A six-parameter space to describe galaxy diversification

Didier Fraix-Burnet, Tanuka Chattopadhyay, Asis Kumar Chattopadhyay, Emmanuel Davoust, Marc Thuillard

Astronomy & Astrophysics manuscript no. multivargal
June 15, 2012

A six-parameter space to describe galaxy diversification

D. Fraix-Burnet[1], T. Chattopadhyay[2], A.K. Chattopadhyay[3], E. Davoust[4], and M. Thuillard[5]

[1] Université Joseph Fourier - Grenoble 1 / CNRS, Institut de Planétologie et d'Astrophysique de Grenoble, BP 53, F-38041 Grenoble cedex 9, France e-mail: fraix@obs.ujf-grenoble.fr

[2] Department of Applied Mathematics, Calcutta University, 92 A.P.C. Road, Kolkata 700009, India

[3] Department of Statistics, Calcutta University, 35 Ballygunge Circular Road, Kolkata 700019, India

[4] Université de Toulouse, CNRS, Institut de Recherches en Astrophysique et Planétologie, 14 av. E. Belin, F-31400 Toulouse, France

[5] La Colline, 2072 St-Blaise, Switzerland

Received January 2012; accepted June 2012

ABSTRACT

Context. Galaxy diversification proceeds by transforming events like accretion, interaction or mergers. These explain the formation and evolution of galaxies that can now be described with many observables. Multivariate analyses are the obvious tools to tackle the datasets and understand the differences between different kinds of objects. However, depending on the method used, redundancies, incompatibilities or subjective choices of the parameters can void the usefulness of such analyses. The behaviour of the available parameters should be analysed before an objective reduction of dimensionality and subsequent clustering analyses can be undertaken, especially in an evolutionary context.
Aims. We study a sample of 424 early-type galaxies described by 25 parameters, ten of which are Lick indices, to identify the most structuring parameters and determine an evolutionary classification of these objects.
Methods. Four independent statistical methods are used to investigate the discriminant properties of the observables and the partitioning of the 424 galaxies: Principal Component Analysis, K-means cluster analysis, Minimum Contradiction Analysis and Cladistics.
Results. The methods agree on six parameters as most structuring: central velocity dispersion, disc-to-bulge ratio, effective surface brightness, metallicity, and the line indices NaD and $OIII$. The partitioning found using these six parameters, when projected onto the fundamental plane, looks very similar to the partitioning obtained previously on a totally different sample and using only the parameters of the fundamental plane. Two new groups are identified here, and we are able to provide some more constraints on the assembly history of galaxies within each group thanks to the larger number of parameters. We also find another "fundamental plane" with the absolute K magnitude, the linear diameter and the Lick index $H\beta$. We confirm that the $Mg\,b$ vs velocity dispersion correlation is very probably an evolutionary correlation, and find that several other such scaling relations are very probably evolutionary correlations as well. Finally, combining the results of our two papers, we obtain a classification of galaxies that is based on the transforming processes that explain the origin of the groups.
Conclusions. By taking into account that galaxies are evolving complex objects and using appropriate tools, we are able to derive an *explanatory* classification of galaxies, based on the physical causes of the diverse properties of galaxies, as opposed to *descriptive* classifications quite common in astrophysics.

Key words. galaxies: elliptical and lenticular, cD - galaxies: evolution - galaxies: formation - galaxies: fundamental parameters - methods: statistical

1. Introduction

Galaxies are complex and evolving objects. Their diversity is exploding with the instrumental improvements that feed huge databases. A good understanding of the physics governing the processes at work within and between the different components of galaxies leads to numerical simulations that produce synthetic populations of hopefully realistic objects. The number of possibilities for the variables involved in this physics renders the morphological Hubble classification and its equivalents obviously too simple. Morphology, as detailed as it could be determined in the visible, is only one component of the physics of galaxies, and ignores many ingredients of galaxy evolution, like kinematics and chemical composition (e.g. Cappellari et al., 2011). In addition, such classifications do not make full use of the wealth of information that observations and numerical simulations provide.

Multivariate partitioning analyses must now be used. One basic tool, the Principal Component Analysis, is relatively well-known (e.g. Cabanac et al., 2002; Recio-Blanco et al., 2006), but this is not a clustering tool in itself. Many attempts to apply multivariate clustering methods have been made very recently (e.g. Ellis et al., 2005; Chattopadhyay & Chattopadhyay, 2006, 2007; Chattopadhyay et al., 2007, 2008, 2009a,b; Fraix-Burnet et al., 2009; Sánchez Almeida et al., 2010; Fraix-Burnet et al., 2010). Sophisticated statistical tools are used in some areas of astrophysics and are developing steadily, but multivariate analysis and clustering techniques have not yet entirely penetrated the community. It is true that the interpretation of the results is not always easy. Some of the reasons are given below.

Before using the available parameters to derive and compare physical properties of galaxies, it is important to check whether they can *discriminate* between different kinds of galaxies. A partitioning of objects into robust groups can only be obtained with structuring parameters. This does not necessarily preclude using other information to help the physical and evolutionary interpretation of the properties of the groups and the relationships between groups. Among the descriptors of galaxies, many come directly from the observations independently of any model. In principle all the information is within the spectrum. But since it is a huge amount of information, it is usually summarized by broad-band fluxes (magnitudes), slopes (colors), medium-band and line fluxes (e.g. the Lick indices). This dimensionality reduction is often guided by observational constraints or some physical a priori, but not by discriminant (i.e. statistical) properties.

Multivariate partitionings group objects according to global similarities. They yield a descriptive classification of the diversity, but do not provide any explanation for the differences in properties between groups. Modelling or numerical simulations must be used to understand physically the partitioning and the relationships between the groups.

However, galaxy properties are in fine explained by evolution. Mass, metallicity, morphology, colors, etc, are all the result of galaxy evolution. It can thus be expected that the relationships between the groups are driven by evolution. In addition, galaxy (formation and) evolution proceeds through a limited number of transforming processes (monolithic collapse, secular evolution, gravitational interaction, accretion/merger, sweeping/ejection, see Fraix-Burnet et al., 2006b,c). Since they depend on so many parameters (initial conditions, nature of the objects involved, impact parameters, ...), the outcomes of each of them vary a lot, so that diversity is naturally created through evolution of the galaxy populations. This is what we call diversification.

It is easy to see that each of these transforming processes follows a "transmission with modification" scheme, because a galaxy is made of stars, gas and dust which are both transmitted and modified during the transforming event (Fraix-Burnet et al., 2006b,c,a). This is why a hierarchical organization of galaxy diversity might be expected with evolutionary relationships between groups. Cladistics has been shown to be an adequate (Fraix-Burnet et al., 2006b,c,a) and quite effective (e.g. Fraix-Burnet et al., 2009, 2010) tool for identifying this hierarchical organization. Instead of a descriptive classification of galaxy diversity, we hope to be able to build an explanatory classification that is physically more informative.

In a previous work (Fraix-Burnet et al., 2010) we found that the fundamental plane of early-type galaxies is probably generated by diversification. We believe that this result can be of such importance as to deserve dedicated studies to assess its robustness. This is why the present work brings several essential novelties:

- a distinct data set is used, for which there are more parameters available, useful for the analyses themselves and also for the subsequent interpretation, once the groups are determined (Sect. 2.1),
- two additional methods, PCA and MCA, are used (Sect. 2.2),
- a new set of parameters is used for the various partitioning methods (Sect. 4) and these are selected in a rather objective way (Sect. 3), unlike ad hoc selection of parameters from longtime conventional wisdom (as followed in Fraix-Burnet et al., 2010),
- measurement errors have been used in the classification in the case of cladistic analysis (Sect. 2.1 and Appendix B.2),
- the combination of the partitionings in the present paper and in Fraix-Burnet et al. (2010) should help devise a new scheme for galaxy classification.

We present the data in Sect. 2.1 before describing the philosophy of our approach with the different methods used to analyse the structuring properties of the parameters (i.e. their ability to discriminate different groups) and the partitioning of the sample in Sect. 2.2. Then we give the results of these analyses and the "winning" set of parameters (Sect. 3) that is used for the partitioning (Sect. 4). We then comment on the structuring parameters (Sect. 5) and detail the group properties (Sect. 6). Scaling relations, correlations and scatterplots are presented in Sect. 7, the well-known fundamental plane of early-type galaxies and the other fundamental plane, which we discovered in this paper, being discussed in Sect. 7.4. Finally, we combine the present result with the one in Fraix-Burnet et al. (2010) on a cladogram that summarizes the inferred assembly histories of the galaxies and thus is a tentative new scheme to classify galaxies (Sect. 8). The conclusion of this study closes this paper (Sect. 9).

2. Data and methods

2.1. Data

We selected the 424 fully documented galaxies from the sample of 509 early-type galaxies of the local Universe of Ogando et al. (2008). As these authors point out, these kinds of samples appear relatively small compared to those at intermediate redshifts obtained with large surveys (like the SDSS for instance), they have the advantage of having higher quality spectroscopic data and more reliable structural information like the effective radius. To describe the galaxies, we took from Ogando et al. (2008) the ten parameters that belong to the set of 25 Lick indices defined by Worthey & Ottaviani (1997): $H\beta$, $Fe5015$, $Mg1$, Mg_2, $Mg\,b$, $Fe5270$, $Fe5335$, $Fe5406$, $Fe5709$, NaD. We computed two other parameters from these Lick indices: [MgbFe]'= $\sqrt{Mgb * (0.72 * Fe5270 + 0.28 * Fe5335)}$ (Thomas et al., 2003) and Mg b/Fe= $Mgb/\left(\frac{1}{2}(Fe5270 + Fe5335)\right)$ (Gonzáles, 1997), which are indicators of metallicity and of light-element abundance respectively. The other parameters taken from Ogando et al. (2008) were the number of companions n_c, the morphological type T, the line index $OIII$, the velocity dispersion ($\log\sigma$) and the linear effective radius ($\log r_e$).

The surface brightness within the effective radius (Bri_e) and the disc-to-bulge ratio (D/B) were taken from Alonso et al. (2003). The absolute magnitude in B (M_{abs}) and the distance of the galaxies were taken from Hyperleda[1], which adopts a Hubble constant of 70 km/s/Mpc. The distances to four galaxies not available in Hyperleda were taken from the literature : NGC 1400 (27.7 Mpc, from Perrett et al. (1997)), NGC 4550 (15.49 Mpc from Mei et al. (2007)) and NGC 5206 (3.6Mpc from Karachentsev et al. (2002)). The color B-R was calculated from the corrected apparent B magnitude in Hyperleda and the total R magnitude given by Alonso et al. (2003). The linear diameter ($\log(diam)$) was computed from logdc given in Hyperleda. The infrared magnitudes and colors were taken or calculated from NED[2].

Altogether, we have 25 parameters to describe the 424 galaxies. However two parameters were removed for the analyses: the number of companions n_c and the morphological type T. They

[1] http://leda.univ-lyon1.fr/

[2] http://nedwww.ipac.caltech.edu/

are both discrete parameters. More importantly n_c is not a property of the galaxies, but of their local environment, while T is qualitative and subjective. Naturally, the full set of 25 parameters is used to interpret the results.

We thus used 23 parameters for the analyses in this paper: three are geometrical: D/B, $\log r_e$, $\log(diam)$; two come from medium-resolution spectra: $\log\sigma$ and $OIII$ in addition to the ten Lick indices ($H\beta$, $Fe5015$, $Mg1$, Mg_2, $Mg\,b$, $Fe5270$, $Fe5335$, $Fe5406$, $Fe5709$, NaD) and $[MgbFe]'$, $Mg\,b/Fe$, the six others are broad-band observables: Bri_e, the absolute magnitudes in B (M_{abs}) and K (K_{abs}), the total colors B-R, J-H and H-K.

The paper by Ogando et al. (2008) gives error bars for the parameters. However, evaluating the influence of measurement errors on the partitioning is difficult because their multivariate distribution function is unknown. This appears to be a big statistical problem. Fuzzy cluster analyses could perhaps be useful but they are quite complicated to implement, and the very good agreement between all our results render such an investment not so urgent at this point. In addition, measurement uncertainties can easily be integrated into the cladistic analysis. We thus limit ourselves to two restricted assessments: the influence of two determinations of the distance of galaxies (needed to determine r_e) on the result, and the cladistic analysis with errors. These are described in Appendix B. In any case, one should consider that the nature of galaxies implies continuous variations of the parameters, hence the partitions are necessarily fuzzy with no rigid boundaries between groups. This implies some uncertainty in the placement of the individual objects in the multivariate parameter space.

2.2. Methods

The philosophy of our approach is to use multivariate tools in a first step to select the parameters that can discriminate different groups within the whole sample. These parameters, called structuring parameters, are then used in a second step to partition the data in several groups.

In this paper, we use four methods, which are described in more details in Appendix A. Three of them are used to analyse the parameters: Principal Component Analysis (PCA, Sect. A.1), Minimum Contradiction Analysis (MCA, Sect. A.2) and cladistics (Sect. A.3), while the groupings are performed with the two latter (MCA and cladistics) together with a cluster analyses (CA, Sect. A.4),

The four approaches are all very different in philosophy and technique. Since there is no ideal statistical method, it is useful to compare results obtained with such independent methods. Convergence improves confidence, but since assumptions behind the different techniques are different, exact agreement cannot be expected. In the end, it is the physics that decides if a partitioning is informative or not.

3. Analyses of the parameters

In this section, we investigate the behaviour of the observables by three of the methods presented above: PCA, MCA and cladistics. These three multivariate techniques use the parameters directly instead of distance measures, as in the cluster analysis also considered in this paper. Hence, a lot of information can be gained about the parameters themselves, such as their correlations (PCA) or their respective behaviour in the partitioning process (MCA and cladistics). From this information, one can infer the discriminant power of all parameters for the studied sample.

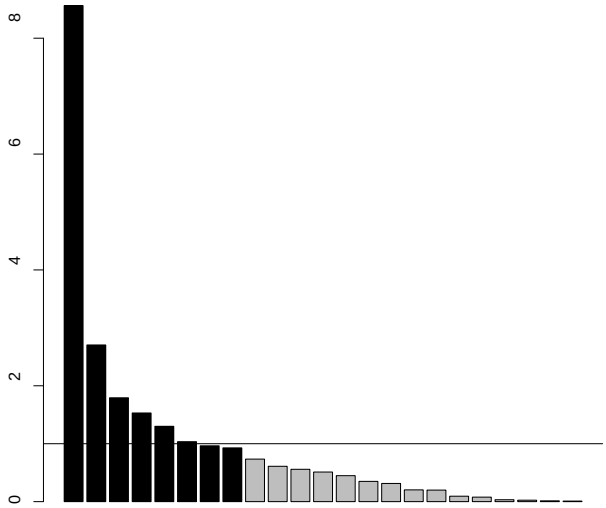

Fig. 1. PCA eigenvectors for the sample with 424 objects and 23 parameters. The eigenvalue 1 is indicated by the horizontal line and the eight eigenvectors higher than 1 or so are darkened.

3.1. Principal Component Analysis

We performed a first PCA analysis (Sect. A.1) on the set of 23 parameters. It shows that six principal components (PC or eigenvectors) have eigenvalues greater than 1 and two others are very close (Fig. 1), so that eight components describe most (82%) of the variance of the sample while the first five ones already account for 69% of the variance.

The loadings (i.e. the coefficients of the parameters composing the eigenvectors, Table A.1) give some indications as to which parameters are correlated, redundant or structuring. The most influent parameters (the first few for the first eigenvector, and the first one for the other PCs) in each PC are:

1. $Mg\,b$, $\log\sigma$, Mg_2, $[MgbFe]'$, NaD, Mg_1
2. Bri_e
3. $OIII$
4. Bri_e
5. H-K
6. J-H
7. D/B
8. $Fe5709$

Since $Mg\,b$, $Mg1$ and Mg_2 are so closely related (Burstein et al., 1984), the three are undoubtedly redundant. Moreover, $[MgbFe]'$ depends very much on $Mg\,b$, and is considered a better estimate of the metallicity of a galaxy. Hence, the most influent a priori non-redundant parameters are: $\log\sigma$, $[MgbFe]'$, NaD, Bri_e, $OIII$, H-K, J-H, D/B and $Fe5709$.

We then performed a second PCA analysis after removal of the supposedly redundant parameters ($Mg\,b$, $Mg1$, Mg_2, M_{abs}) and $\log r_e$ and $\log(diam)$that are affected by the uncertainties due to the distance determination (see Appendix B.2). We also disregarded J-H which has an outlier and otherwise quite constant values. Note that somewhat paradoxically this behaviour could explain why J-H appears in the sixth principal component above. We are now left with 16 parameters: $\log\sigma$, Bri_e, D/B,

$H\beta$, $Fe5015$, $Fe5270$, $Fe5335$, $Fe5406$, $Fe5709$, NaD, $OIII$, $H\text{-}K$, $B\text{-}R$, K_{abs}, $[MgbFe]'$, $Mg\,b/Fe$. Five eigenvectors have eigenvalues higher than 1 and account for 68% of the variance. The most influent parameters are now found to be:

1. $[MgbFe]'$, $\log\sigma$, NaD
2. $Fe5015$
3. Bri_e
4. $Fe5709$
5. $H\text{-}K$

The agreement is very good with $[MgbFe]'$, $\log\sigma$, NaD, Bri_e, $Fe5709$ and $H\text{-}K$ still present, while $OIII$ disappeared but has loadings very close to those of $Fe5015$ and $Fe5709$ in components 2 and 4 respectively. D/B do not appear as very influent in this analysis.

3.2. Minimum Contradiction Analysis

The MCA analysis (Sect. A.2) uses all parameters and explores them regarding the best order that can be obtained. It is possible to derive the structuring capacity of the parameters according to their respective behaviour as formalised in Thuillard & Fraix-Burnet (2009). We find that:

- $\log\sigma$, $Fe5270$, NaD, $[MgbFe]'$, Bri_e, $B\text{-}R$, $OIII$, D/B appear as structuring parameters.
- $\log\sigma$, K_{abs}, $\log(diam)$ are strongly correlated.

