statistical subtraction, we emphasize the intrinsic difficulty of their approach. Tanaka et al. (2005) emphasize that the use of photometric redshift is known to possibly produce biased and colour-dependent estimates of the photometric redshift (Kodama, Bell & Bower 1999; Tanaka et al. 2005). Furthermore, even in the absence of biases, De Lucia et al. use the likelihood of the data as a proxy for probability of the hypothesis (i.e. exchange posterior with likelihood), which is a potentially dangerous practice (see also section 3.3.6 in APG). Unfortunately, we cannot quantify the importance of the involved approximations because quantitative details of the De Lucia et al. scheme for rejecting interlopers has not been published. Nevertheless, our results on MS1054 allow to reject the claimed universality (De Lucia et al. 2004) of the deficit of faint red galaxies in clusters at $z \sim 0.8$.

De Lucia et al. (2004) use the luminous-to-faint ratio to conclude that there is a 3σ difference between the Coma and $z \sim 0.75$ clusters. However, their comparison is blessed by an inaccuracy in the computation of the Coma ratio: panel b of their fig. 3 shows that the number of bright and faint galaxies are similar. In fact, counting them gives a ratio of $\sim 0.7 = \frac{\sim 62}{\sim 93}$ versus a quoted 0.34 ± 0.06. After our revision, the Coma cluster has a luminous-to-faint ratio that differs by less than 1σ from the quoted $z \sim 0.75$ EDiSCs clusters (0.81 ± 0.18). The claimed difference on the luminous-to-faint ratio of Coma and $z \sim 0.75$ EDiSCs clusters is the only quantitative evidence De Lucia et al. present in favour of a build-up of the red sequence. Because, instead, the two ratios are equal, their data favour, if any, the opposite scenario: all (or at least most) of the red galaxies were already in place in the $z \sim 0.75$ EDiSCs clusters.

Clusters studied in Kodama et al. (2004) are much poorer than MS1054. Because the former ones show a deficit, while the latter do not, the presence of a deficit may be correlated to the cluster richness. However, the evidence for the above is far from conclusive. Tanaka et al. (2005) show that the LF of the red galaxies has the same slope (within the errors) in clusters and in groups, and this holds both at $z = 0$ and 0.83. Similarly, Barrientos & Lilly (2003) do not find any evidence for a richness-dependent slope in their study of eight clusters at $z \sim 0.4$, whose x-ray luminosity spans a factor 200. Finally, poor clusters display a contrast over the background lower than rich clusters, and therefore the determination of a deficit is more difficult and subject to larger uncertainties. In fact, systems studied in Kodama et al. (2004) are so poor that spectroscopic evidence (Yamada et al. 2005) suggests that at least two of the Kodama et al. (2004) systems are line-of-sight superpositions of even poorer environments. To summarize, the deficit found in

Kodama et al., if genuine, may be related to lower density environments than clusters.

Goto et al. (2005) studied the very same cluster studied in this paper, MS1054. For the background subtraction they rely on spectroscopic data and the assumption that spectroscopically observed galaxies are a random sampling of the galaxies in the MS1054 line of sight. Our determination of M^* agrees with Goto et al., whereas we disagree on the existence of a deficit of faint red galaxies. However, at faint magnitudes their completeness is low (20 per cent) and therefore subject to large uncertainties important in ascertaining the existence of a deficit. Furthermore, judging from their fig. 8, their suggested deficit is just an $\approx$1–2σ effect, which is not confirmed by our data going 1 mag deeper. Finally, our determination of the LF slope is 4 times more precise than theirs.

4.1 Important caveats

The analysis of this paper is based on a single high-redshift cluster that is compared to some nearby or $z \sim 0.4$ clusters. The issue of the cluster being representative arises: can our results be generalized to the whole class of clusters? It is difficult to give an answer to this question. MS1054 has a large velocity dispersion, but not the largest among the clusters studied in this context, because one of the clusters studied in De Lucia et al. (2004) has a larger velocity dispersion (Poggianti et al. 2006). MS1054 has a bright L_X, but at least 15 more luminous clusters are known and therefore MS1054 is not extreme in its X-ray luminosity either. Furthermore, MS1054 is not more extreme in any of their properties than the clusters previously studied in our contest, because MS1054 is a part of this latter sample (and a part can never be more extreme than itself).

