Globular Cluster Subpopulations in Early-Type Galaxies: Insights into Galaxy Formation

Jean P. Brodie

University of California Observatories/Lick Observatory, University of California, Santa Cruz, CA 95064, U.S.A.

Abstract.   New and archival HST images of the globular cluster systems of 17 relatively nearby galaxies and 6 more distant galaxies have been used to establish the characteristics of GC subpopulations over a range of host galaxy morphological type from E5 to Sa. GC color/metallicity, size and luminosity distributions have been examined in the nearby galaxies and color distributions have been determined for the more distant sample. Correlations with parent galaxy properties and trends with galactocentric radius have been explored. Supplemented with Keck spectroscopy, our results are best explained by an *in situ* formation scenario in which *both* GC subpopulations formed at early times within the potential well of the protogalaxy, in multiple episodes of star formation. We have also discovered a third population of clusters, fundamentally distinct from the compact red and blue clusters common in early type galaxies.

1.   Introduction

Since their discovery in early-type galaxies nearly a decade ago, globular cluster (GC) subpopulations have been widely recognized as holding important clues about both GC and host galaxy formation. The presence of subpopulations is apparent from the bimodal color distributions in many early-type galaxies, implying multiple mechanisms or epochs of GC formation in a galaxy's history.

Explanations for bimodality fall into three categories: Mergers, *in situ*/multiphase collapse and accretion scenarios. The merger scenario, in which elliptical galaxies form from gaseous mergers of spiral galaxies (Schweizer 1987; Ashman & Zepf 1992), was the only one to actually predict bimodality rather than explaining it after the fact. The GC system of the resulting galaxy consists of a blue (metal poor) population from the progenitor spiral galaxies and a new population of red (metal rich) clusters formed in the merger event itself out of enriched gas. The *in situ*/multi-phase scenario was put forward by Forbes, Brodie & Grillmair (1997) in light of an increasing body of observational evidence not easily explained under the merger model. In this scenario the blue GCs form first in a clumpy protogalactic medium. Star and cluster formation is halted after most of the blue clusters and a few stars have formed, perhaps because of the ejection of gas from these clumps due to supernova explosions. There follows a dormant period of a few Gyr while the gas expands, cools and falls back into the now more fully formed galaxy potential. At this point star and cluster for-

218

mation starts again, this time from enriched gas, forming the red GCs and the bulk of the galaxy stars. Harris et al. (1999) came to similar conclusions based on their HST study of GCs and halo stars in NGC 5128. In accretion scenarios (e.g. Côté et al 1998) the red GCs form *in situ* in large elliptical galaxies and the blue GCs are accreted along with lower luminosity galaxies or they can be stripped from neighboring galaxies. Some new cluster formation may also be required (Hilker et al. 1999).

Clearly, mergers take place and massive clusters are formed in the process. This is frequently observed (e.g. Schweizer, 1997). Equally clearly, more massive galaxies accrete less massive ones, along with their retinues of more metal poor GCs. These phenomena must influence the final characteristics of a GC system, but to what extent? The questions we need to address are: Is there a dominant mechanism determining the global properties of GC systems in ellipticals? Is that mechanism the same for all galaxies. How do spirals fit into this picture? If GC systems are the result of a mix of processes how does that mix vary with host galaxy properties? A point to remember, though, is that at high enough redshift the distinction between the scenarios is quite blurred.

We have explored the red and blue GC subpopulations of 23 galaxies observed with the WFPC2 camera on HST. Much of the data is from our own HST programs but we have supplemented with data from the HST archive where these are of sufficient depth. The result is a large, *homogeneous* set of GC system data observed with the same telescope (HST) and instrument (WPFC2) through the same filters (except for two galaxies) and all subjected to the same reduction and analysis techniques. This homogeneity is the key to uncovering subtle correlations which can be masked by systematics when comparing data sets from the literature which have been observed and analyzed in a variety of different ways.

Our approach has been to determine the key properties of the GCs (ages, metallicities, sizes, luminosity function turnover magnitudes), relate them to characteristics of the parent galaxy (luminosity, color, environment) and try to identify trends with galactocentric radius. In addition to the HST data we have used Keck spectroscopy to provide more accurate age and abundance estimates for selected clusters. This information has allowed us to put some useful constraints on the formation history of GC systems and their host galaxies.