Contrary to PCA, the correlations are not automatically removed, some or all of them may remain. In the present case, the three strongly correlated parameters are not obviously redundant since they are not related by a direct causal relation (see Fraix-Burnet, 2011). However, keeping all of them for the MCA analysis does not bring more structuring information than keeping only one (Thuillard & Fraix-Burnet, 2009). As a consequence, since $\log\sigma$ is listed as one of the structuring parameters, K_{abs} and $\log(diam)$ can be disregarded in the analysis.

Consequently, the MCA analysis finds eight structuring parameters.

3.3. Cladistic Analysis

Each cladistic analysis (Sect. A.3) uses and investigates a given set of parameters. To better understand the 23 parameters, it would be necessary to analyse all possible subsets. Since this takes too much computing time, we decided to eliminate obvious redundancies ($Mg\,b$, $Mg1$, Mg_2, M_{abs}, $\log(diam)$). $H\beta$, which is an age indicator for stellar populations older than a few hundred Myr, is problematic for cladistics because age is a property of all groups. This parameter might be able to trace recent transformative events accompanied by starbursts, if it were not for the degeneracy between the age of a younger stellar component and its relative contribution to the total stellar mass or luminosity. Guided by the PCA and MCA analyses, we also disregarded $Fe5335$, $Fe5406$, $H\text{-}K$, $Mg\,b/Fe$ and $J\text{-}H$. It is remarkable that $\log r_e$ is not found in PCA and MCA analyses as a structuring parameter. We thus disregarded it also here, but we kept it for a specific analysis of the fundamental plane together with $\log\sigma$, Bri_e and Mg_2 (see Sect. 7.4.1 and Fraix-Burnet et al., 2010).

Finally, we studied in more details the remaining eleven parameters: $\log\sigma$, D/B, NaD, $[MgbFe]'$, Bri_e, $OIII$, $Mg\,b$, $Fe5015$, $Fe5270$, $Fe5709$, $B\text{-}R$. To find the most structuring ones in this list, we examined the relative robustness of the trees obtained by cladistic analyses using the eight subsets of these parameters listed in Table 1. Analyses of each parameter subset were performed with the full sample and several subsamples and the results compared. The details of our procedure are presented in Sect. A.3.

Table 1. Subsets of parameters used to determine the most structuring parameters from cladistic analyses. The names of the subsets include the number of parameters.

Subset	Parameters
4cA	$\log\sigma$ D/B NaD $[MgbFe]'$
5c	$\log\sigma$ D/B NaD $[MgbFe]'$ Bri_e
5cA	$\log\sigma$ D/B NaD $[MgbFe]'$ $Mg\,b$
6c	$\log\sigma$ D/B NaD $[MgbFe]'$ Bri_e $OIII$
6cA	$\log\sigma$ D/B NaD $[MgbFe]'$ $Mg\,b$ Bri_e
7c	$\log\sigma$ D/B NaD $[MgbFe]'$ Bri_e $OIII$ $Fe5015$
8c	$\log\sigma$ D/B NaD $[MgbFe]'$ Bri_e $OIII$ $Fe5015$ $Mg\,b$
10c	$\log\sigma$ D/B NaD $[MgbFe]'$ Bri_e $OIII$ $Mg\,b$ $Fe5270$ $Fe5709$ $B\text{-}R$

Five or six structuring parameters are favored by the cladistic analysis because the results are then more stable. The trees from subsets 5c and 6c are in very good agreement, that of 6c being nearly entirely structured, and the result with 6c is in very good agreement with the cluster analysis (Sect. 4).

We conclude that the most structuring set of parameters from the cladistic analyses is that of 6c: $\log\sigma$, D/B, NaD, $[MgbFe]'$, Bri_e, $OIII$.

3.4. Final set of structuring parameters

The PCA, MCA and cladistic analyses agree on the following five parameters: $\log\sigma$, $[MgbFe]'$, NaD, Bri_e and $OIII$, while they globally favor five to eight structuring parameters. The cladistic and MCA analyses point to D/B as an important parameter, which appears only weakly in PCA. $B\text{-}R$ is structuring in MCA only, while the iron indices appear with $Fe5015$ and $Fe5709$ on one side (PCA), and $Fe5270$ on the other (MCA). None of these four parameters are preferred in the cladistic analyses.

Hence, we select the consensual six parameters $\log\sigma$, D/B, NaD, $[MgbFe]'$, Bri_e, $OIII$ for partitioning analyses of our sample.

4. Partitioning of the sample galaxies

We now compare the partitioning obtained with four methods: a cluster analysis using eight principal components (Sect. 4.1), a cluster (Sect. 4.2) and a cladistic (Sect. 4.4) analysis using six parameters, and a MCA optimisation with the same six parameters (Sect. 4.3). The partitionings are compared at the end of this section and in Fig. 2. The order of the groups for cladistics is essentially dictated by the tree (and its rooting, see Sect. 4.4), while for the other methods, the order has been arbitrarily chosen to correspond as much as possible to the cladistic order.

4.1. PCA plus Cluster Analysis

A cluster analysis (Sect. A.4) was performed using the eight PCs obtained by PCA (Sect. 3.1 and Sect. A.1). This analysis will be noted PCA+CA in this paper. Three groups are found and labelled PCACA1, PCACA2 and PCACA3. The first two have about 100 objects, while the third is about twice as big (Fig. 2).

As noted in Sect. A.1, the use of principal components in multivariate clustering very likely misses a significant part of

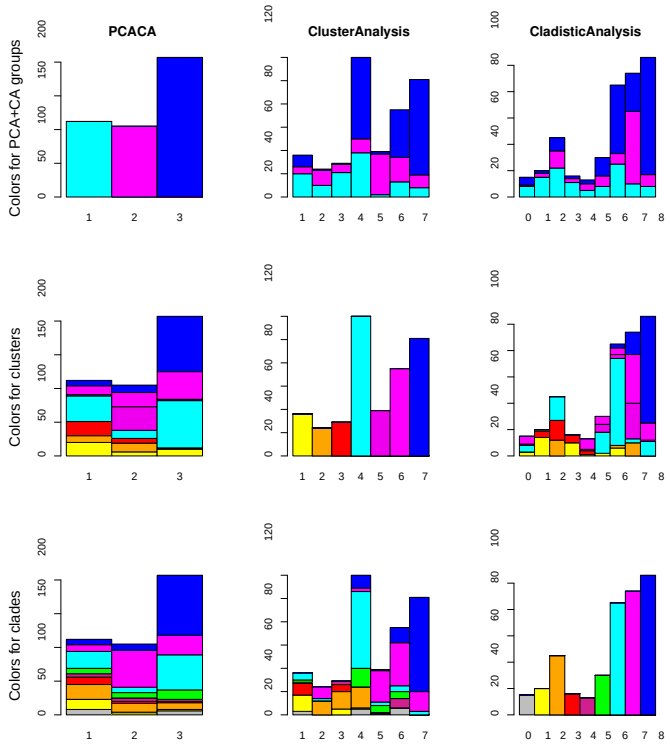

Fig. 2. Comparison of analyses with three different methods: PCA+CA (left panels), cluster analysis (middle panels) and cladistics (right panels). The MCA result is not shown since it can be easily compared to the cladistic result (see Sect. 4.5). In the first row, the colors identify the eight groups found in the cladistic analysis; in the second row, they identify the seven groups of the cluster analysis and, in the third row, the three groups found in PCA+CA. The colors for the cluster and PCA+CA groups are chosen to more easily visualize the agreement with the cladistic partitioning.

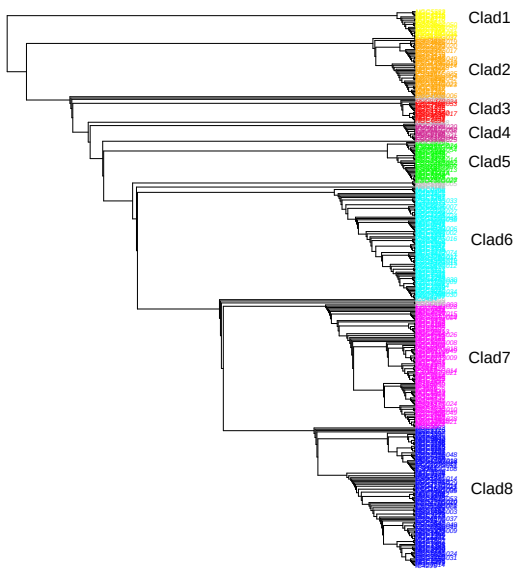

Fig. 3. Most parsimonious tree found with cladistics with the identification of the eight groups and their corresponding colors.

the underlying physics since it suppresses all correlations, even those that are due to hidden parameters or independent evolutions (see Sect. 6). We present this result here mainly as an illustration of this point.

4.2. Cluster Analysis

A cluster analysis (Sect. A.4) was performed with the six parameters listed in Sect. 3.4. Seven groups were found and named Clus1 to Clus7. There are three large groups, with about 80 to 120 objects. The other ones have around 20 to 40 objects (Fig. 2).

4.3. Minimum Contradiction Analysis

With the six parameters listed in Sect. 3.4, the MCA analysis (Sect. A.2) performs an optimisation of the order to minimise the contradiction. The result is four groups, and maybe two others. Globally, the groups are very fuzzy, i.e. they have no sharp limits. This is expected because of the continuous nature of the parameters, and because of uncertainties and measurement errors. This is an important point essentially overlooked by the other methods, and that should be kept in mind.

As we will see below, these four groups are easily identified with the groups obtained by cladistics, and for this reason they are not given labels in this paper.

4.4. Cladistic Analysis

The cladistic analysis, performed with the six parameters selected in Sect. 3.4, produces a most parsimonious tree (shown in Fig. 3), on which we can identify groups. There is no absolute rule to define groups on a cladogram. However, substructures in the tree are a good guide.

We have identified eight groups on this tree, three large ones with more than 80 objects, an intermediate group with about 50 objects, and four smaller ones with fewer than 30 members (Fig. 3). These groups are named "Clad1" (the most ancestral one, at the top) to "Clad8" (at the bottom) (Fig. 3). This numbering and presentation of the tree should not a priori be seen as a diversification arrow since branches can be switched graphically. It is the physical interpretation that both confirms the possible ancestrality of group Clad1 and gives the right order of diversification.

Rooting of the tree (i.e. the choice of the objects that appear graphically at the top of the tree and are supposed to be the closest to the common "ancestor species" of all the objects of the sample) is necessary to define the direction of diversification and in general affects the contours of the groups. The tree is here rooted with the group of lowest average metallicity as measured by $[MgbFe]'$ according to our assumption for primitiveness: low metallicity is more ancestral. This guess is monovariate and might not represent the best choice in a multivariate study like the present one. However, we do not yet have a better multivariate criterion for primitiveness. The rooting of the tree can be changed easily.

Contrary to the other partitioning methods, some objects appear isolated on the tree and are consequently not easily grouped with others. They could indeed represent a class by themselves, but, for the sake of simplicity, we decided not to identify them

with specific colors. We simply gather all such objects as Clad0, give them a grey color on plots or simply disregard them in the discussions that deal with the statistical properties of groups.

4.5. Comparison of the four partitionings

The four methods produce three (PCA+CA), four (MCA), seven (cluster analysis) and eight (cladistics) groups. They thus all agree on a relatively small number of groups.

The agreement between cladistics and PCA+CA is quite good (see Fig. 2), if we identify the three following groups: (Clad1,Clad2,Clad3,Clad4,Clad5 and a part of Clad6), (Clad7) and (Clad6,Clad8) with PCACA1, PCACA2 and PCACA3 respectively.

The agreement between PCA+CA and cluster analysis is also quite good with PCACA3 composed of Clus4, Clus7 and a part of Clus6, PCACA2 composed essentially of Clus5 and partly Clus6, and PCACA1 mainly composed of Clus1, Clus3 and part of Clus4.

Cluster and cladistic partitionings agree very well on the three big groups (Clus4≃Clad6, Clus5+Clus6≃Clad7, Clus7≃Clad8). The situation is slightly more complicated with the other groups, but still convergent with Clus1≃Clad1+Clad3, and Clad2 split into Clus2, Clus3 and Clus4 dominant in the first two. Conversely, Clus2 is mainly in Clad2 and also in Clad7.

The four groups from MCA are in very good correspondence with groups Clad6, Clad7, Clad8 and Clad1+Clad3. On the other hand, Clad2+Clad5 does not seem well justified from the MCA result. Interestingly, Clad6 and Clad8 are not quite independent, in agreement with PCACA3 being mainly composed of Clad6 and Clad8 as seen above.

The PCA+CA identifies a lower number of groups as compared with the other partitionings. This was expected, because of the effect of the PCA analysis, which eliminates too many of the correlations (Sect. A.1). The MCA result reinforces the significance of groups Clad6, Clad7, Clad8 and Clad1+Clad3, which are all also identified in the cluster analysis. The other groups from cladistics and cluster analyses are less robust or more fuzzy.

We conclude that the number of groups is at least four, and probably seven or eight. In the following, we consider the cladistic result with eight groups because it provides the very important evolutionary relationships between them.

Supplementary figures in Appendix C are given for the cluster partitioning and can be used to check that the interpretation does not depend on the detailed boundaries of the groups. In addition, two complementary cladistic analyses were performed in order to check the influence on the result of two determinations of the distance of galaxies (needed to determine r_e) and of measurement errors are presented in Appendix B.

5. Structuring descriptors of galaxies

Among the initial 23 quantitative parameters (Sect. 2.1), only six are structuring and actually yield a relatively robust partitioning (Sect. 3.4). The 17 remaining parameters do not yield enough information to discriminate different classes of objects, either because they are intrinsically not informative, bear the same redundant information as the structuring ones, or are not discriminant for the sample under study.

It is remarkable that the global luminosity of the galaxies (M_{abs} or K_{abs}) is not structuring. It is usually used as an indicator of mass and chosen as a main criterion of a priori classification. Luminosity is also often taken as characterising the level of evolution (for instance, it is implied in the so-called "downsizing effect"). However, from a diversification point of view, the absence of the global luminosity is expected, since mass is a global property that can be acquired by different processes, i.e. accretion or merging, which have different timescales and perturbing powers. Such parameters, which show too much convergence, are not well suited to establish phylogenies, that is, they are not good tracers of the assembly history of galaxies. Mass is bound to increase, it is thus not specific to a particular assembly history, which could distinguish different kinds of galaxies. Nevertheless, mass is not entirely absent and is represented somehow in $\log\sigma$ and Bri_e, which are certainly better tracers of the way mass has been assembled than mass itself.

$OIII$ which tends to decrease in more metallic galaxies, is a structuring parameter, but $H\beta$ which is often used as an age indicator, is not. This is not so surprising since age is not an indicator of diversity, as it is shared by all objects (see a discussion in Fraix-Burnet et al., 2009). Age, even more than mass or size, is bound to increase independently of the assembly history. Anyhow, defining an age for a galaxy is tricky and is often taken as the average age of the stellar populations, which is a poor tracer of the assembly history.

The size parameters $\log r_e$ and $\log(diam)$ are not structuring. They are probably merely scaling factors somewhat similar to mass, bound to increase whatever the sequence of transforming events that occur during the assembly history of galaxies. But size does not seem to be represented at all, and if so probably weakly in $\log\sigma$ and Bri_e, or even in some hidden correlation, which we will study later on in this paper.

However, one may wonder why Fraix-Burnet et al. (2010) found a robust partitioning using only four parameters, two of which are in our list of six ($\log\sigma$ and Bri_e) and one (Mg_2) is very similar to $[MgbFe]'$. But the fourth is $\log r_e$, which is not a structuring parameter in the present analysis. There are several reasons for this.

First, in Fraix-Burnet et al. (2010), the four parameters were not the result of a multivariate and objective selection, they were chosen because common wisdom suggests that they may be important to characterize the physics of galaxies. The very positive result obtained with these four parameters strongly supports this a priori, but the present paper demonstrates that only three of them are really structuring parameters.

Second, three parameters out of four are structuring, so that the partitioning signal is borne by these three. Unless the fourth parameter ($\log r_e$) is strongly erratic or contradictory, this signal is not expected to be entirely destroyed (see Sect. 7.4.1).

Third, the structuring parameters may in principle be different from one sample to another, if the diversity of objects is not covered equally. They may also depend on the initial set of parameters, if more structuring ones are present in a larger list. This is probably the case for $\log r_e$ which has been replaced by better observables.

It is thus not surprising that the four parameters used in Fraix-Burnet et al. (2010) yielded a robust partitioning and that we find more structuring parameters in the present study. The six parameters selected in the present analyses will not necessarily be the best ones for other samples, which may require new partitioning analyses.

6. Group properties

The groups identified by the partitioning methods must be understood in the light of their statistical and comparative properties.

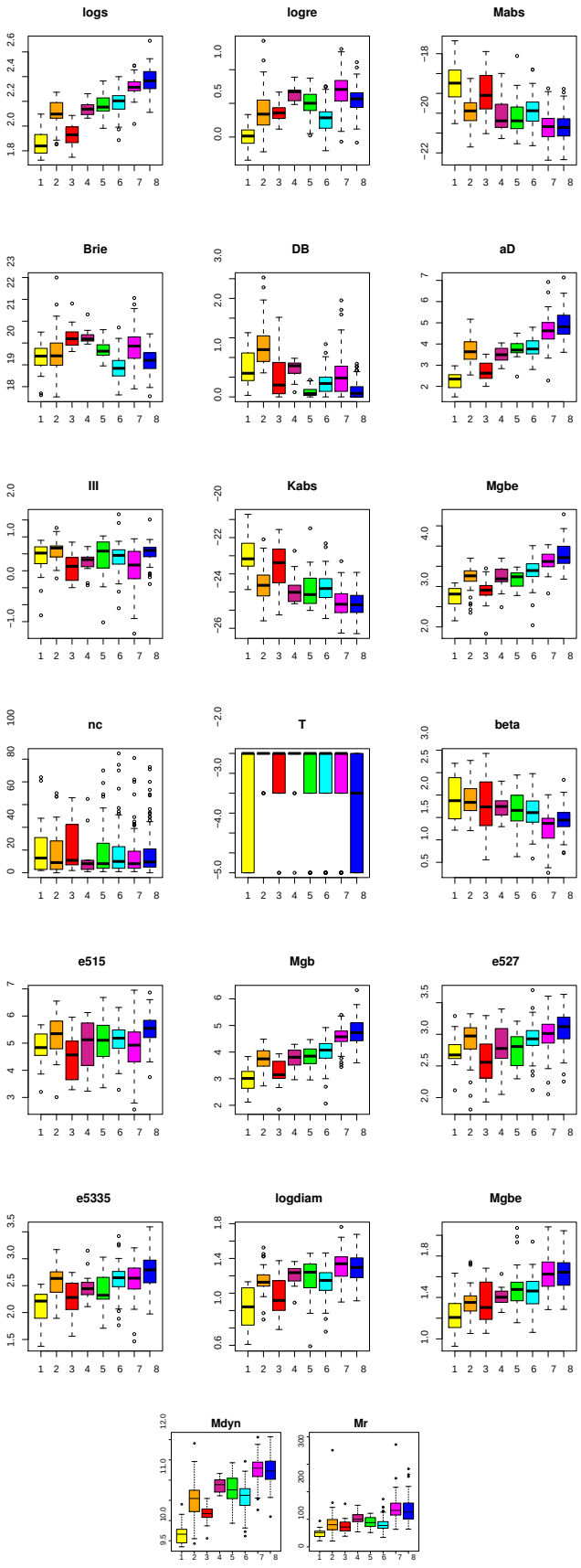

Fig. 4. Selection of the most interesting boxplots.