Overall, the problem of generalizing from a single example (or very few examples) is often the rule in the context of the build-up of the red sequence. De Lucia et al. have only a single reference (zero-redshift) comparison cluster. Goto et al. (2005) have, like us, one single cluster (the same we studied) at high redshift. Tanaka et al. (2005) has, like us, only one cluster at $z \sim 0.8$. Therefore, our result on MS1054 has the same poor generalization power of other published works: we are all in the process of building cluster samples from individual examples. However, after our work, there are six $z \sim 0.8$ clusters that show evidence *against* a build-up of the red sequence, MS1054, the four EDisCs clusters (after our revision) and also the $z \sim 0.8$ Tanaka et al. (2005) cluster, and none are in favour, giving us some confidence that the build-up of the red sequence in clusters is surely not a universal phenomenon at these redshifts, contrary some previous claims.

Beside the issue of sample representativity, we warn also about the methodology used in some literature to ascertain the existence of a build-up of the red sequence. Perhaps prompted by the covariance between errors on M^* and α, several authors prefer to measure the ratio between the number of luminous and faint galaxies at different redshifts, in order to look for an epoch dependency of the luminous-to-faint ratio. The ratio definition requires to consistently separate bright and faint galaxies in a coherent way at low and high redshift, raising the circular problem of inferring something about the luminosity evolution (from the measured ratios) *assuming* a luminosity evolution (required to define how the designations of magnitude ranges as 'bright' or 'faint' change with redshift). Also, if we know how the luminosity evolves (because this is needed to define the magnitude ranges of the ratio), there is no need to perform the experiment (why measure the ratios?). Once an author reaches a conclusion in disagreement with the hypothesis under which it has been derived (for example, a differential evolution between bright and faint galaxies having evolved by the same amount of evolution as the ranges of bright and faint galaxies) should he not revise the hypothesis?

Besides the logical inconsistency, the assumption of a level of evolution underestimates uncertainties by a significant factor, because, instead of marginalizing over the uncertain parameter (the unknown level of evolution), the uncertain parameter is taken to be known (a given level of evolution is assumed in determining the absolute magnitude ranges appearing in the ratio definition) violating the Bayes theorem and axioms of probability.

A statistical and logical correct way to measure the build-up of the red sequence strictly parallels section 5.2 in Andreon (2006) and its application to a large sample of high-redshift clusters is presented in a forthcoming paper (Andreon, in preparation).

5 CONCLUSIONS

Using one of the deepest data sets available, we determine that the red sequence of MS1054 is well populated at *all* studied magnitudes, showing no deficit of faint (down to $M^* + 3.5$) red galaxies: the faint end of the colour–magnitude relation is neither empty nor underpopulated. In order to quantify the presence of an eventual deficit, we determine the characteristic magnitude and slope of the LF of red galaxies in MS1054 at $z = 0.83$. Our determination of the faint end slope is a few times more precise than previous studies at the same redshift (e.g. Goto et al. 2005; Tanaka et al. 2005).

We found a flat slope in both the $\approx B$ and $\approx J$ rest frame, showing that there is no deficit of faint red galaxies, i.e. the abundance of red galaxies at faint magnitudes is similar to the one at intermediate magnitudes, down to the magnitude of completeness (M^*+3.5). Comparison with LFs at various redshifts suggests that the slope of the LF is not changed, within the errors, from $z = 0.83$ to $z \sim 0.4$ and $z = 0$. Therefore, the analysis of the LF shows no evidence for a decreasing (with magnitude or redshift) number of faint red galaxies. The presence of faint red galaxies in high-redshift clusters disfavours scenarios where the evolution of red galaxies is mass-dependent, because the mass dependency should differentially depopulate the red sequence, while the MS1054 colour–magnitude relation is populated as in nearby and $z \sim 0.4$ clusters. The presence of abundant faint red galaxies in the high-redshift cluster MS1054 restricts the room for allocating descendants of Butcher–Oemler galaxies, because descendants should change the faint end slope of the LF of red galaxies, while instead the same faint end slopes are observed in MS1054−0321, at $z \sim 0.4$ and in nearby clusters. In the rich MS1054 cluster, the colour–magnitude relation seems to be fully in place at $z = 0.831$ and therefore red galaxies of all magnitudes were wholly assembled at higher redshift. The largest uncertainty of our and literature conclusions resides in the small number of clusters studied, especially at high redshift, and their unknown representativity. Neverthess, six clusters (MS1054, the four EDisCS clusters and one cluster in Tanaka et al. 2005) support our conclusions and none supports a build-up of the red sequence at $z < 0.8$.