2. Nearby Galaxies

Our nearby sample comprises 17 galaxies out to the distance of the Virgo cluster. It is discussed in detail in Larsen et al. (2001) so I will provide only a summary and a few updates here. In Larsen, Forbes & Brodie (2001) we present additional information on the Sombrero (Sa) galaxy. Included in our nearby sample are 1 Sa, 4 S0, 11 E galaxies and 1 cD galaxy. They cover a range in M_B from -18.6 to -22.1 mag. Some are found in groups, some in rich clusters and a few are in relatively isolated environments.

2.1. Color Magnitude diagrams

The examples in Figure 1 illustrate the point that some color-magnitude diagrams are obviously bimodal to the eye, while others are not. Nonetheless, a
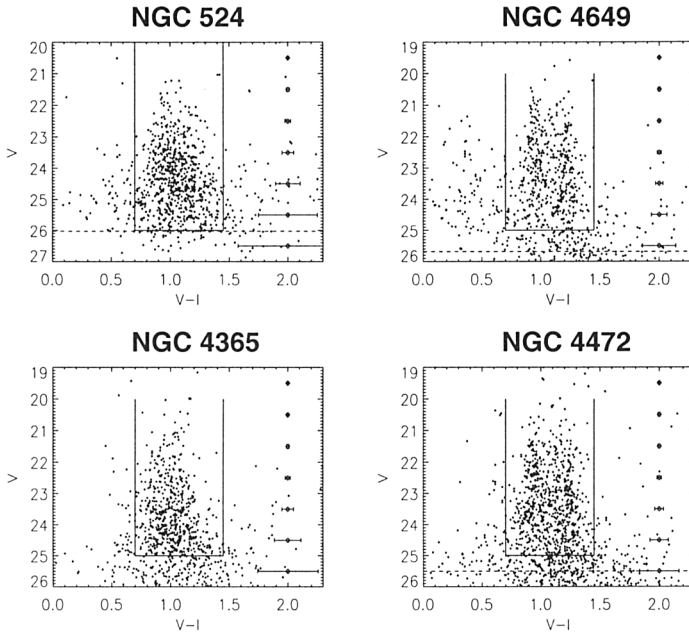
Figure 1.   Selected CMDs. The right panels are obviously bimodal while the left panels are not. Only NGC 4365 is found to be unimodal by the KMM test.

KMM test (Ashman, Bird & Zepf 1994) indicates, with high probability, that two Gaussians are a better fit to the data than a single one in nearly all cases. Interestingly, the peak $(V-I)_o$ colors are consistent even in galaxies where the CMDs and $(V-I)_o$ histograms show weak/no evidence for bimodality. One galaxy, NGC 4365, appears to have only one (broadened) peak, at $(V-I)_o = 0.98$, typical for the metal poor populations in bimodal systems. This galaxy, a bright elliptical, would be expected to possess a significant metal rich population under the accretion scenario.

The average blue and red peak $(V-I)_o$ colors are 0.95±0.02 and 1.18±0.04, corresponding to $[Fe/H] = -1.4$ and -0.6 respectively (Kissler-Patig et al. 1998). This is remarkably similar to the metallicities of the Milky Way metal poor (halo) and metal rich (disk/bulge) GCs (Zinn, 1985).

2.2.   Globular Cluster Luminosity Functions

The GC luminosity function (GCLF) is usually assumed to have a Gaussian form (although a students t_5 function is actually a better fit) whose peak appears to occur at a constant absolute magnitude, modulo metallicity effects (Ashman, Conti & Zepf 1995). The physical basis for the observed GCLF is unknown as is the degree to which it is universal. Its shape may have been set up at the time of formation or it may be the result of dynamical evolution. For example, dy-
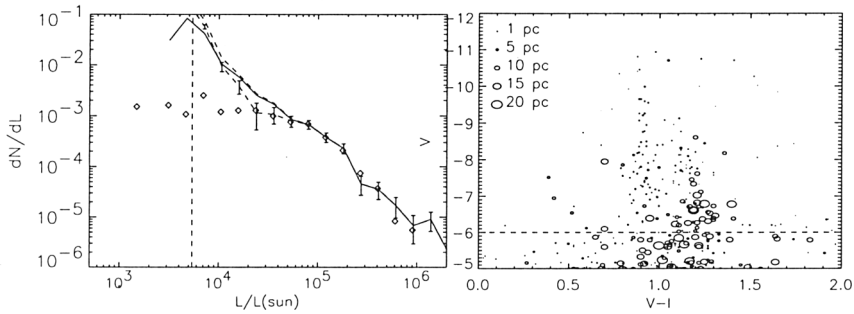
Figure 2. Left panel: 2a) The GCLF of NGC 1023 (solid line) is significantly different from that of the Milky Way (open circles) at low luminosities. Right panel: 2b) A population of faint, red, extended objects is apparent in this CMD where symbol size is proportional to object size.

namical effects could have eroded the faint end of a power-law initial luminosity function, resulting in the GCLF we observe today.