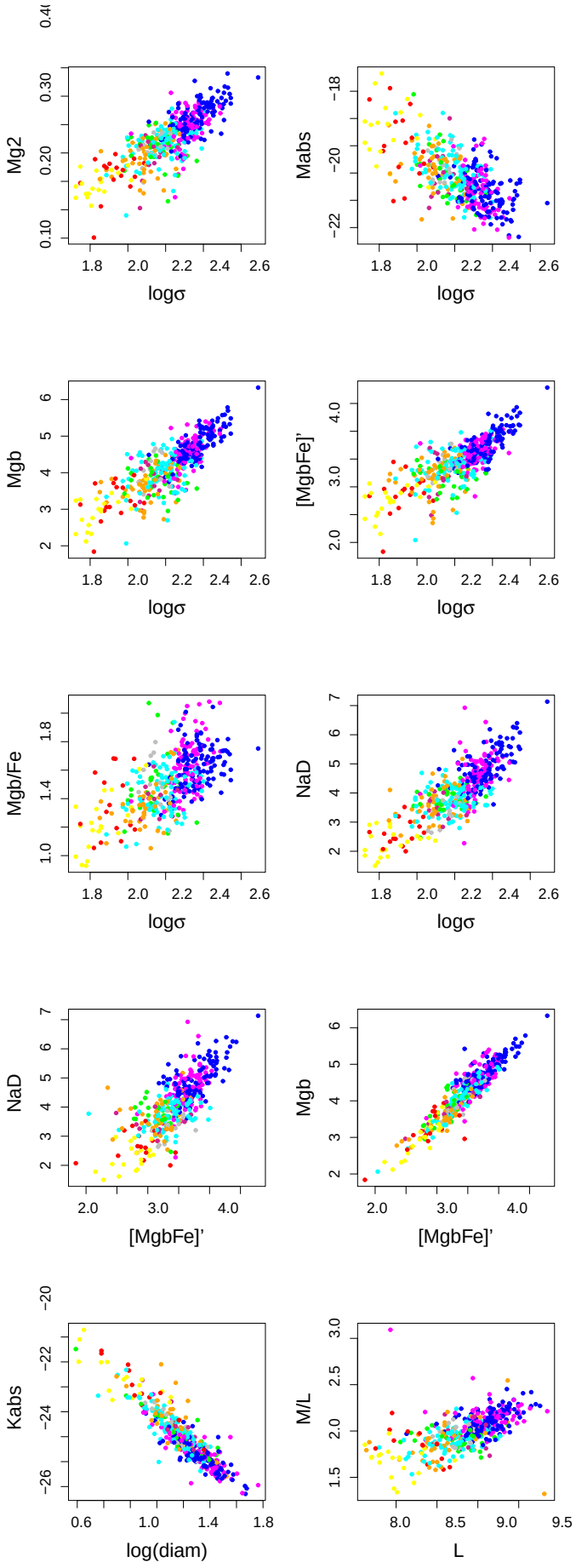

Fig. 5. Scatter plots showing evolutionary correlations. Colors are the same as in Fig. 4 and Fig. 3.

In this section, we first identify the main trends along diversification and then describe the distinctive properties of the groups.

For this purpose, we use boxplots, which give the four quantiles of each parameter for the eight groups. We consider two additional parameters. The dynamical mass is defined as $M_{dyn} \simeq A\sigma^2 R_e/G$ with $A = 3.8$ (with M_{dyn} in solar mass, σ in km s^{-1} and R_e in kpc) according to Hopkins et al. (2008) as done in Fraix-Burnet et al. (2010). This makes $M_{dyn} \simeq 5.95 * \sigma^2 R_e$. Using this mass, we compute the mass to light ratio M/L_r (using $Bri_e = -2.5 \log(L_r/\pi r_e^2) + 4.29$).

We show the most informative boxplots in Fig. 4; the other ones do not show significant differences between groups. We show the boxplots for the cluster partitioning in Fig. C.1 in Appendix C.

Figure 4 shows that $\log\sigma$, $\log r_e$, NaD, $[MgbFe]'$, $Mg\,b$, $\log(diam)$, $Fe5270$, $Fe5335$, $Mg\,b/Fe$, M_{dyn} and M/L_r are essentially increasing along the diversification rank defined on the tree of Fig. 3, while $H\beta$, M_{abs} K_{abs} and possibly D/B decrease. As already mentioned (Sect. 4.4), this rank is not necessarily as linear as it seems. Anyhow, the adopted rooting of the tree gives a very sensible result: globally galaxies tend to become more metallic, more luminous, more massive and larger with increasing diversification. At the same time, they acquire a larger central velocity dispersion, often related to the mass increase, and NaD is also known to increase along with mass and velocity dispersion. Also, the decrease of $H\beta$ indicates that the average age of the stellar populations increases with diversification.

$Mg\,b/Fe$ increases with diversification. The index [α/Fe] is known to increase with galaxy mass and age, because successive mergers and accretions trigger more intense star formation over shorter time scales. Clearly these events participate in the diversification of galaxies, confirming our observed increase of $Mg\,b/Fe$.

The $OIII$ index does not show any trend with diversification, but has a lower median value in three groups: Clad3, Clad4 and Clad7.

There is no systematic trend in environment with diversification, n_c shows a large range in all groups, except in Clad4 where it is small. Since an observed galaxy is the result of a long and multiple sequence of transforming events, it is probably the past environment, rather than the observed one, that plays a role in the diversification process.

On average, the most diversified groups (Clad5 to Clad8) have a lower D/B ratio, suggesting that transforming events, like accretion, interaction and merger, tend to destroy discs and build larger bulges, presumably by randomizing stellar orbits.

The morphological type is unevenly distributed among groups, Clad2 and Clad 4 having nearly only galaxies with $T = -2$, whereas $T = -5$ galaxies are found mainly in Clad1 and Clad8 (respectively the most ancestral and one of the most diversified groups).

Apart for the general trends with diversification, the groups have distinctive properties, otherwise there would be no reason for separate groups. These distinctive properties are the following, where the number of members is given in parentheses:

- Clad 1 (20 objects): the galaxies have the properties expected from an ancestral group: they are small, faint, discy, of low metallicity. They are young (although with a large spread in $H\beta$) and have a large spread in morphological type. They have very low M_{dyn} and $\log\sigma$, and a low M/L_r.
- Clad2 (45): the galaxies have the same average $H\beta$, Bri_e and $OIII$ as Clad1. They are larger, brighter, more massive, more metallic and have a much higher $\log\sigma$ than both Clad1 and Clad3. They are all of morphological type $T = -2$, and have the highest D/B of all groups by far.
- Clad3 (16): the galaxies show a large range in many parameters ($H\beta$, M_{abs} and K_{abs}, $\log(diam)$, n_c, $OIII$, $Mg\,b/Fe$), but not in $[MgbFe]'$, $\log r_e$, $\log\sigma$, Bri_e, NaD or M_{dyn}. They have a small $\log(diam)$ similarly to Clad1, but a much higher $\log r_e$. They are relatively faint with a relatively low $\log\sigma$ and $OIII$, and a high Bri_e.
- Clad4 (13): the galaxies resemble those of Clad2 in most respects (see discussion below on the respective placements of Clad2 and Clad3). In particular, they are all of morphological type $T = -2$. The main differences are that Clad4 objects are large (the highest $\log r_e$ after Clad7), of low surface brightness (high Bri_e) , with higher M_{dyn} and M/L_r, and slightly less discy.
- Clad5 (30): the galaxies are very much like the Clad4 ones, except that they have a low Bri_e and a very low D/B, the lowest of all groups with Clad8.
- Clad6 (85): it is one of the three big groups which are also the most diversified. Its galaxies have unexpectedly low values of M_{dyn}, $\log r_e$ and Bri_e. Interestingly, they have very similar properties to those of Clad2 galaxies, except for a much lower D/B, slightly lower Bri_e and $H\beta$, and a slighty higher $Mg\,b/Fe$.
- Clad7 (94): the galaxies are the largest in this sample. They are the most luminous and have the highest $\log\sigma$, M_{dyn} and M/L_r together with Clad8. Clad7 galaxies have a higher Bri_e, slightly lower $H\beta$ and $OIII$, and a slightly higher D/B than Clad6 and Clad8.
- Clad8 (106): their distinctive properties are a very low D/B and a very large spread in morphological type, like Clad1. They show a higher $\log\sigma$ and a lower Bri_e than galaxies of Clad7. They have high M_{dyn} and M/L_r as Clad7.

The Clad2 group often departs from the general trend along diversification (Fig. 4), which would seem smoother if Clad2 and Clad3 were inverted. We have noted (Sect. 4.5) that Clad2 is split between Clus2, Clus3 and also Clus4. It is significantly higher than expected in $\log\sigma$, D/B, NaD, $[MgbFe]'$, $Mg\,b$, $\log(diam)$, $Fe5270$, $Fe5335$, and lower in M_{abs} and K_{abs}. This means that, because of some parameters, it seems misplaced in the diversification scenario for other parameters. This behaviour is visible in the partitioning from the cluster analysis (Fig. C.1) since Clus2 and Clad2 are partially similar. So why has Clad2 been placed so early by the cladistic analysis, while it is more diversified in four of the six parameters used for the analysis?

The diversification scenario given on the tree of Fig. 3 is obtained from the parsimony criterion, which chooses the simplest combined evolution of all parameters. Taken individually, the simplest evolutionary curve of each variable is monotonic with as few reversals as possible. For instance, on the $\log\sigma$ boxplot of Fig. 4, one would expect Clad2 and Clad3 to be inverted to avoid the Clad2 box to "peak". However, this is a multivariate compromise, and since Clad2 would be better placed in the very first position on the D/B plot, to "smooth" the evolutionary curve, it is understandable that this is the most parsimonious placement on the tree. In addition, while the two structuring parameters Bri_e and $OIII$ show a variable behaviour, they would nevertheless induce us to place Clad2 before Clad3.

The conclusion is that Clad2 is correctly placed in second position, because this is a multivariate analysis, which seeks a compromise among several parameters. This shows the importance of selecting the parameters objectively, with multivariate

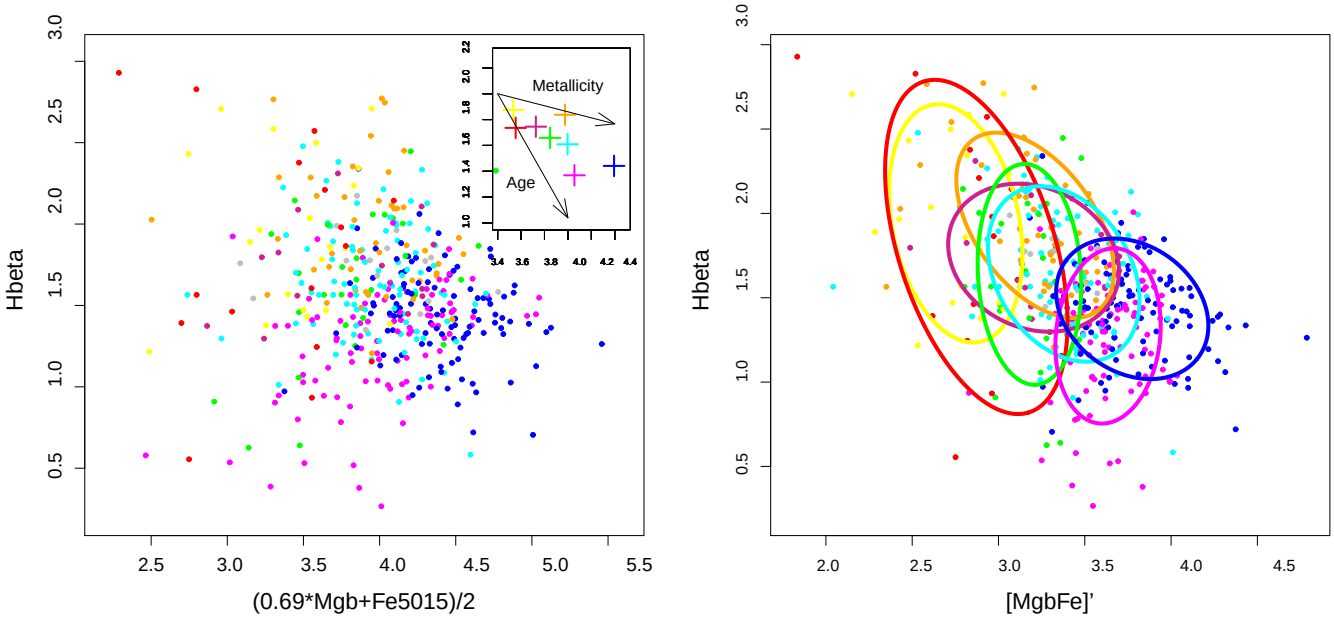

Fig. 6. Left: Equivalent of Figure 9 by Thomas et al. (2011) or Figures 3 and 8 by Kuntschner et al. (2010). The inset shows the median for each group and the arrows indicating the direction of increase of age and metallicity from single stellar population evolutionary models as shown by these authors. Right: Group inertia ellipses for the $H\beta$ vs $[MgbFe]'$ scatter plot.

tools. Otherwise, with too many redundant parameters, the peculiar properties of Clad2 could have been easily lost.

The relative and distinctive properties of the galaxies from the different groups obviously cannot be summarized with one or two physical parameters only. The relative properties of the groups show that the evolution of galaxies is not linear. The global trend in some properties (like mass or metallicity or $H\beta$) may appear roughly linear globally, but a detailed analysis, and especially the distinctive properties within each group, give many clues to understand the assembly history of the corresponding galaxies.

Having highlighted the group properties, we examine the possible correlations between them, by way of scatter plots, in the next two sections.

7. Scatter plots and correlations

Scatter plots must be examined with the partitioning to look for different behaviours between groups or within groups.

In the first case, the distribution of the groups traces the projected evolutionary track given by the tree. The fundamental plane is one example (Sect. 7.4.1 and Fraix-Burnet et al., 2010). We will however focus on cases showing a roughly linear track, where the groups are approximatively ordered along a linear correlation. We call them evolutionary correlations (Sect. 7.1) since these groups are related in cladistics by evolutionary relationships. They are important since they imply that the observed relation can be mainly generated by evolution as found in Fraix-Burnet et al. (2010) and formalised in Fraix-Burnet (2011).

In the second case (Sect. 7.3), some correlation may or may not be present within a given group, independently of the global behaviour between groups.

7.1. Evolutionary correlations

We confirm that the evolutionary nature of the $Mg_2 - \log\sigma$ correlation found by Fraix-Burnet et al. (2010) and find several other cases, the best ones being shown in Fig. 5. Such evolutionary correlations are revealed by the succession of groups ordered along the correlation with the most ancestral group (Clad1) at one end and the most diversified ones (Clad7 and Clad8 here) at the other end.

Several of these evolutionary correlations involve the following set of parameters: $\log\sigma$, $\log r_e$, M_{abs} (and K_{abs}), $H\beta$, $Mg\,b$ (and $Mg1$, Mg_2, $[MgbFe]'$ and $Mg\,b/Fe$), NaD and $\log(diam)$. Some relations are particularly tight (like $\log(diam)$ vs K_{abs}or $Mg\,b$ vs $[MgbFe]'$). The iron Lick indices $Fe5270$, $Fe5335$ and $Fe5406$, as well as H-K, also show an evolutionary correlation with M_{abs} although quite loose.

In all cases, except for $\log(diam)$ vs K_{abs} and $Mg\,b$ vs $[MgbFe]'$ (which are discussed in Sect. 7.3), the correlation is not present within each group. This is a clear sign that there is no direct causal physical link between the two variables, but simply a change on average with galaxy diversification.

Thomas et al. (2005) discuss the origin of the $Mg\,b$ vs $\log\sigma$ correlation. They find that metallicity, not age, is the main driving factor. This would again justify the use of metallicity as a reasonable tracer of diversification. But we find rather that the $Mg\,b$ vs $\log\sigma$ correlation is an evolutionary correlation, implying that diversification is indeed the real driver: metallicity, like central velocity dispersion, is bound to change on average as the galaxies evolve. This could explain why investigations find that this correlation appears so sensitive to several parameters (it has been proposed to be driven by metallicity, age, and relative abundance of different heavy elements, see Matković et al., 2009, for references). This sensitivity more probably points to an underlying, hidden and confounding factor, which creates the apparent correlation (Fraix-Burnet, 2011).