ACKNOWLEDGMENTS

I thank Gabriella De Lucia, Sperello di Serego Alighieri and Taddy Kodama for their comments on an early version of this paper. I thank the referee, Alfonso Aragon-Salamanca, whose comments urged me to find a solution for the problem described in Appendix A. I also thank Giulio D'Agostini, Roberto Trotta, Andrew Liddle, Sabrina Gaito, Bruno Apolloni and Dario Malchiodi for useful discussion and for their statistical advice. Finally, I am grateful for the use of the facilities of the *Hubble Space Telescope* and VLT.

REFERENCES

Andreon S., 2006, A&A, 448, 447
Andreon S., Pelló R., Davoust E., Domínguez R., & Poulain P., 2000, A&AS, 141, 113
Andreon S., Punzi G., Grado A., 2005, MNRAS, 360, 727 (APG)
Andreon S., Cuillandre J.-C., Puddu E., Mellier Y., 2006, MNRAS, submitted
Barrientos L. F., Lilly S. J., 2003, ApJ, 596, 129
Bell E. F., de Jong R. S., 2001, ApJ, 550, 212
Bruzual A. G., Charlot S., 2003, MNRAS, 344, 1000
Butcher H., Oemler A., 1984, ApJ, 285, 426 (BO)
Colless M. et al., 2001, MNRAS, 328, 1039
D'Agostini G., 2003, Bayesian Reasoning in Data Analysis – A Critical Introduction. World Scientific Publishing, Singapore
D'Agostini G., 2005, preprint (arXiv:physics/0511182)
De Lucia G. et al., 2004, ApJ, 610, L77
Donahue M., Voit G. M., Gioia I., Lupino G., Hughes J. P., Stocke J. T., 1998, ApJ, 502, 550
Förster Schreiber et al., 2006, AJ, 131, 1891
Garilli B., Maccagni D., Andreon S., 1999, A&A, 342, 408
Gioia I. M., Maccacaro T., Schild R. E., Wolter A., Stocke J. T., Morris S. L., Henry J. P., 1990, ApJS, 72, 567
Goto T. et al., 2005, ApJ, 621, 188
Jee M. J., White R. L., Ford H. C., Blakeslee J. P., Illingworth G. D., Coe D. A., Tran K.-V. H., 2005, ApJ, 634, 813
Kajisawa M. et al., 2000, PASJ, 52, 61
Kodama T., Bell E. F., Bower R. G., 1999, MNRAS, 302, 152
Kodama T. et al., 2004, MNRAS, 350, 1005
Kraft R. P., Burrows D. N., Nousek J. A., 1991, ApJ, 374, 344
Labbé I. et al., 2003, AJ, 125, 1107
Metropolis N., Rosenbluth A., Rosenbluth M., Teller A., Teller E., 1953, J. Chem. Phys., 21, 1087
Nakata F. et al., 2001, PASJ, 53, 1139
Oemler A. Jr., 1974, ApJ, 194, 1
Poggianti B. M. et al., 2006, ApJ, in press (astro-ph/0512391)
Sandage, A., Tammann G. A., Yahil A., 1979, ApJ, 232, 352 (STY)
Schechter P., 1976, ApJ, 203, 297
Secker J., Harris W. E., Plummer J. D., 1997, PASP, 109, 1377
Strateva I. et al., 2001, AJ, 122, 1861
Tanaka M., Kodama T., Arimoto N., Okamura S., Umetsu K., Shimasaku K., Tanaka I., Yamada T., 2005, MNRAS, 362, 268
Tanaka M., Kodama T., Arimoto N., Okamura S., Umetsu K., Shimasaku K., Tanaka I., Yamada T., 2006, MNRAS, 366, 1551
Tran K.-V. H., Kelson D. D., van Dokkum P., Franx M., Illingworth G. D., Magee D., 1999, ApJ, 522, 39
Yamada T. et al., 2005, ApJ, 634, 861
York D. G. et al., 2000, AJ, 120, 1579
Wilks S., 1938, Ann. Math. Stat., 9, 60
Wilks S., 1963, Mathematical Statistics. Princeton Univ. Press Princeton
Zwicky F. 1957, Morphological astronomy. Springer, Berlin

APPENDIX A: STATISTICAL INFERENCE: SLOPE, INTERCEPT, INTRINSIC SCATTER OF A CORRELATION IN THE PRESENCE OF A BACKGROUND POPULATION

We are here faced with a parameter estimation problem of a density distribution function given by the sum of two distribution functions, one carrying the signal (the cluster colour–magnitude distribution) and the other being due to a background (background colour–magnitude distribution) from the observations of many individual events (the galaxies luminosities and colours), without knowledge of which event is the signal and which one is background. The problem is complicated by the presence of a correlation that can be different for cluster and background galaxies, and by 'smearing' effects due to uncertainties on colours and magnitudes, as well as by complications as those derived by a non-vanishing intrinsic scatter between variables (e.g. the intrinsic scatter of the colour–magnitude relation).

Here, we present the solution of the full problem, also including options unused in this paper (but used in other papers e.g. Andreon et al. 2006), in order to present the method once.

Our approach starts from laws of probabilities and strictly parallels D'Agostini (2003, 2005). By only using the chain rule of probabilities, we derive a single likelihood function that accounts simultaneously for all the available data, cluster and background. In particular, D'Agostini (2005) derives the likelihood for a related problem that can be read, with obvious change of variable names, as determining the parameters and the uncertainty of a linear fit between colour and magnitude with normal errors on magnitudes, σ_m, on colours, σ_{col}, plus an intrinsic dispersion around the colour–magnitude relation, $\sigma_{\rm intr}$, and in absence of any background. The likelihood of the ith galaxy is given by his equation (52):

$$\mathcal{L}_i^{\rm no\,bkg} \propto \frac{1}{\sqrt{2\pi}\sqrt{\sigma_{\rm intr}^2+\sigma_{col_i}^2+a^2\sigma_{m_i}^2}} \times \exp\left[-\frac{(\mu_{col_i}-am_i-c)^2}{2\left(\sigma_{\rm intr}^2+\sigma_{col_i}^2+a^2\sigma_{m_i}^2\right)}\right], \quad \text{(A1)}$$

where a and c are the slope and intercept of the colour–magnitude relation and μ_{col_i} is the (unknown) true value of the colour of the ith galaxy.

D'Agostini (2005) does not address the problem of interlopers due to a background, which, instead, we address here. We need to combine the likelihood above with the likelihood that a galaxy of magnitude m_i is drawn from the correlation described by equation (A1) or from another distribution/correlation due to background galaxies. Finally, we should account for the background correlation being known with finite certainty. We also would like to use the information contained in the control field direction (were we know for sure that the correlation due to MS1054 cluster galaxies is not there, because the field images a distant part of the sky). Therefore, the likelihood of the ith galaxy, $\mathcal{L}_i$, is given by the sum of two terms.These are a distribution function gently varying (in mag m and col) to account for the background contribution plus a term due to cluster galaxies given by the product of the likelihood function in equation (A1) (because we assume that cluster galaxies obey a linear colour–magnitude relation) and a Schechter function (because we assume that cluster galaxies are distributed in magnitude according to such a function):

$$\mathcal{L}_i \propto \delta_{\rm c}\Omega_j\mathcal{L}_i^{\rm no\,bkg}Schechter(m_i) + \Omega_j\left[d+e\times m_i+f\times col_i+g\times m_i^2\right], \quad \text{(A2)}$$

where $\delta_{\rm c}=1$ for cluster data sets, $\delta_{\rm c}=0$ for the other data sets, d, e, f and g describe the shape of the distribution of galaxy in the colour–magnitude plane in the reference field direction (taken in this example as a second-order polynomial), Ω_j is the studied solid angle and *Schechter* is the usual Schechter function (equation 1), with (unknown) α, M^* and ϕ^* parameters. If background galaxy counts have a more complex colour–magnitude distribution than a gently decreasing distribution like our parametrization, more coefficients (or any other parametrization) can be used to describe its shape. Similarly, if the cluster LF is deemed to be more complicated, or the colour–magnitude relation is supposed to be potentially more complex than linear in *mag* (e.g. curved, as we considered in Andreon et al. 2006 for another data set), it is enough to plug into the equations (A1) and (A2) the aimed relations.