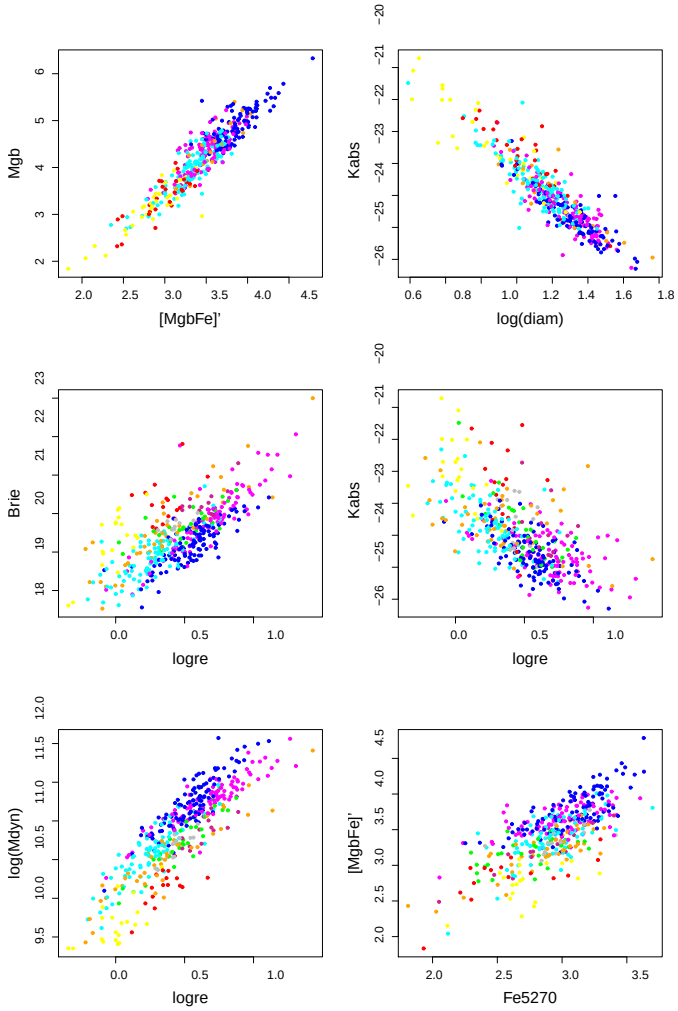

Fig. 7. Scatter plots showing correlations within groups.

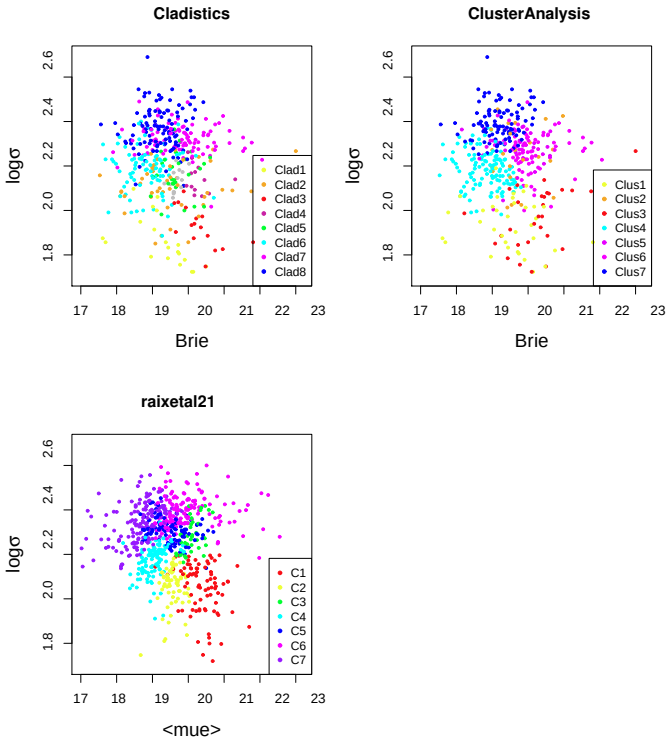

Fig. 8. Projection on the fundamental plane. Comparison between the partitionings obtained by cluster and cladistic analyses and that of Fraix-Burnet et al. (2010). See also FIg.C.3

.

The correlations between $Mg\,b$ and $[MgbFe]'$ and between NaD and $[MgbFe]'$ are clearly evolutionary, with diversification increasing from left to right in Fig. 5, while the correlations found by Thomas et al. (2011) are driven by total metallicity, which, for given age and light-element ratio, increases from left to right in their Figs 6 and 8. The dispersion of the NaD vs $[MgbFe]'$ relation is higher than that of $Mg\,b$ vs $[MgbFe]'$ in our data and theirs and not well accounted for by their model, presumably because of the fixed age and light-element ratio assumptions. Nevertheless, there is agreement between our result and their model since the average metallicity of galaxies obviously increases with diversification.

The $[MgbFe]'$ vs $\log\sigma$, $Mg\,b/Fe$ vs $\log\sigma$ and $Mg\,b$ vs $\log\sigma$ correlations (Fig. 5) can be compared with the Z/H vs $\log\sigma$ and α/H vs $\log\sigma$ in Figure 16 in Kuntschner et al. (2010). The same correlation is present, but we clearly show its evolutionary nature. Kuntschner et al. (2010) ask the question: *What drives the [α/Fe] – $\log\sigma_e$ (or mass) relation?* Our answer is simply: diversification. Indeed, they arrive at the same conclusion because they find that, in their sample, *there is evidence that the young stars with more solar-like [α/Fe] ratios, created in fast-rotating disc-like components in low- and intermediate-mass galaxies, reduce the global [α/Fe] and thus significantly contribute to the apparent [α/Fe] – $\log\sigma_e$ relation*. These galaxies belong to our Clad1 group as stated above, but they are not the sole responsible cause for this "apparent" relation, since all our groups are aligned along the same trend.

The Faber-Jackson relation, M_{abs} vs $\log\sigma$ (Fig. 5), also appears to be a purely evolutionary correlation: the sequence of evolutionary groups are aligned along this correlation. If mass had been the hidden parameter, then a similar correlation should exist within each group. This is not the case. This result is corroborated in an independent way by Nigoche-Netro et al. (2011).

7.2. Diversification or ageing?

There is a well-known degeneracy between the age (measured by $H\beta$) and the metallicity (measured by $[MgbFe]'$ or $(0.69*Mgb+Fe5015)/2$) of stellar populations, which models for the stellar evolution have tried to lift (e.g. Tripicco & Bell, 1995). In particular, Thomas et al. (2011) and Kuntschner et al. (2010) show evolutionary tracks from different models applied to galaxy observations of $H\beta$ as a function of a metallicity indicator defined by $(0.69*Mgb+Fe5015)/2$. We reproduce the same figures in Fig. 6 (left) showing with crosses the median for each group and with arrows the principal direction of increase for age and metallicity. We emphasize that the stellar evolution models by Thomas et al. (2011) and Kuntschner et al. (2010) used for the figures are single population models with fixed solar value of [α/Fe], while the data used in this paper (Sect. 2.1) are integrated over the whole galaxy, mixing together the contributions of possibly several different stellar populations.

Our groups are clearly arranged according to diversification following an increase in both age and metallicity. The spread of the correlation is large, and, within each individual group, the range in age and metallicity is large as well. This dispersion

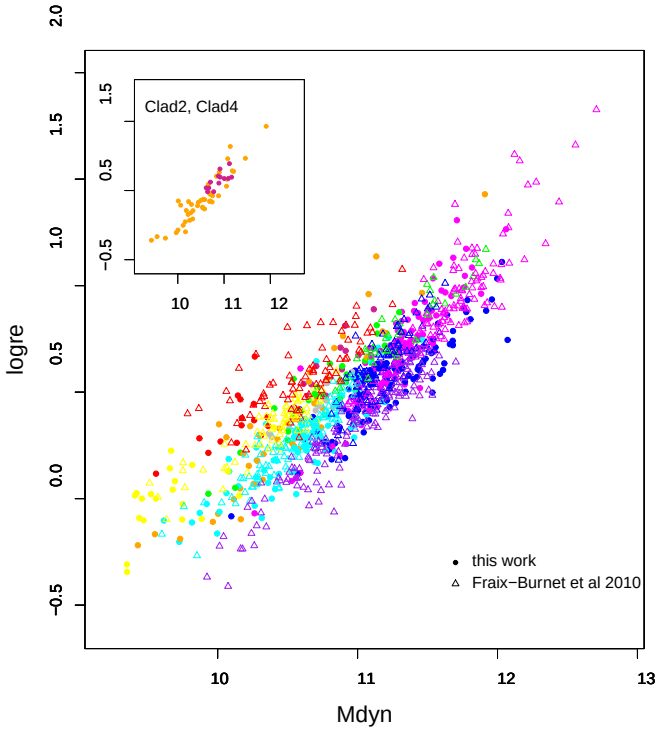

Fig. 9. Comparison between the partitioning obtained by cladistics and that of Fraix-Burnet et al. (2010). In the inset are plotted the groups Clad2 and Clad4, which are difficult to see in the main graph. See also FIg.C.4.

could certainly be explained by many factors, such as an extended horizontal branch, which can increase $H\beta$ (e.g. Greggio, 1997; Matković et al., 2009). Anyhow, the median for each group is nearly perfectly aligned between the two axes for age and metallicity, ordered as in Fig. 3, except for Clad2 and Clad7, which depart from the main alignment. Indeed this kind of plot merely tells us that age, metallicity, and $H\beta$ of galaxies evolve with time on average.

But we find no correlation between age and metallicity *within* individual groups. This is best shown in Fig. 6 (right), which plots $H\beta$ vs $[MgbFe]'$, which is a better indicator of metallicity than $Mg\,b$. It is striking that the elongated inertial ellipses for Clad3, Clad5 and Clad7 are well aligned with the $H\beta$ axis, with relatively little spread in metallicity, the one for Clad1 is slightly inclined and the one for Clad2 is along the global trend. Since the other ellipses are very round, Clad1 and Clad2 are the sole groups that might show a barely significant correlation between $H\beta$ and metallicity indicator.

The wider range in $H\beta$ for the less diversified groups, especially Clad1 and Clad3, appears clearly in Fig. 6 and can also be seen in Fig. 4. It probably corresponds to the well known wider range in age for the low-mass objects (e.g. Matković et al., 2009). Our interpretation is not that the low-mass galaxies have had a longer star formation history. Rather, such objects formed (appeared) over a longer time scale in the Universe's history and the older ones did not change much, apart from the ageing of stars. On average, larger galaxies necessarily took more time to assemble and complexify (diversify), so that it is very unlikely to find young and very diversified galaxies. However, the notion of galaxy age must be questioned.

Figure 6 illustrates the fundamental difference between diversification and age. The Clad1 group, assumed to be ancestral because of its low metallicity, appears as the youngest on average according to stellar evolution models. If a galaxy is to resemble the most pristine objects, it should not have been transformed too much even by secular evolution. This is why the most pristine objects are necessarily relatively young and metal-poor. Conversely, the most diversified objects have a higher average stellar age, which gives no information on the epoch of the transforming events that gave them their observed properties. As a consequence, old galaxies are not obvious ancestors. Also, the spread in age within each group is generally large and overlaps the spread in age of the other groups. Age is thus not a good landmark of evolution. From the point of view of astrocladistics, the so-called downsizing effect results from a confusion between age and level of diversification (Fraix-Burnet et al., 2006b,c, 2009).

So the age of a galaxy or a group of galaxies is probably not so important and may even be meaningless (Serra & Trager , 2007). We even find this term misleading, and suggest to replace it by "average stellar age". The diversification state should be used instead, as it reflects the actual assembly history of a galaxy.

7.3. Correlations within groups ("specific correlations")

As previously seen (Sect. 7.1 and Fig. 5), the two scatterplots log($diam$) vs K_{abs} and $Mg\,b$ vs $[MgbFe]'$ show both a global evolutionary correlation and correlations within the groups (which we will call "specific correlations"). Four other scatterplots only show specific correlations, the global correlation being less obvious and/or more dispersed : Bri_e and K_{abs} vs $\log r_e$, $\log(M_{dyn})$ vs $\log r_e$, $[MgbFe]'$ vs $Fe5270$. These six specific correlations are shown in Fig. 7.

Diversification within each group is determined by the structure of the tree in Fig. 3. If we examine the evolution of the parameters involved in the correlations along each branch (thus each group) of the tree, we find that:

- K_{abs} increases slightly with diversification within Clad6 and Clad8,
- $[MgbFe]'$ might possibly increase in Clad8,
- $\log r_e$ might possibly decrease in Clad6,
- M_{dyn} decreases slightly in Clad6 and might possibly increase in Clad5 and Clad8.

This is clearly not enough to explain all the observed specific correlations with evolution within the groups. However, the difference between objects of a same group is weaker than for the whole sample and thus would require refined cladistic analyses with possibly additional descriptors. Let us examine in some details the scatterplots in Fig. 7.

The correlation, particularly tight and linear, between $Mg\,b$ and $[MgbFe]'$, holds also within each group. Since the first parameter is largely dependent on the α-elements, while the second is essentially independent of it (Thomas et al., 2011), these specific correlations can probably be explained, like the global one, by evolution. The correlations between $[MgbFe]'$ and $Fe5270$ or $Fe5335$ are more dispersed.

The correlation between log($diam$) and K_{abs} seems to be present within each group. This is however not a proof that it is causal since the bigger the more luminous can still be due to evolution within each group or to some other confounding parameter (Fraix-Burnet, 2011). In addition, all correlations, global or specific, are approximately similar, which suggests a same explanation for all.

The Kormendy relation (Bri_e vs $\log r_e$) clearly appears to depend on the group, and is much less dispersed for the most diversified groups. There is an "evolution" of the correlation curve following diversification, the different relations appearing stacked on each other. At first glance, galaxies are brighter when more diversified, but this is not so simple if we look at Clad7 and Clad8: galaxies from the first group are globally larger and fainter. The correlation is also more dispersed for Clad1 and Clad3.

The K_{abs} vs $\log r_e$ relation is quite dispersed, but there are slightly more convincing correlations for some groups, at least for the most diversified ones. For Clad1, there is little variation in $\log r_e$ so that there is no real correlation. The difference between this relation and the K_{abs} vs $\log(diam)$ one is striking.

The M_{dyn} vs $\log r_e$ relation is tight, with very clear correlations specific to each group and with relatively little overlap between them. This plot can be usefully compared to numerical simulations (e.g. Robertson et al., 2006) as done in Fraix-Burnet et al. (2010, see Sect. 7.4.1 and Sect. 8).

The $[MgbFe]'$ vs $Fe5270$ relation is quite dispersed, despite the square-root relation linking both parameters. Correlations can be easily seen within groups, and they appear generally different (in particular for the slope) from the global relation.

To summarize, there are several cases where the specific correlations are present in all or some of the groups, whether the global correlation exists or not, Why is it so?

If the correlation is present both globally and within groups, then we can guess that it is for the same reason. In the two cases here ($\log(diam)$ vs K_{abs} and $Mg\,b$ vs $[MgbFe]'$), the global correlation is evolutionary (Sect. 7.1) and since all correlations appear to have approximately the same slope, then the specific correlations should be evolutionary as well.

In the other cases where there is no obvious global correlation, the reason must be specific to the group, and probably different from one group to another. The correlations might be explained by a direct physical cause, or by a confounding parameter, which can still be evolution. Note that the confounding factor may depend on the group.

Anyhow, the origin of the correlations and their properties is quite complex. In the case of the M_{dyn} vs $\log r_e$ relation, numerical simulations show that it is determined by several variables involved in the assembly history, like the epoch of last merger, the level of dissipation, the number of accretion events, the impact parameters and so forth (Robertson et al., 2006). Thanks to a good diversity of simulated galaxy populations, Fraix-Burnet et al. (2010) were able to derive the history assembly of each group. The specific correlations are then explained by several drivers from the physical point of view, or by "cosmic variance" within each group from the observational point of view, or by confounding factors from a statistical point of view.

Consequently, the two scatterplots showing both global and specific correlations are probably driven by a dominant general evolutionary factor (like perhaps dynamical evolution for $\log(diam)$ vs K_{abs} and chemical evolution for $Mg\,b$ vs $[MgbFe]'$) affecting all galaxies of the sample, while the other ones have multiple and necessarily specific factors, like in the M_{dyn} vs $\log r_e$ relation. For instance, the importance of merger events applies only to those galaxies that have experienced such a catastrophic transforming process during their assembly history.

7.4. Fundamental planes

7.4.1. The Fundamental Plane of early-type galaxies

The well-known and intensively studied correlation between Bri_e, $\log\sigma$ and $\log r_e$ is called the fundamental plane. The first multivariate analysis of this relation was performed recently by Fraix-Burnet et al. (2010). The present sample and that of Fraix-Burnet et al. (2010), both at low redshift, have no galaxy in common and the parameters used for partitioning the sample into groups are different.

The present partitioning is in excellent agreement with the result by Fraix-Burnet et al. (2010) as illustrated in Fig. 8, which shows the projection onto the fundamental plane ($\log\sigma$ vs Bri_e) of the partitionings obtained by cladistics and cluster analysis in the present paper, and the partitioning obtained by Fraix-Burnet et al. (2010). . The structures within the fundamental plane found in Fraix-Burnet et al. (2010) are thus confirmed.

There is a good correspondence between the groups in the two studies as seen in Fig. 8 and in more detail in Fig. C.3: C1 includes Clad4 and a large part of Clad3, C4 $\simeq$ Clad6, C3, C5 and C6 are essentially included into Clad7, C7 $\simeq$ Clad8. The $\log r_e$ vs M_{dyn} diagram (Fig. 9 and Fig. C.4) confirms these equivalences, pointing out that C1 overlaps both Clad3 and Clad4 and is distributed more like Clad3. There are however some differences.

Clad1 seems to occupy a region of the fundamental plane ($\log\sigma$ vs Bri_e) not very well covered in the sample used by Fraix-Burnet et al. (2010).

Clad2 has no equivalent in Fraix-Burnet et al. (2010) when projected onto the fundamental plane (Fig. 8). This group is also plotted separately in Fig. 9 to show that it follows the same correlation as the other groups, spanning nearly the full range in both $\log r_e$ and M_{dyn}.

Clad5 is also absent in Fraix-Burnet et al. (2010). Since it appears in the very center of both Fig. 8 and Fig. 9, we believe it is identified here because of the larger number of parameters used here amounts to a higher-resolution analysis. Its properties were also found to be quite similar to Clad4 (Sect. 6), so that Clad4 is quite different from C1 (see above).

Group C2 of Fraix-Burnet et al. (2010) is absent in the present partitioning. This is perhaps because the corresponding regions on the fundamental plane (Fig. 8) are not very populated in our sample, or because of the different sets of parameters used in the two studies.