Given j data sets (say, cluster no. 1, cluster no. 2, ... field no. 1, field no. 2, ...) each composed of N_j galaxies, the likelihood $\mathcal{L}$ is given by the formula

$$\ln \mathcal{L} = \sum_{\text{data sets } j} \left(\sum_{\text{galaxies } i} \ln \mathcal{L}_i - s_j \right), \tag{A3}$$

where $\mathcal{L}_i$ is given by equation (A2) and provides the *unnormalized* (because integral is not 1) likelihood of the ith galaxy of the jth data set to have m_i, col_i.

s_j is the integral of likelihood function over the colour and magnitude ranges. Given as a formula,

$$s_j = \int \int \mathcal{L} \, \mathrm{d}m \, \mathrm{d}col. \tag{A4}$$

The integral should be performed on the appropriate colour and magnitude ranges (those accessible to the data) and it is equal by the *expected* number of galaxies (see APG for some examples). Section 3.3.5 in APG clarifies the danger of mistakenly taking s_j to be equal to the *observed* number of galaxies. The s_j term disfavours models that predict a number of galaxies very different from the observed one.

As usual, if errors on magnitude are not negligible, *Schechter* in equation (A4) should be replaced by the convolution between the Schechter function and the error function.

The next step is to summarize the result of the computation above with a few numbers. If regularity conditions are satisfied and the sample size is large, the analysis may proceed making reliance on the Likelihood Ratio theorem in order to estimate the 'best fit' and confidence contours, as we proceeded in APG. Alternatively, one may strictly follow laws of probability, as we did in Andreon et al. (2006): we first assume a prior, and then we compute the posterior and we summarize it quoting a few numbers. Technically, we use a Markov Chain Monte Carlo (Metropolis et al. 1951) with a Metropolis sampling algorithm in order to efficiently explore the 10-dimensional parameter space.

The proposed method does the following.

(i) It derives from first principles, i.e. from laws of probabilities, and does not violate them, while other methods do (examples are given in the following items).

(ii) It provides sensible results in all conditions (in presence of a large background and low cluster signal, for example), when other methods fail or provide results of unknown meaning. For example, the proposed method does not produce complex or negative values of the intrinsic scatter and does not return implausible values for the errors (errors crossing the physical border $\sigma_{\text{intr}} = 0$), oddities that instead bless and are frequent in other approaches.

(iii) It gives error bars that have properties that physicians expect them to have. For example, error bars are larger in the presence of background than in its absence, a property that other approaches do not always guarantee (see, e.g. the discussion in Kraft, Burrows & Nousek 1991, for an astronomical paper about the estimate of a Poissonian signal in the presence of a background in a simple case).

(iv) In contrast with other methods, we do not make reliance on the rule of sums in quadrature in regimes where it is known that the rule cannot be used (near borders or when regularity conditions are not satisfied).

The largest disadvantage of the present method is that it requires some effort of understanding because it is new in this context (but not in other ones), some coding and the computation of integrals, i.e. more work than a quick and dirty estimation computed with other tools. In exchange for some effort, the method returns numbers that can be trusted, instead of numbers of which we are unsure about their correctness.

Using such formalism, it is straightforward to compute the relative evidence of a linear or bended colour–magnitude relation, as we did in Andreon et al. (2006) for Virgo galaxies.

In this paper, we have only one cluster and one control field area. We consider only galaxies having $1.6 < V - I < 2.3$ mag, because we do not need to model a colour region where red objects are not, and $K > 16.5$ mag (or $I > 20$ mag for the I band), again in order not to model a region of the parameter space where MS1054 cluster galaxies are not, and, of course, we consider only magnitudes brighter than the completeness magnitude. Furthermore, we reparametrize the (linear) colour–magnitude relation as in the formulae below, in order to reduce the covariance between errors.

We quote averages and dispersions of the posterior. We take uniform priors over a large parameter range fully enclosing the region where the likelihood is non-vanishing and we further put to zero the prior in unphysical regions of the parameter and data space. Because the sample size is large and the parameters are well determined by our data, other choices of the prior would lead to indistinguishable results.

We found

$$V - I = 1.88 \pm 0.01 - (0.086 \pm 0.010)(K - 19.54)$$

with a total (i.e. intrinsic + photometric) scatter of 0.10 ± 0.01 mag and

$$V - I = 1.88 \pm 0.01 - (0.085 \pm 0.012)(I - 22.97)$$

with a total scatter of 0.14 ± 0.01 mag.

This paper has been typeset from a TEX/LATEX file prepared by the author.