To complete the comparison between the two studies, we have performed the same analysis as in Fraix-Burnet et al. (2010), thus using the same four parameters $\log\sigma$, $\log r_e$, Bri_e and Mg_2 (Appendix B). Naturally, since these four parameters are not all structuring for the present sample, the resulting tree and the corresponding partitioning are slightly less robust. However, the agreement is still quite good. The sample used in the present paper (Ogando et al., 2008) is globally at a lower redshift than the one used in Fraix-Burnet et al. (2010) (Hudson et al., 2001). This renders the determination of the distance, and thus $\log r_e$ less accurate (see Ogando et al., 2008). This could partly explain why $\log r_e$ has not been found as structuring or why the four-parameter result here is slightly less robust than in Fraix-Burnet et al. (2010).

We thus confirm the result by Fraix-Burnet et al. (2010) that there are structures within the fundamental plane. Since the organisation of the groups on this plane defines very similar evolutionary paths corresponding to a clear trend in diversification as defined by Fig. 3, we confirm that this global relation in the three-parameter space is mainly driven by diversification. For the sake of clarity, we stress that this evolutionary interpretation

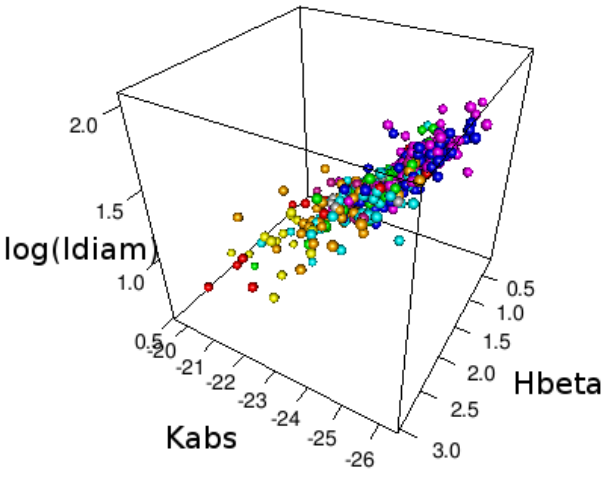

Fig. 10. Another "fundamental plane" in the parameter space defined by $H\beta$, K_{abs} and $\log(diam)$.

of the fundamental plane holds for the groups, not within the groups, since neither the present paper nor that by Fraix-Burnet et al. (2010) tackle this question. This would certainly deserve further specific studies because Fraix-Burnet et al. (2010) found that the tightness of the fundamental plane strongly depends on the group considered, the least diversified ones showing a very loose – if significant at all – correlation.

7.4.2. Other correlation planes

Concerning the fundamental plane, the groups follow one another along a path of diversification in the $\log r_e$ vs $\log \sigma$ (edge-on projection) scatterplot, whereas they are well distinguishable and distributed in no obvious order in the Bri_evs $\log r_e$ (face-on projection) scatterplot .

Hence, one can expect to find other "fundamental planes" by looking at the behaviour of groups on two scatterplots made with a set of three parameters. This is the case for $H\beta$, K_{abs} and $\log(diam)$ (Fig. 10). Obviously, the two latter parameters are reminiscent of Bri_e and $\log r_e$, while $H\beta$ is of a different nature from $\log \sigma$, so that this fundamental plane is not redundant with the classical one. Replacing $H\beta$ by D/B, one also obtains a nice correlation, and thus another fundamental plane.

It is interesting to note that the parameters of these fundamental planes might be easier to observe than for the classical one, but still allow the distance determination thanks to $\log(diam)$.

There might be other similar surfaces in higher-dimension parameter spaces but with more complex projections. Are they important to discover and are they more useful than the classical fundamental plane?

The answer to these questions is twofold. First, they are identical to the classical fundamental plane, in the sense that they are essentially evolutionary correlations driven by diversification (Fraix-Burnet, 2011). All are simply projections of the tree in a sub-parameter space. They are thus not more and not less informative. Second, they can be useful once the confounding factor (here evolution) has been taken into account. This is beyond the scope of the present paper.

8. The assembly history of early-type galaxies

In a previous paper (Fraix-Burnet et al., 2010) the assembly history of a completely different sample of early-type galaxies was uncovered by analysing its properties in the fundamental plane and in a mass-radius diagram with the help of numerical simulations from the literature. That sample was composed of galaxies in clusters, while the present one is composed of galaxies in the field, groups, or clusters, and has the advantage of more properties being documented. Emission-line galaxies were excluded from both samples. We repeat here the exercise on both sets of groups merged together following the correspondence detailed in Sect. 7.4.1.

- Clad1: these galaxies have low metallicity and in many respects look quite primitive, with a very low dynamical mass. They are less metallic and more primitive than the Clad3/C1 ones which were chosen as the most primeval group in Fraix-Burnet et al. (2010). Clad1 galaxies could be the remains of a simple assembly through a monolithic collapse with little dissipation and their somewhat discy nature probably requires significant feedback and few perturbations (Benson, 2010).
- Clad2: this group shows a steep correlation between $\log r_e$ and M_{dyn}. This could indicate some merger processes (e.g. Ciotti et al., 2007) but the galaxies are discy. However, they also show a high $\log \sigma$. This suggests that these galaxies are quite primitive objects like Clad1 but being more massive, that they underwent a more significant secular evolution, maybe like "pseudo-bulges" (Kormendy & Kennicutt, 2004; Benson, 2010).
- Clad3/C1: chosen as the most primeval group because of its low average Mg_2 in Fraix-Burnet et al. (2010), this group is here surpassed by the still lower Mg_2 of Clad1. The galaxies of C1 were found to possibly be the remains of a simple assembly through a monolithic collapse with little dissipation, and they were probably perturbed by interactions. We here rather propose accretion as the main perturbations because the Clad3 galaxies are small, not very much concentrated and have a low $\log \sigma$.
- C2: they were found in Fraix-Burnet et al. (2010) to be less massive and smaller than the ones of Clad3/C1, and they have a slightly higher Mg_2. They are also somewhat brighter. They could be the remains of wind stripping of some kind of more diversified objects because of a strong interaction.
- Clad4: this group is very similar to Clad2, but since the Clad4 galaxies are larger and have a higher M_{dyn}, they could more probably have been perturbed by a strong interaction that yielded a more massive central black hole.
- Clad5: being very similar to Clad4 objects, they could be these discy galaxies seen edge-on but this is statistically untenable since there are three times more Clad5 objects than Clad4 ones. Since they have a very low D/B, a better explanation is that Clad5 galaxies could be Clad4 perturbed members. These perturbations are probably mergers since they have lost the disciness of Clad4 objects, but to preserve similar properties, little gas should be involved, pointing to dry mergers as the most probable transforming events.
- Clad6/C4: three scenarii were proposed for C4 in Fraix-Burnet et al. (2010): these objects could simply be galaxies in which star formation has been continuous, or they could

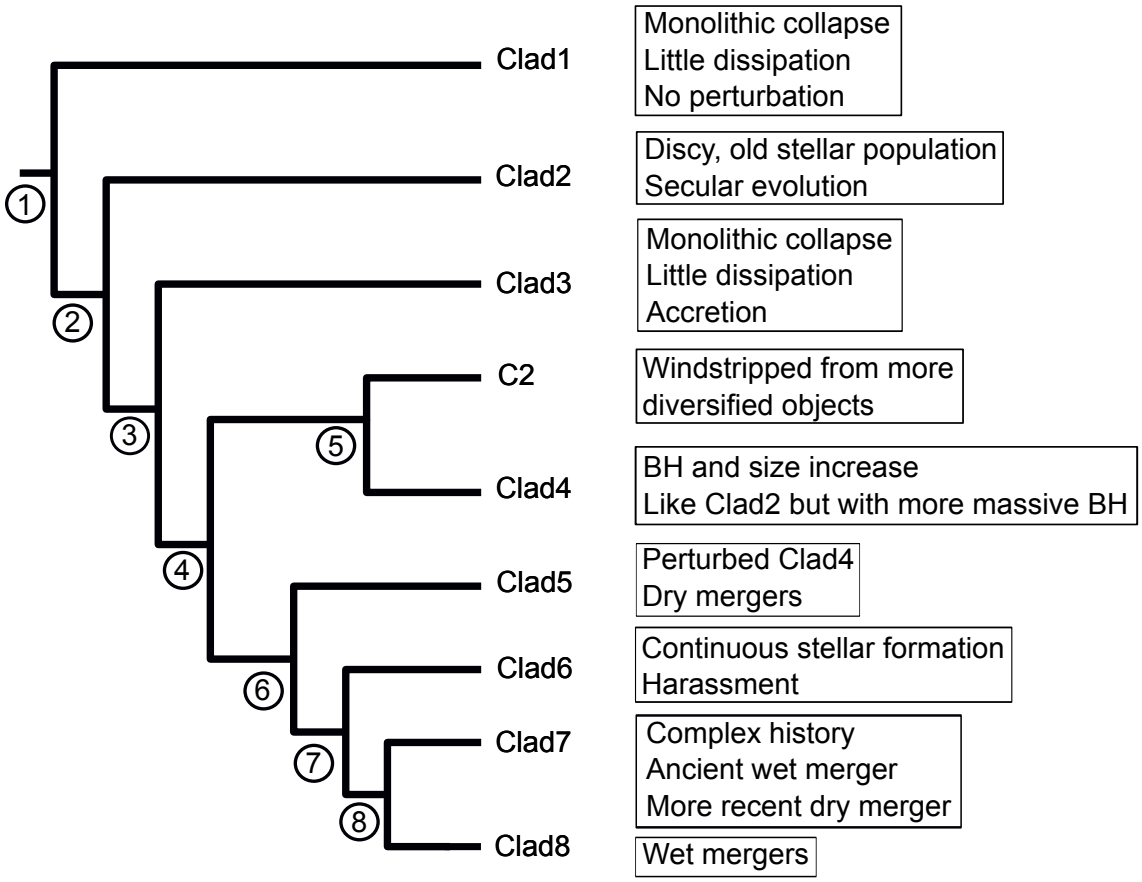

Fig. 11. Combination of the tree from this paper (Fig. 3) and the one in Fraix-Burnet et al. (2010), showing the proposed assembly history. Numbers indicated at nodes are referred to in the text.

be C1 galaxies in which the initially richer gas has not been swept out, or they could also be the remnants of several minor mergers and accretion. The Clad6/C4 galaxies have unexpectedly low M_{dyn}, Bri_e and $\log r_e$, and to a lesser extent a low $\log(diam)$. Otherwise they do not look odd, so that a continuous star formation with little external perturbations could also be a reasonable explanation. However, we find that they are very similar to Clad2, which we proposed to have undergone a significant secular evolution, but with a much lower D/B. This suggests that many interactions, such harassment (Moore et al., 1996), could be the culprit.

– Clad7: because Clad7 includes groups C3, C5 and C6, their history might be complex according to Fraix-Burnet et al. (2010), with many transforming events (accretions, minor mergers, together with more or less dissipational major mergers). They are probably the remains of both wet and dry mergers, the most recent ones being of the latter kind. The low $H\beta$ which would indicate that the last star formation event is relatively ancient, reinforces this interpretation. They could represent a kind of end state of galaxy diversification.
– Clad8/C7: C7 galaxies were found to be small and very metallic with a high surface brightness, and they define a tight FP. They seemed to be associated with the remains of a dissipative (wet) merger, with very little or no dry mergers. They could also have formed through minor mergers and accretions but the tight FP favors the dissipative wet-merger scenario. The low D/B found here for Clad8/C7 tends to confirm this conclusion. We believe that they could well define another possible end state for galaxy diversification.

We summarize the above histories on a single cladogram (Fig. 11) combining the trees obtained in the two studies.

The best way to interpret the evolutionary scenario depicted in this cladogram is to identify, for each node of the tree, a particular transforming event that could characterize all groups related by branches and sub-branches starting from this node (Fraix-Burnet et al., 2006c). The sequence of nodes downwards starting at the upper left of the tree thus defines a sequence of "innovations" that occurred in a common ancestor and were transmitted to all its descendant species. In principle, these innovations are properties of the galaxies that remain as imprints transmitted through subsequent transforming processes. Here, we consider the transforming events as innovations since they are the origin of the modifcations of the properties of galaxies. However, it should be kept in mind that parallel evolutions (events occuring independently on two different lineages), convergences (different pathways leading to the same parameter value) and reversals (backward parameter evolution) are probably present, making this exercise currently quite tentative. These behaviours (called homoplasies) are supposedly not too numerous here since the parcimony optimisation of the cladistic analysis minimizes these kinds of parameter evolutions. We do not discuss them in this very first attempt, especially because the properties at hand are few.

We attempt to identify these innovations on Fig. 11 using the possible histories of each group identified previously. They appear on the tree when they first occur in the history of the Universe. It is thus expected that the most basic transforming events occur very early and the more complex ones later, making

the sequence of "innovations" along the tree a representation of the so-called cosmic evolution.

Finally, let us remind the reader that if the level of diversity goes along the vertical axis on Fig. 11, the horizontal axis, especially the branch lengths, has no particular meaning in this representation. It is important to keep in mind that we are dealing with present-day galaxies and properties, not those that prevailed at the time of the transforming event indicated at the node from which a given branch emerges. Simply speaking, this means that the galaxies of, say Clad1, are present-day galaxies that passively evolved from a less diversified initial state than those of Clad6 or Clad7.

The nodes are identified by numbers on Fig. 11. The corresponding proposed events are the following:

1. collapse
2. secular evolution
3. accretion
4. interaction
5. "gaseous" interaction
6. dry merger
7. harassment
8. wet mergers

The first transforming event (node 1), which marked the history of all the galaxies, is most likely monolithic collapse. This is probably the simplest process to form a self-gravitating ensemble of gas, stars and dust that we can call "galaxy". The galaxies of Clad1 evolved passively after exhausting their gas reservoir.

The next three events (nodes 2-4) must have been gentle, since the discy morphology is well preserved until Clad4 and it is now well established that minor mergers generally preserve the structure of discy galaxies, while mergers of galaxies of comparable masses generally do not preserve it. We first have (node 2) the secular evolution, defined as the evolution of a galaxy in isolation, that is expectedly very much frequent and can modify significantly the structure and properties of galaxies. Then, for node 3, accretion must be invoked to increase the masses of galaxies. A more complex event follows, interaction, an external perturbation which, with the wealth of possible impact parameters and galaxy properties, is probably the main driver of galaxy diversity, especially during the first Gyr of the Universe (Benson, 2010).

For node 5 between C2 and Clad4, interaction involving gas must be invoked to strip the gas in C2 galaxies by ram pressure or to feed the central black hole in Clad4 objects.

Mergers must be advocated at node 6 since more diversified groups have lost their disciness. For Clad5 galaxies, the most probable transforming event is a major merger without much gas (dry mergers).

Important star formation must have occured for Clad6 galaxies, and several properties point to repeated perturbations, so that harassment is a good candidate for node 7 (Moore et al., 1996). Harassment is the cumulative effect of high-speed galaxy encounters, which heats the disc ($\log\sigma$ increases) and favors gas inflow to the galaxy center. This kind of transforming events plays on a longer timescale than the dry merger at node 6 which explains why it appears "later" on the cladogram.

The two most diversified groups, Clad7 and Clad8, are found to have a complex history certainly including wet mergers (node 8).

Many associated processes, like feedback and quenching of star formation (e.g. Bundy et al., 2006; Benson, 2010) are not proposed here because we concentrate on the more generic events. A significant difficulty of such an exercise is to identify some properties with transforming events which are very complex, involving diverse impact parameters and various chemical, physical and dynamical processes. We believe it is somewhat illusory to associate a particular feature to any such events, and only statistical analyses of simulated cases could provide average properties that can be compared to statistical analyses of real objects like the one we have performed.

9. Conclusion

In this paper, we have used several multivariate tools, first to select the most discriminant parameters from the 25 initially available for the sample of 424 fully documented galaxies of Ogando et al. (2008), and second to partition the sample into groups. The three partitioning methods yield similar number and composition for the groups, considering that some fuzziness is expected.

Our first result is that among the initial 23 quantitative parameters available in this study, only six are structuring and actually yield a relatively robust partitioning for this sample. Among the ten Lick indices, only two ($[MgbFe]'$ and NaD) are structuring, together with $\log\sigma$, Bri_e, $OIII$ and D/B.

The global evolutionary scenario found by astrocladistics gives a very sensible result: globally galaxies tend to become more metallic, more luminous (more massive) and larger with increasing diversification. At the same time, they acquire a larger central velocity dispersion, often related to the mass increase, and NaD also increases along with mass and velocity dispersion, as expected. These are global statistical trends explained by general basic physical and chemical processes as a function of time since the Big Bang.

As a consequence, the many properties of galaxies that are bound to evolve on average with galaxy diversification explain the several evolutionary correlations found in this paper. In particular, we confirm the evolutionary nature of the Mg_2 vs $\log\sigma$ correlation found by Fraix-Burnet et al. (2010) on a different sample and using a different set of parameters. Rather interestingly, we also find some correlations that are specific to some groups only. These can be explained either by a direct physical cause or by a confounding factor specific to some groups (like the epoch of last merger, the level of dissipation, the number of accretion events, the impact parameters, or the number of mergers).

One of the most important results of our work is that the structures defined by the partitioning, when projected onto the fundamental plane ($\log\sigma$ vs Bri_e) or on the $\log r_e$ vs M_{dyn} diagram, are very similar to those found by Fraix-Burnet et al. (2010) on a totally distinct sample with different parameters.

The fundamental plane of early-type galaxies appears to be very probably generated by diversification. In support of this, we also find another "fundamental plane", a three-dimensional correlation between $H\beta$, K_{abs} and $\log(diam)$. Basically, all scatter plots are simply projections into a sub-parameter space of the partitioning established in the six-parameter space. Thus there is less information in these scatter plots or "fundamental planes" than in the multivariate partitioning and the evolutionary tree obtained with cladistics.

Another important result is that six parameters – not fewer – are needed to describe the diversity of this sample. The three parameters of the fundamental plane ($\log\sigma$, Bri_e and $\log r_e$), plus the index Mg_2, do not yield as robust a partitioning here, although they did in a previous study on a distinct sample (Fraix-Burnet et al., 2010). We argue that there is no contradiction here, essentially because three structuring parameters extracted in the present paper are present in our previous study ($\log\sigma$, Bri_e, Mg_2

being replaced by $[MgbFe]'$). The multivariate analyses generally depend on the objects in the sample and on their initial set of descriptors, both of which are different in the two studies. Nevertheless, as a consequence, similar analyses will have to be conducted on other samples with other descriptors, since our fairly small sample of nearby galaxies cannot represent the diversity of galaxies throughout the Universe, and the available parameters are here restricted to the visible domain.

We have combined the results of Fraix-Burnet et al. (2010) and the ones in the present paper on a single cladogram, showing a possible assembly history for each group. From this cladogram, we try to identify the transforming events that are at the origin of galaxy diversification. The transforming events indicated as "innovations" are tentative, because the information at hand is insufficient to identify them with certainty. These proposed events show that the use of sophisticated statistical tools yields a very sensible classification. Figure 11 is the basis of an *explanatory* classification which links the objects to the fundamental transforming processes, i.e. to the physics, rather than a *descriptive* classification like most current classifications of galaxies. In this respect, it should be remarked that the Edwin Hubble classification is of the latter type, using morphology, while his tuning fork diagram (often called sequence nowadays) is explanatory since it gives the links between the classes. Nearly one century later, we know that galaxies are much more than just morphology, so that we need to generalize the Hubble diagram to a multivariate picture of galaxy diversification. Figure 11 found by cladistics is one step in this direction.

Hence, increasing both the sample size and the number of descriptors is an absolute requirement. The six-parameter space needed to describe the diversity of the sample of the present paper is probably a minimum space because of the complexity of galaxies and of their assembly history. The nature of the structuring parameters might also change with the input of more observables. Also, the number of groups and their boundaries will certainly change. This is a double quest: classifying galaxies in objective, evolutionary and intelligible groups, and finding the parameter space in which these groups are identified. This quest is necessarily progressive, and will probably never ends. However, one can hope a convergence to occur sometimes.

A limitation of the present work is that cladistics cannot be applied directly to very large samples as the necessary computer time would be excessive. However, once the most discriminant parameters are identified, it is possible to repeat the cladistic analysis on many subsamples, and subsequently combine the trees to define classes of galaxies. The ultimate goal is to gather the huge number of galaxies in the Universe into a tractable number of groups and establish the corresponding evolutionary relationships.

Acknowledgements. This research has made use of the HyperLeda database (http://leda.univ-lyon1.fr) and of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

References

Alonso, M. V., Bernardi, M., da Costa, L. N., et al. 2003, AJ, 125, 2307
Babu, G.J., Chattopadhyay, T., Chattopadhyay, A.K. & Mondal, S. 2009, ApJ, 700, 1768
Benson, A. J. 2010, Phys. Rep., 495, 33
Bernardi, M., Alonso, M. V., da Costa, L. N., Willmer, C. N. A., Wegner, G., Pellegrini, P. S., Rité, C. & Maia, M. A. G. 2002, AJ, 123, 2159
Bundy, K., Ellis, R. S., Conselice, C. J., Taylor, J. E., Cooper, M. C., Willmer, C. N. A., Weiner, B. J., Coil, A. L., Noeske, K. G. & Eisenhardt, P. R. M., 2006, ApJ, 651, 120
Burstein, D., Faber, S. M., Gaskell, C. M., & Krumm, N. 1984, ApJ, 287, 586
Cabanac, R. A., de Lapparent, A., & Hickson, P. 2002, Astronomy & Astrophysics, 389, 1090
Cappellari, M., Emsellem, E., Krajnović, D., McDermid, R. M., Serra, P., Alatalo, K., Blitz, L., Bois, M., Bournaud, F., Bureau, M., Davies, R. L., Davis, T. A., de Zeeuw, P. T., Khochfar, S., Kuntschner, H., Lablanche, P.-Y., Morganti, R., Naab, T., Oosterloo, T., Sarzi, M., Scott, N., Weijmans, A.-M. & Young, L. M. 2011 MNRAS 416, 1680
Chattopadhyay, A., Chattopadhyay, T., Davoust, E., Mondal, S., & Sharina, M. 2009a, Astrophysical Journal, 705, 1533
Chattopadhyay, T., Babu, J., Chattopadhyay, A., & Mondal, S. 2009b, Astrophysical Journal, 700, 1768
Chattopadhyay, T. & Chattopadhyay, A. 2006, The Astronomical Journal, 131, 24522468
Chattopadhyay, T. & Chattopadhyay, A. 2007, Astronomy & Astrophysics, 472, 131
Chattopadhyay, T., Misra, R., Naskar, M., & Chattopadhyay, A. 2007, Astrophysical Journal, 667, 1017
Chattopadhyay, T., Mondal, S., & Chattopadhyay, A. 2008, Astrophysical Journal, 683, 172
Chattopadhyay, T.,Sharina, M. & Karmakar, P. 2010, Astrophysical Journal, 724, 678
Ciotti L., Lanzoni B., Volonteri M., 2007, ApJ, 658, 65 (arXiv:astro-ph/0611328)
Ellis, S. C., Driver, S. P., Allen, P. D., Liske, J.b Bland-Hawthorn, J., & De Propris, R. 2005, Monthly Notices of the Royal Astronomical Society, 363, 1257
Fraix-Burnet, D. 2011, Monthly Notices of the Royal Astronomical Society: Letters, 416, L36
Fraix-Burnet, D., Choler, P., & Douzery, E. 2006a, Astronomy and Astrophysics, 455, 845, 13 pages 5 figures with 3 online only.
Fraix-Burnet, D., Choler, P., Douzery, E., & Verhamme, A. 2006b, Journal of Classification, 23, 31, 16 pages, 6 figures
Fraix-Burnet, D., Davoust, E., & Charbonnel, C. 2009, MNRAS, 398, 1706
Fraix-Burnet, D., Douzery, E., Choler, P., & Verhamme, A. 2006c, Journal of Classification, 23, 57, 14 pages, 4 figures
Fraix-Burnet, D., Dugué, M., Chattopadhyay, T., Chattopadhyay, A. K., & Davoust, E. 2010, MNRAS, 407, 2207
Gonzáles J., 1993, PhD thesis, University of California
Greggio, L. 1997, MNRAS, 285, 151
Hopkins, P. F., Cox, T. J., & Hernquist, L. 2008, ApJ, 689, 17
Hudson, M. J., Lucey, J. R., Smith, R. J., Schlegel, D. J., & Davies, R. L. 2001, MNRAS, 327, 265
Karachentsev I. D., Sharina M. E., Dolphin A. E., Grebel E. K., Geisler D., Guhathakurta P., Hodge P. W., Karachentseva V. E., Sarajedini A. & Seitzer P. 2002, A&A, 385, 21
Kormendy, J. & Kennicutt, Jr., R. C. 2004, ARA&A, 42, 603
Kuntschner, H., Emsellem, E., Bacon, R., et al. 2010, MNRAS, 408, 97
MacQueen,J. 1967, Fifth Berkeley Symp.Math.Statist.Prob., 1, 281
Maddison, W. P. & Maddison, D. R. 2004, Mesquite: a Modular System for Evolutionary Analysis
Matković, A., Guzmán, R., Sánchez-Blázquez, P., et al. 2009, ApJ, 691, 1862
Mei S., Blakeslee J. P., Côté P., Tonry J. L., West M. J., Ferrarese L., Jordán A., Peng E. W., Anthony A. & Merritt D. 2007, ApJ, 655, 144 (arXiv:0702.510)
Milligan, G.W. 1980, Psychometrika, 45, 325
Moore, B., Katz, N. , Lake, G., Dressler, A. & Oemler, A., 1996, Nature, 379, 613
Murtagh, F., & Heck, A. 1987, Multivariate Data Analysis (Dordrecht: Reidel).
Nigoche-Netro, A., Aguerri, J. A. L., Lagos, P., et al. 2011, A&A
Ogando, R. L. C., Maia, M. A. G., Pellegrini, P. S., & da Costa, L. N. 2008, AJ, 135, 2424
Perrett K. M., Hanes D. A., Butterworth S. T., Kavelaars Jj, Geisler D. & Harris W. E. 1997 AJ, 113, 895 (arXiv:9611052)
Recio-Blanco, A., Aparicio, A., Piotto, G., De Angeli, F., & Djorgovski, S. 2006, Astronomy & Astrophysics, 452
Robertson, B., Cox, T. J., Hernquist, L., et al. 2006, ApJ, 641, 21
Sánchez Almeida, J., Aguerri, J. A. L., Muñoz-Tuñón, C., & de Vicente, A. 2010, ApJ, 714, 487
Serra P. & Trager S. C. MNRAS, 374, 769 (arXiv:astro-ph/0610.343)
Sugar, A. & James, G. 2003, JASA, 98, 750
Swofford, D. L. 2003, PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)
Thomas, D., Maraston, C., & Bender, R. 2003, MNRAS, 339, 897
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
Thomas, D., Maraston, C., & Johansson, J. 2011, MNRAS, 412, 2183
Thuillard, M. 2007, Evolutionary Bioinformatics, 3, 267
Thuillard, M. 2008, Evolutionary Bioinformatics, 4, 237
Thuillard, M. & Fraix-Burnet, D. 2009, Evolutionary Bioinformatics, 5, 33

Thuillard M. & Moulton V. 2011, Bioinformatics and Computational Biology, 9, 453
Tripicco, M. J. & Bell, R. A. 1995, AJ, 110, 3035
Whitmore B.C. 1984, ApJ, 278, 61
Worthey G., Faber S. M., González J. J. & Burstein D. 1994, ApJS, 94, 687
Worthey, G. & Ottaviani, D. L. 1997, ApJS, 111, 377

Appendix A: Methods

A.1. Principal component analysis

Principal Component Analysis (PCA) is well-known to astronomers. It is not a partitioning method, its aim is to reduce the dimensionality of the parameter space. From the correlation matrix, PCA builds eigenvectors (the principal components), which are orthogonal and linear combinations of the physical parameters. These eigenvectors usually have no physical meaning. In general, most of the variance of the sample can be represented with only a few principal components (those having an eigenvalue greater than 1). They thus give a simpler representation of the data by eliminating the correlations between physical parameters. Strongly correlated parameters are gathered in the same eigenvector, and the most influent parameters (with respect to variance) are the ones with the highest coefficient (loading) in each eigenvector. The physical interpretation must be done back in the real parameter space.

PCA is thus very efficient at reducing the parameter space to supposedly uncorrelated components and helps in detecting the most structuring or discriminating parameters. The number of significant eigenvectors gives an idea of the number of parameters necessary to describe the sample. Principal components can also be used for subsequent cluster or cladistic analyses.

There is however a caveat to be kept in mind. PCA eliminates all correlations, be they causal or not. It is extremely useful to remove redundancies, as well as physical correlations between two parameters indicating the same underlying process. But PCA also removes evolutionary correlations (also called "spurious" or confounding in statistics, Fraix-Burnet, 2011), for instance between two parameters that are independent but vary with time. The $\log\sigma$-Mg_2 correlation for early-type galaxies (see Fraix-Burnet et al., 2010) is a good example. Such independent evolutions are lost through the PCA reduction of dimensionality.

A.2. Minimum Contradiction Analysis

Partitioning objects is to put some order. In some cases, i.e. in hierarchical clustering or in cladistics, the arrangement of the objects can be represented on a tree. A tree is a graph representing the objects as the leaves with a unique path between any two vertices. A bifurcating tree has all its internal vertices of degree at most 3 (at most 3 branches connect to any such vertex).

By indexing circularly all the leaves of a planar representation of a weighted binary tree, one obtains a perfect order, meaning that the corresponding ordered distance matrix fulfills all Kalmanson inequalities. Generally speaking the Kalmanson inequalities are fulfilled if the ordered distance matrix corresponds to a weighted binary tree or a superposition of binary trees (Thuillard & Moulton, 2011). The difference between the perfect order and the order one obtains with a given dataset is called the contradiction. The minimum contradiction is the best order one can get.

The Minimum Contradiction Analysis (Thuillard, 2007, 2008, MCA,) finds this best order. It is a powerful tool to explore whether the parameters can lead to a tree-like arrangement of the objects (Thuillard & Fraix-Burnet, 2009). Using the parameters that fulfill this property, the method then performs an optimisation of the order and provides groupings with an assessment of their robustness.

For taxa indexed according to a circular order the distance matrix

$$Y^n_{i,j} = \frac{1}{2}(d_{i,n} + d_{j,n} - d_{i,j})$$

fulfills the so-called Kalmanson inequalities (Kalmanson, 1975):

$$Y^n_{i,j} \geq Y^n_{i,k}, Y_{k,j}{}^n \geq Y^n_{k,i} \quad (i \leq j \leq k) \tag{A.1}$$

with $d_{i,j}$ the pairwise distance between taxon i and j. The matrix element $Y^n_{i,j}$ is the distance between a reference node n and the path i-j. The diagonal elements $Y^n_{i,i} = d_{i,n}$ correspond to the pairwise distance between the reference node n and the taxon i.

The contradiction on the order of the taxa can be defined as:

$$C = \sum_{k>j\geq i} \left(\max\left(\left(Y^n_{i,k} - Y^n_{i,j}\right), 0\right)\right)^2 + \sum_{k\geq j>i} \left(\max\left(\left(Y^n_{i,k} - Y^n_{j,k}\right), 0\right)\right)^2 \tag{A.2}$$

for any $i, j, k \neq n$.

The best order of a distance matrix is, by definition, the order minimizing the contradiction (Thuillard, 2007, 2008).

Thuillard & Fraix-Burnet (2009) showed that the perfect order is linked to the convexity of the variables in the parameter space, and is obtained for specific properties of the variables along the order. It is then possible to detect the structuring potentiality of the variables. This is exactly what is done in Sect. 3.2.

A.3. Cladistic Analysis

Cladistics seeks to establish evolutionary relationships between objects. It is a non-parametric character-based phylogenetic method, also called maximum parsimony method. It does not use distances, because there is no assumption about the metrics of the parameter space. Rather, the "characters" are traits, descriptors, observables, or properties, which can be given at least two states characterizing the evolutionary stage of the objects for that character. The use of this approach in astrophysics is known as astrocladistics (for details and applications see Fraix-Burnet et al., 2006b,c, 2009, 2010). Simply speaking, the characters here are the parameters, the (continuous) values of which are supposedly evolving with the level of diversification of the objects. The maximum parsimony algorithm looks for the simplest arrangement of objects on a bifurcating tree. The complexity of the arrangement is measured by the total number of "steps" (i.e. changes in all parameter values) along the tree.

The success of a cladistic analysis much depends on the behaviour of the parameters. In particular it is sensitive to redundancies, incompatibilities, too much variability (reversals), parallel or convergent evolutions. It is thus a very good tool to investigate whether a given set of parameters can lead to a robust and pertinent diversification scenario.

In the present study, we used the same kind of analysis as in our previous papers on astrocladistics. We discretized the parameters in 30 equal-width bins, which play the role of discrete evolutionary states. This choice of 30 bins is justified by a fair representation of diversity, a stability of the analysis in the sense that the result does not depend on the number of bins, and a bin width corresponding roughly to the typical order of magnitude of the uncertainties (i.e. 7%, see Fraix-Burnet et al., 2009).

Table A.1. Loadings on the eight principal components of the PCA analysis made on the set of 23 parameters (see Sect. 3.1) .

	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Comp8
Mgb	-0.9143	-0.122395	0.1594	-0.1593	-0.1215	0.04717	0.11541	0.03549
logs	-0.9067	-0.012224	0.0674	0.0950	-0.0443	0.02837	-0.05006	0.04723
Mg2	-0.8983	-0.158860	0.1190	-0.1121	-0.1455	0.05519	0.07249	-0.00744
mgbfe	-0.8879	-0.303358	0.0656	-0.1870	0.0543	-0.04251	0.00546	-0.04366
NaD	-0.8659	-0.099002	0.0527	0.0657	-0.0496	-0.00899	-0.00360	-0.03957
Mg1	-0.8470	-0.132117	0.1739	-0.0977	-0.1787	0.01250	0.06330	-0.05674
Kabs	0.7780	-0.357711	0.1878	-0.3577	-0.1884	0.00503	0.04219	-0.10201
mabs	0.7255	-0.359334	0.2095	-0.4395	-0.2075	0.06904	-0.06490	-0.11826
ldiam	-0.7248	0.409603	-0.2557	0.3327	0.1887	-0.09360	0.02929	0.06032
mgb.fe	-0.6742	0.245147	0.2938	-0.0296	-0.3977	0.17776	0.26559	0.15522
logre	-0.6512	0.570850	-0.3297	-0.2541	-0.0686	-0.01075	-0.01336	-0.01325
Fe5335	-0.5455	-0.406391	-0.1312	-0.2574	0.3127	-0.11438	-0.13735	-0.09773
hbeta	0.5389	-0.223947	-0.4974	0.2296	0.0104	0.18101	-0.01882	-0.00539
Fe5406	-0.4906	-0.465277	-0.1337	-0.0838	0.1044	-0.13418	0.01741	-0.11408
Fe5270	-0.4900	-0.522075	-0.1240	-0.1365	0.2660	-0.18686	-0.15191	-0.15975
H.K	-0.3065	0.117644	0.0945	0.2266	0.4477	0.48040	-0.20626	-0.28296
Fe5015	-0.2763	-0.550090	-0.5230	0.1094	-0.2198	0.17362	0.05076	-0.00969
D.B	0.2024	0.134962	-0.3041	-0.1834	0.2342	-0.15225	0.66215	-0.38279
B.R	-0.1338	0.363115	-0.2925	-0.4631	-0.1578	0.13444	-0.51784	-0.06306
Brie	-0.0673	0.614240	-0.4327	-0.4875	-0.0751	-0.08047	0.03262	-0.07633
Fe5709	0.0602	-0.172935	-0.1399	-0.3042	0.4027	-0.14195	0.06450	0.74089
OIII	-0.0311	-0.395042	-0.6303	0.1494	-0.3828	0.25483	0.08224	0.13941
J.H	-0.0175	-0.000856	-0.0977	0.3589	-0.3548	-0.69952	-0.23475	-0.11213

Table A.2. Fitness of parameters on the cladograms obtained for each subset as represented by the Rescaled Consistency Index (RCI).

Subset	Order from RCI	RCI
4cA	$D/B\ \log\sigma\ [MgbFe]'\ NaD$	0.102 0.086 0.086 0.075
5c	$D/B\ \log\sigma\ NaD\ [MgbFe]'\ Bri_e$	0.077 0.063 0.063 0.059 0.051
5cA	$[MgbFe]'\ Mg\,b\ \log\sigma\ D/B\ NaD$	0.098 0.090 0.080 0.075 0.066
6c	$\log\sigma\ D/B\ NaD\ [MgbFe]'\ Bri_e\ OIII$	0.059 0.055 0.055 0.050 0.040 0.039
6cA	$[MgbFe]'\ Mg\,b\ \log\sigma\ D/B\ NaD\ Bri_e$	0.076 0.073 0.061 0.060 0.054 0.043
7c	$D/B\ \log\sigma\ OIII\ NaD\ [MgbFe]'\ Fe5015\ Bri_e$	0.053 0.051 0.044 0.041 0.040 0.037 0.031
8c	$Mg\,b\ [MgbFe]'\ \log\sigma\ NaD\ D/B\ OIII\ Fe5015\ Bri_e$	0.055 0.052 0.050 0.044 0.041 0.038 0.033 0.030
10c	$[MgbFe]'\ Mg\,b\ NaD\ \log\sigma\ D/B\ Fe5270\ Bri_e\ B\text{-}R\ OIII\ Fe5709$	0.075 0.055 0.044 0.042 0.034 0.033 0.025 0.025 0.023 0.020

We also adopted the parsimony criterion, which consists in finding the simplest evolutionary scenario that can be represented on a tree. The maximum parsimony searches were performed using the heuristic algorithm implemented in the PAUP*4.0b10 (Swofford, 2003) package, with the Multi-Batch Paup Ratchet method (http://mathbio.sas.upenn.edu/mbpr). The results were interpreted with the help of the Mesquite software (Maddison & Maddison, 2004) and the R-package (used for graphics and statistical analyses).

Making cladistic analyses with different sets of parameters both helps find the most robust result and gives interesting information on the behaviour of the parameters themselves. The robustness of cladograms is always difficult to assess objectively, so we use a criterion similar to that of other statistical distance analyses: if a similar result is found by using different conditions or methods, then it can be considered as reasonably robust. There are four possible tests here:

1. occurrence of a branching pattern among most parsimonious trees: with so few parameters, many equally parsimonious trees are found, often arbitrarily limited to 1000. The majority-rule consensus of all of them yields a percentage of occurrence for each node. The higher this percentage, the higher the probability that this node is "robust".
2. agreement of branching patterns between sub-sample analyses, which can be called "internal consistency": by making analyses of several sets of arbitrarily selected sub-samples, we can check whether a given pattern is present on trees found with larger samples, including the full tree.
3. comparison between different sets of parameters: the result should preferably not depend too much on a single parameter. Adding or removing a parameter should not change drastically the tree.
4. comparison with cluster analysis: distance-based methods are totally independent, so an agreement can give a fair confidence in the result.

Since we have many more objects than parameters, a lot of "flying" objects are expected between different analyses, and the above tests should be done with statistics in mind. The first test is always positive in this study: percentages are higher than 70-75%, and most often they are above 95%. This is already an indication that some structure is present in the data. The other three tests are described below.

The full sample of 424 galaxies was divided into three subsamples with 105 objects each and a fourth one with 109 objects. It turns out that the first and fourth subsamples belong exclusively to clusters 1 and 3 respectively of the cluster analysis. The

diversity of the first subsample is less than for the other ones, so that the resulting tree is generally less resolved. The two first subsamples were also gathered to make a 210-object subsample, and the two last ones to make a 214-object subsample. Analyses were made with these six subsamples, as well as with the full sample. Then we estimated the internal consistency by comparing the seven trees two by two and by eye (with the help of the program cophyloplot in the R-package which connects a given object on the two trees).

This procedure was applied for each of the eight-parameter subsets given in Table 1. Subsets 5c, 5cA et 6cA and 3c show a rather good internal consistency, then 4c, 7c and 6c are fairly good, and finally 8c et 10c are not very good.

This already shows that the optimal number of parameters is around 5 or 6, at most 7. This is in excellent agreement with the PCA analysis (Sect. A.1).

If we compare the trees obtained with the full sample for the eight-parameter subsets, we find that subset 5c is very consistent with 6c, 7c and 8c. Also, 5c, 6c, 5cA and 6cA are in good mutual agreement, while this is not the case for 6c, 7c and 8c.

In Table A.2, we show for each tree the Rescaled Consistency Index (RCI) which measures the fitness of a parameter on the phylogeny depicted by the tree. The higher RCI is (indeed the closer to 1), the more structuring the parameter is. In other words, parameters with higher RCI are the most responsible for the tree structure. The absolute value depends on the number of objects and parameters, so it cannot be used to compare trees obtained with different data. Here, we can only use it to compare parameters for a given tree. In Table A.2, the parameters are ordered according to RCI.

When $Mg\,b$ and $[MgbFe]'$ are together in a subset, they dominate the shape of the tree (sets 5cA, 6cA, 8c and 10c), $\log\sigma$ and D/B being right after them. $Mg\,b$ and $[MgbFe]'$ are obviously redundant because they are very well correlated and are more or less the same measure. Hence they cannot be used simultaneously in the cladistic analysis, and the trees we find are more linear than the others. On the contrary, $\log\sigma$ and D/B are not at all correlated, but are always together, and dominate the tree shape when $Mg\,b$ is not present together with $[MgbFe]'$. Also, NaD is very structuring, and only roughly correlated with $\log\sigma$ and $[MgbFe]'$.

If we compare the clusters obtained with the clustering analysis, the agreement decreases roughly for 6c, 7c, 5A, 3c, 5c, 4cA, 8c, 10c, the winner being undoubtedly 6c. The corresponding tree with the groups is shown on Fig. 3.

A.4. Cluster Analysis

In the present study K-means partitioning algorithm of clustering has been considered following (MacQueen, 1967). This method constructs K clusters using a distance measure (here Euclidean). The data are classified into K groups around K centres, such that the distance of a member object of any particular cluster (group) from its centre is minimum compared to its distance from the centres of the remaining groups. The requirement for the algorithm is that each group must contain at least one object and each object must belong to exactly one group, so there are at most as many groups as there are objects. Partitioning methods are applied (Whitmore, 1984; Murtagh, 1987; Chattopadhyay & Chattopadhyay, 2006, 2007; Babu et al., 2009; Chattopadhyay et al., 2009a; Chattopadhyay et al., 2010) if one wants to classify the objects into K clusters where K is fixed. Cluster centres have been chosen on the basis of group average method, which makes the process almost robust (Milligan, 1980).

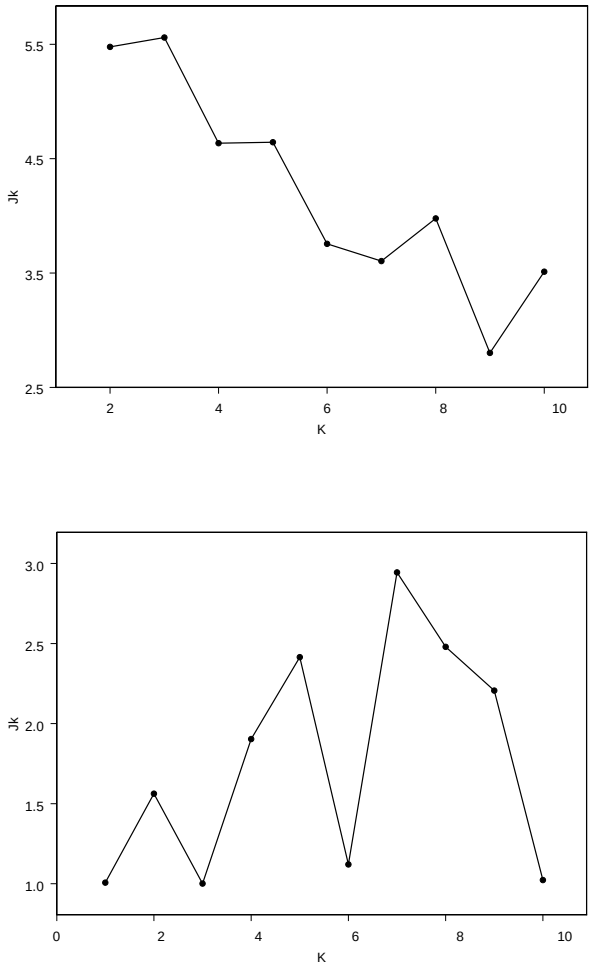

Fig. A.1. Plots showing the jumps as defined in Sect. A.4. Top: jumps for the PCA+CA analysis (Sect.4.1). Bottom: jumps for the cluster analysis with the six parameters (Sect.4.2).

For an optimum choice of K, the algorithm is run for K = 2, 3, 4, etc. For each value of K, the value of a distance measure d_K (called distortion) is computed as $d_K = (1/p)\ min_x E[(x_K - c_K)'(x_K - c_K)]$, which is defined as the distance of the x_K vector (values of the parameters) from the centre c_K where p is the order of the x_K vector. If d'_K is the estimate of d_K at the K^{th} point, then the optimum number of clusters is determined by the sharp jump in the curve $J_K = (d_K'^{-\mathbf{p/2}} - d_{K-1}'^{-\mathbf{p/2}})$ vs K (Sugar & James, 2003). The jumps as a function of K for our PCA+CA and CA analyses are shown on Fig. A.1.

Appendix B: Analysis with $\log\sigma$, $\log r_e$, Bri_e and Mg_2, and error bars

B.1. Analysis with $\log\sigma$, $\log r_e$, Bri_e and Mg_2

We have complemented the study presented in this paper with the analysis of our sample with the four parameters ($\log r_e$, $\log\sigma$, Bri_e and Mg_2) as in Fraix-Burnet et al. (2010). The same three multivariate techniques (cluster analysis, Miminimum Contradiction Analysis, and cladistics) as presented in Sect. 2.2 and Appendix A are used.

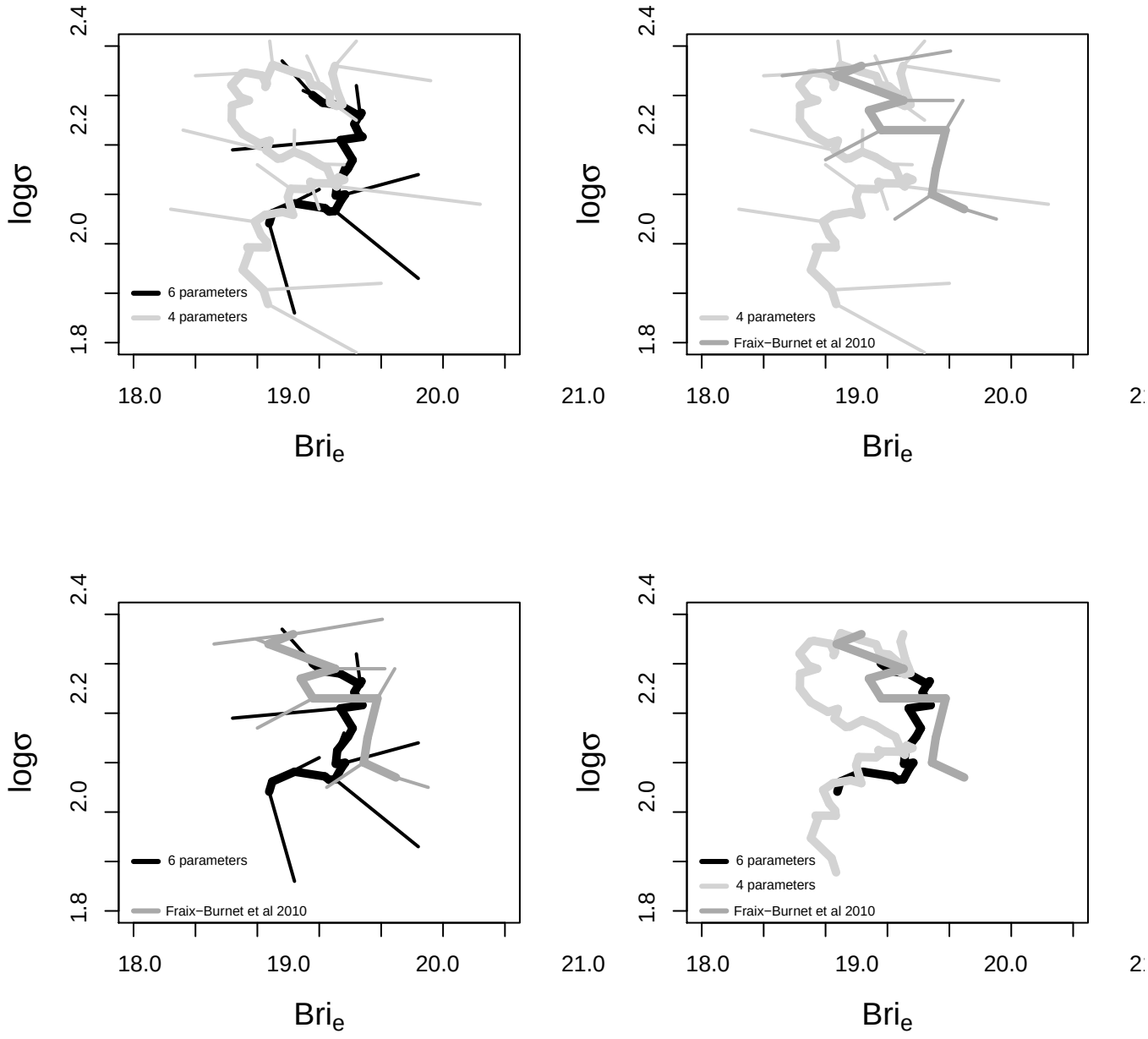

Fig. B.1. Projection of the trees onto the fundamental plane for three cases: the analysis of this Appendix B and the one by Fraix-Burnet et al. (2010) both using the four parameters of the fundamental plane, and the principal study of the present paper with six parameters. Thick lines represent the "trunk" of the trees, while the small branches relate the trunks to the mean of each group. For clarity, results are compared two by two, and only the trunks are shown for the three studies on the lower right diagram. These are evolutionary tracks in the sense of diversification, and not the path of evolution for a single galaxy.

The resulting tree is less structured (more galaxies lie on individual branches) than the one obtained in the present paper using six parameters. This can be explained by the fact that $\log r_e$ and Mg_2 have not been found to be structuring parameters for the considered sample. It is also less structured than in Fraix-Burnet et al. (2010) that uses the same four parameters. This probably can be explained by the problems in the determination of $\log r_e$.

To summarize the results, we show the projection of the three trees – the one obtained in this paper with six parameters, the one obtained here with four parameters and the one by Fraix-Burnet et al. (2010) – onto the fundamental plane ($\log \sigma$ vs Bri_e) without the data points (Fig. B.1). Globally, there are in good agreement and the groupings are consistent. However, the projected tree from the present Appendix departs from the other two on the top half. This is due to the fact that it is less structured than the other ones, so that instead of having one or two groups at this level, there is a sequence of single branches that makes the trunk of the tree to "follow" more closely individual objects.

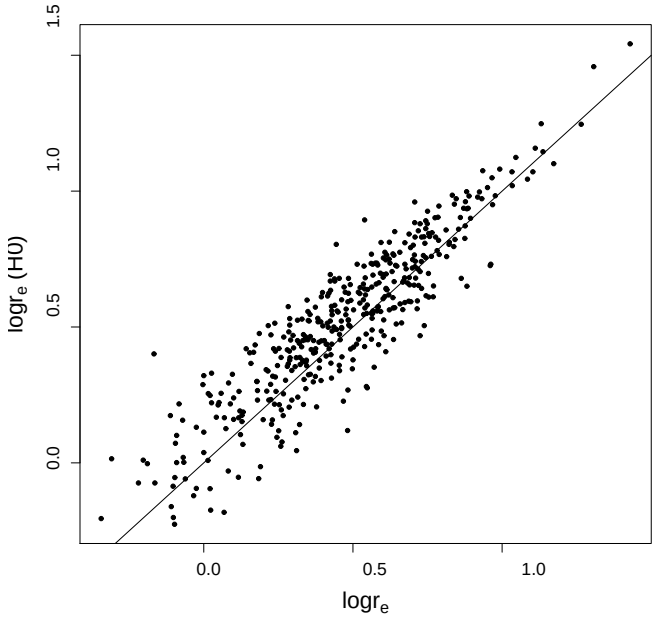

Fig. B.2. Correspondence between the effective radius as computed in two ways.

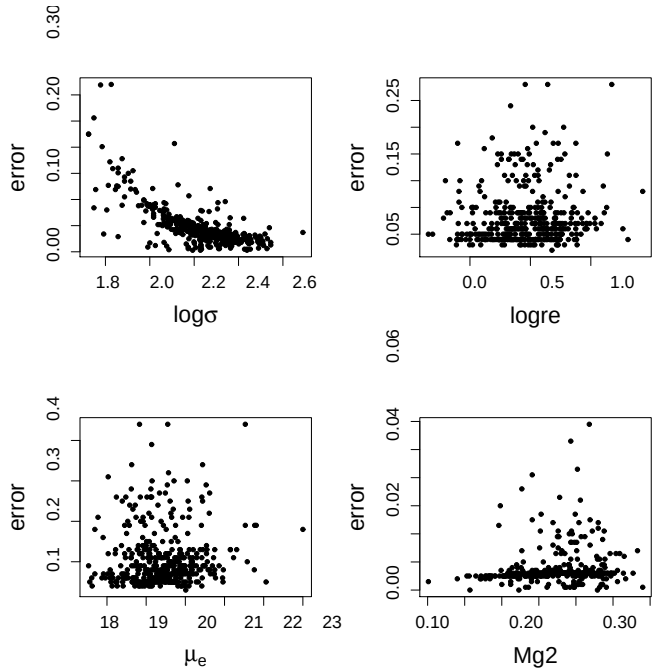

Fig. B.3. Errors for $\log\sigma$, Bri_e and $\log r_e$ (taken for the errors on D/B).

B.2. Influence of r_e and error bars on the partitioning

It turns out that the $\log r_e$ in our sample is recomputed through a statistical relation between the linear diameter of the galaxy (D_n) and its velocity dispersion (σ) which was determined in another paper (Bernardi et al., 2002). The reason given by Ogando et al. (2008) is that, due to the very low redshift of the galaxies in the sample, *the conversion of re in arcseconds to kpc needs a reliable determination of the galaxy distance (D). Considering just the redshift to calculate D, we may incur in error due to the peculiar motion of galaxies. Thus, we adopted D given by the D_n vs σ relation (Bernardi et al., 2002) to calculate r_e in kpc.* However, this relation has been obtained with some hypotheses (like identical properties of galaxies in several clusters) and introduces a dependence of $\log r_e$ (through D) on $\log\sigma$.

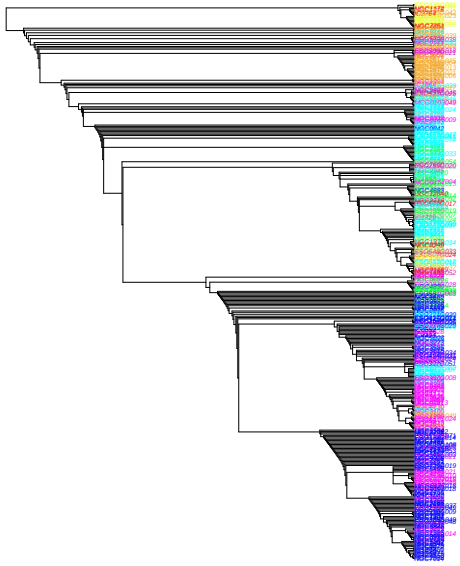

Fig. B.4. Most parsimonious tree found with cladistics taking uncertainties on the parameters into account. The colors correspond to the groups defined in Fig. 3.

The two radii (Fig. B.2) are quite well correlated, but the dispersion is relatively large. We have performed two cladistic analyses with the four parameters of the fundamental plane ($\log r_e$, $\log\sigma$, Bri_e and Mg_2) as above using the two determinations of the effective radius. The agreement between the two results is only fair. This can be explained by the relatively important discrepancy between the two different values of r_e (median difference of 10%). This however is similar to the uncertainty on $\log r_e$, but much larger than for the other parameters. In addition, the radius or dimension of galaxies does not appear as a structuring parameter in the study presented in this paper. Hence it is not so suprising that analyses using this parameter are not very stable.

We now consider the robustness of our clustering result for the six-parameter analysis when taking error measurements into account. This is statistically a very challenging task to assess this influence. However, cladistics can easily input error bars since the optimisation criterion in all analyses performed so far in astrocladistics use the parsimony criterion: among all the possible arrangements of the objects on trees, the simplest evolutionary scenario is retained. The parcimony is measured by using the number of "steps", that is the total number of changes in parameter values along all the branches of the tree. If a missing value or an uncertain one (given by a range of values) is included in the data matrix, all possible values are considered and the ones corresponding to the simplest tree is favored. This simply increases the number of possible cases to consider. Note that all possible values within the range alllowed by measurement uncertainties

are given the same weight wheras it is generally expected that the probability distribution is higher at the central value (ideally gaussian).

We have performed a cladistic analysis like for Fig. 3 using the error bars given in Ogando et al. (2008) and Alonso et al. (2003) for $\log\sigma$and Bri_e, and for D/B we take the error given for $\log r_e$ in Alonso et al. (2003), There errors are shown in Fig. B.3. For NaD, $[MgbFe]'$ and $OIII$ we take a face value of 10% which is the upper limit estimated by Ogando et al. (2008) for the Lick index values in general.

The resulting tree shown in Fig. B.4 is slightly less structured than the one in Fig. 3 but most groups are grossly preserved. Clad3 appears mixed with Clad1 and Clad5 with Clad6. In addition, Clad7 and Clad8 are somewhat split into each other. Interestingly, these behaviours are similar to those issued from the comparison with the partitioning from the cluster analysis. Also, the agreement is quite satisfactory given the high uncertainties on half of the parameters (the Lick indices), a face value given to these uncertainties, and the equal probability given to all values within the range of uncertainty.

This results shows that the cladistic analysis is relatively robust to measurement errors, equivalently to the comparison with different clustering methods.

Appendix C: Supplementary figures

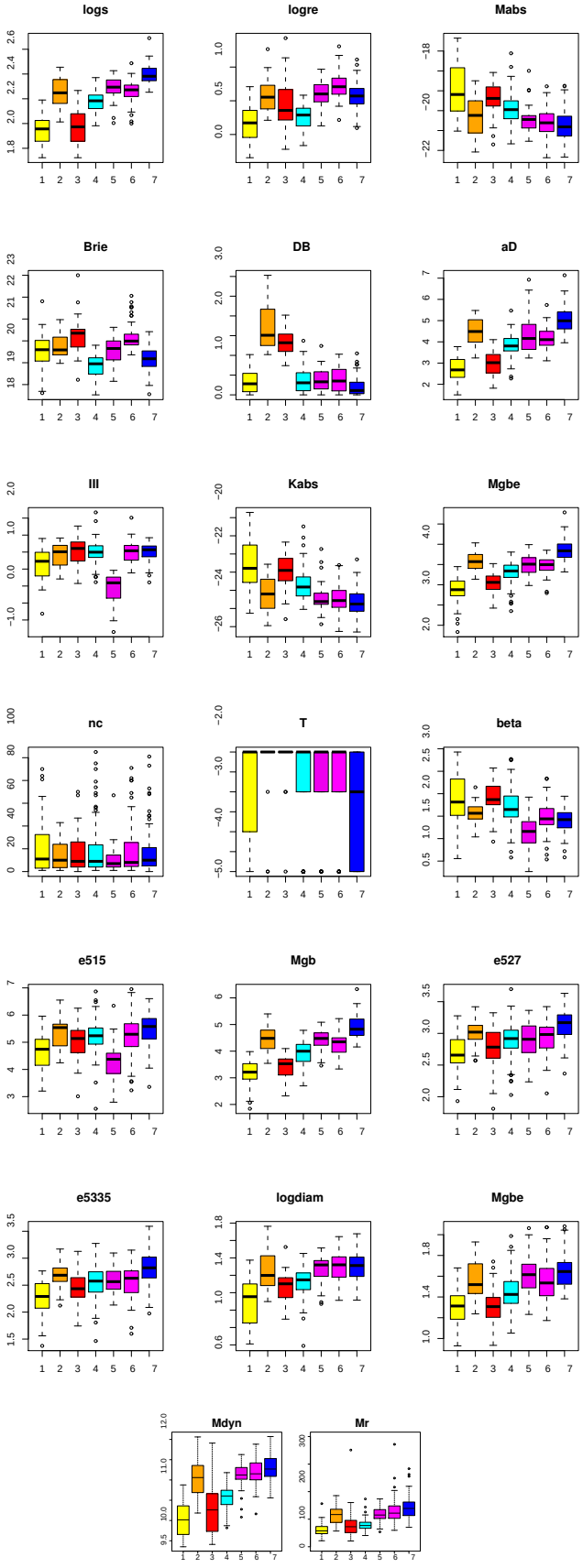

Fig. C.1. Same boxplots as in Fig. 4 but for the cluster partitioning. Colors are the one given in Fig. 2.

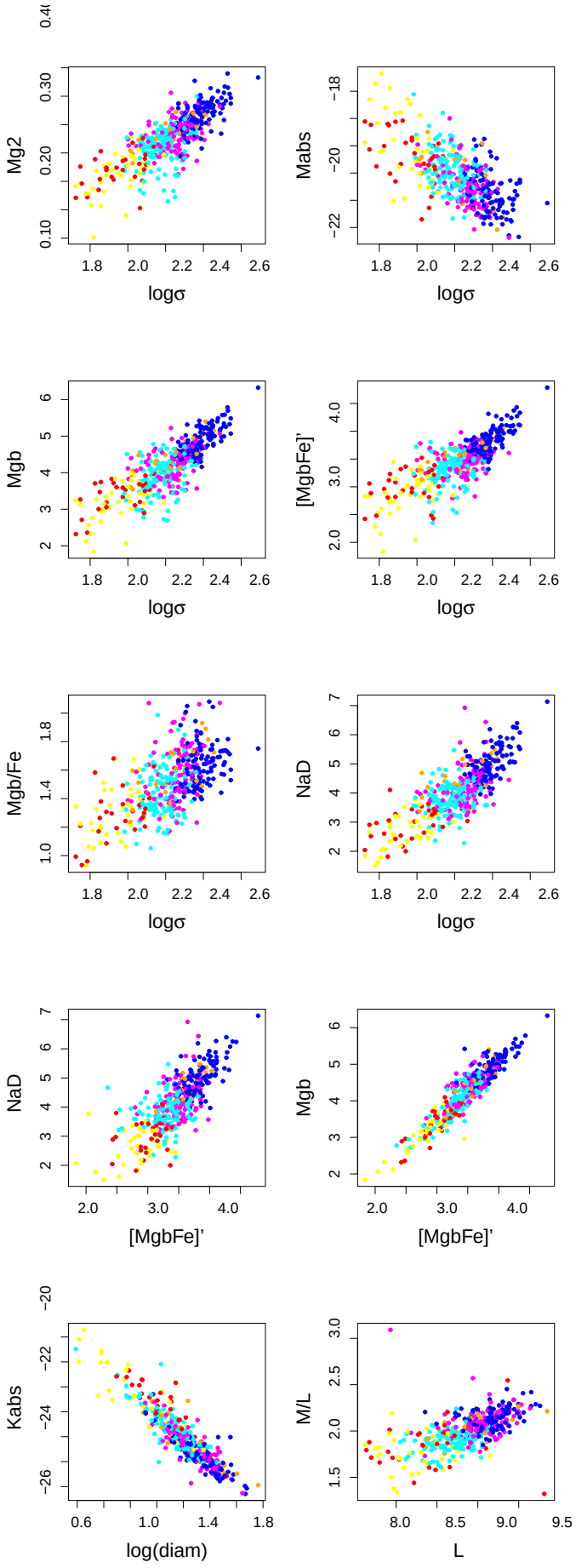

Fig. C.2. Scatter plots showing evolutionary correlations, like Fig. 5, but for the cluster partitioning. Colors are the same as in Fig. C.1.

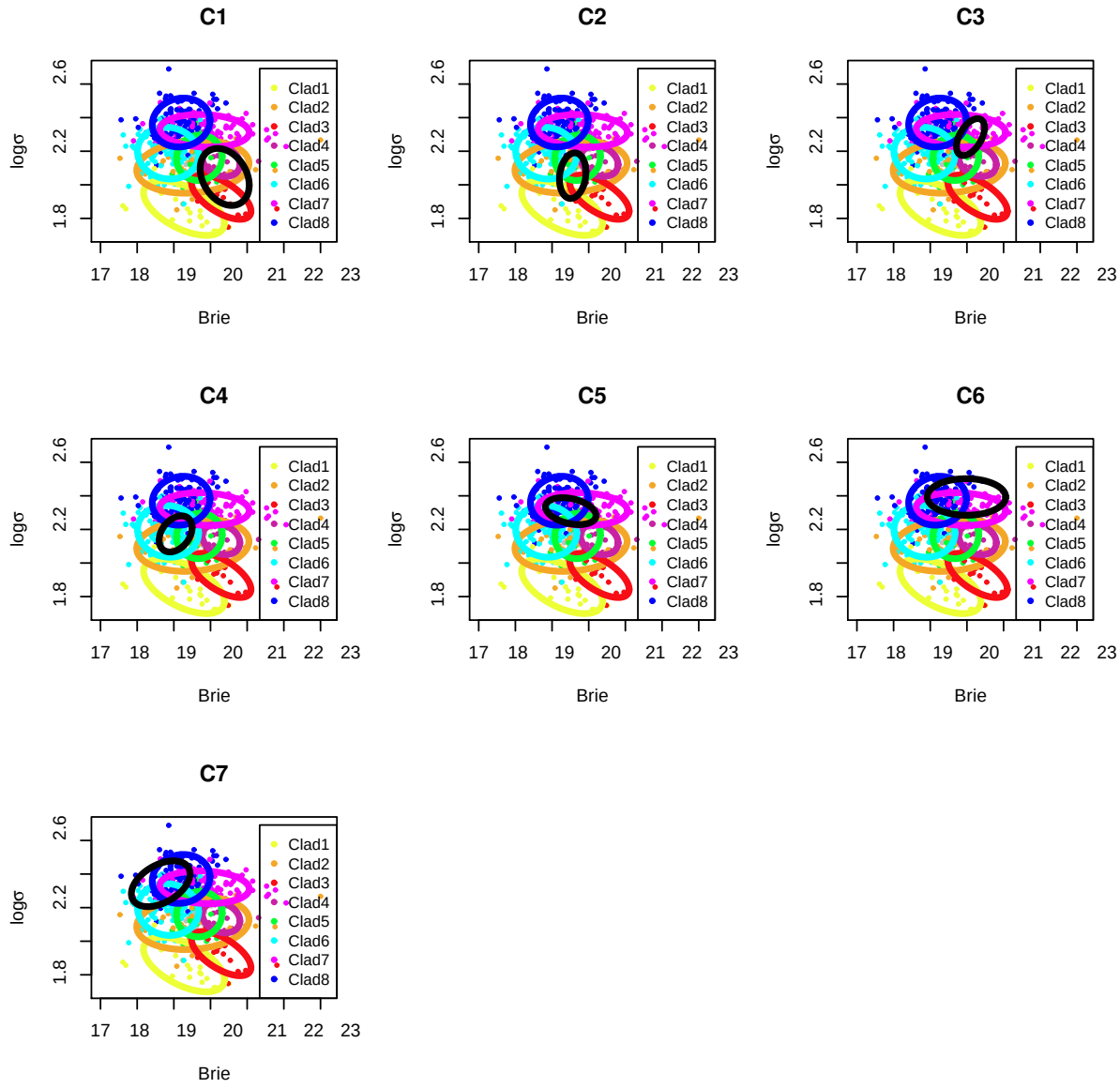

Fig. C.3. Comparison of the positions of the groups found in Fraix-Burnet et al. (2010) and those of the present paper, as projected onto the fundamental plane. The color-coded ellipses are the inertia ellipses for each group from the present paper, the black ellipse is the one for the group from Fraix-Burnet et al. (2010) indicated on top each graph. See also Fig.8.

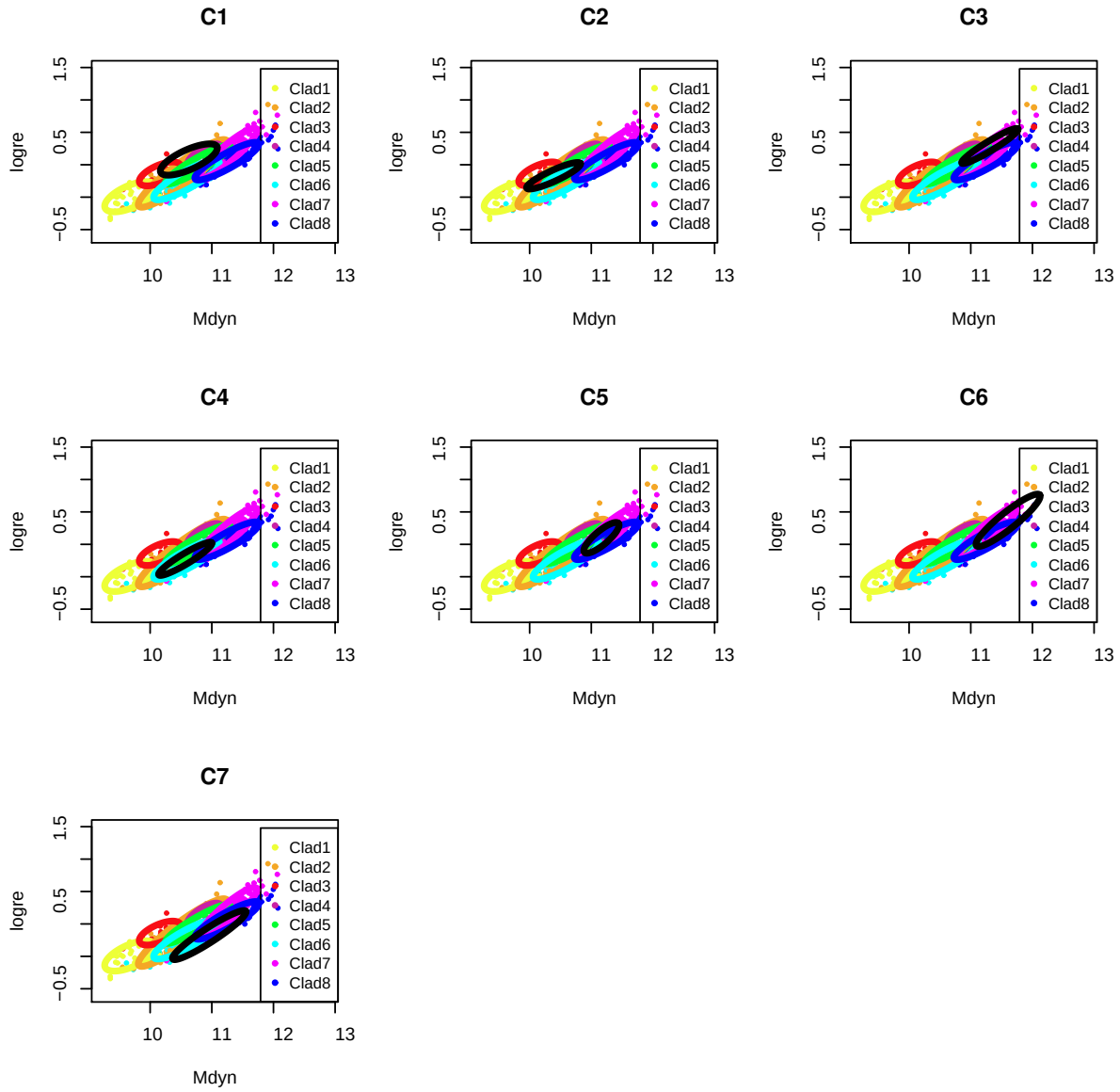

Fig. C.4. Comparison of the positions of the groups found in Fraix-Burnet et al. (2010) and those of the present paper, as projected onto the log r_e vs M_{dyn} diagram. The color-coded ellipses are the inertia ellipses for each group from the present paper, the black ellipse is the one for the group from Fraix-Burnet et al. (2010) indicated on top each graph. See also FIg.9.