Differences between the turnover magnitudes of the blue and red subpopulations (Δm_V^{TO}) would be indicative of differences in their formation and/or subsequent evolution. Assuming the same mass function for the blue and the red GCs, similar old ages and metallicities of [Fe/H] = −1.6 and −0.6, 1996 versions of the Bruzual & Charlot SSP models predict a difference in m_V^{TO} between the blues and the reds of 0.26 mag.

We have examined the GCLF turnover magnitudes for the galaxies in our nearby sample and find $< \Delta m_V^{TO} >$=0.47 mag., or 0.30±0.16 mag. if we exclude systems for which the error on Δm_V^{TO} is >0.25. Note, though that this may introduce a bias because less populous systems belonging to lower-luminosity galaxies tend to be excluded. The fact that our observational result is close to the theoretical expectation suggests that both populations are indeed old *if* their mass functions are identical. However, if these populations formed at different epochs under different environmental conditions, differences in their mass functions would not be altogether surprising.

2.3. The Discovery of a Third Class of Cluster

Observational limitations have meant that little has been known until now about the faint wing of the GCLF, yet this is where the signatures of formation and dynamical evolution should be the strongest. At $m - M$ ~30, the nearby S0 galaxy, NGC 1023 is close enough that the entire faint wing of the GCLF is accessible for high S/N imaging with HST and spectroscopy with 8–10m class telescopes.

As Figure 2a shows, the faint wing of the GCLF in NGC 1023 deviates significantly from that of the Milky Way. This due to the presence of a third population of clusters in addition to the normal compact blue and red subpopulations commonly found in early type galaxies (Figure 2b). These hitherto undetected faint, red extended objects are aligned with the galaxy isophotes and thus appear to be associated with its disk. These "faint fuzzies" are dis-
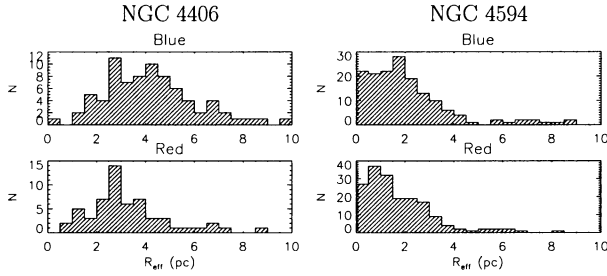
Figure 3.   GC size (R_{eff}) distributions. Blue GCs (upper panels) are roughly 20% larger than red GCs.

cussed in detail in Larsen & Brodie (2000). They have no analogs in the Milky Way and it is impossible at present to assess how common they might be. They are obviously difficult to detect, being extended low surface brightness objects fainter than the GCLF turnover and, of the galaxies studied with HST in sufficient depth for accurate size measurements, they would have been detectable in only 4 cases. They do appear to be present in NGC 3384, another lenticular galaxy, but are definitely absent in the populous lenticular NGC 3115.

In speculating about the origin of these objects we note that there is a nearby dwarf companion galaxy containing two brighter blue objects for which we have Keck spectra. These objects are young clusters (ages ~300 Myr) possibly formed as a result of the interaction with NGC 1023. Perhaps the faint fuzzies are older remnants of similar interactions with long-ago digested companion galaxies.

2.4.   Sizes

Sizes differences between blue and red GC populations had been noticed previously in four galaxies by ourselves (NGC 4472: Puzia et al. 1999; NGC 1023: Larsen & Brodie 2000) and others (NGC 3115: Kundu & Whitmore 1998; M87: Kundu 1999). With our new large sample we could assess how common this phenomenon might be. We found that the blue GCs are always larger than the reds by about 20%. Figure 3 gives examples. Although suitable multiple pointings exist for only a few galaxies, the size difference between the blue and red GCs persists at all radii in these systems. Interestingly, the same size difference is seen in the Milky Way, and at all radii. Note again, the similarities between the early type galaxies and the spirals (Milky Way and Sombrero).

2.5.   Correlations with Parent Galaxy Properties

Brodie & Huchra (1991) showed that GC mean metallicity correlates with parent galaxy luminosity. With the subsequent discovery of multiple GC populations the question naturally arises as to whether the correlation exists for one or both of the subpopulations separately.

In Figure 4 we show that the peak colors of *both* the red and the blue GCs correlate with parent galaxy M_B at the 2-3 σ confidence level. A similar relation
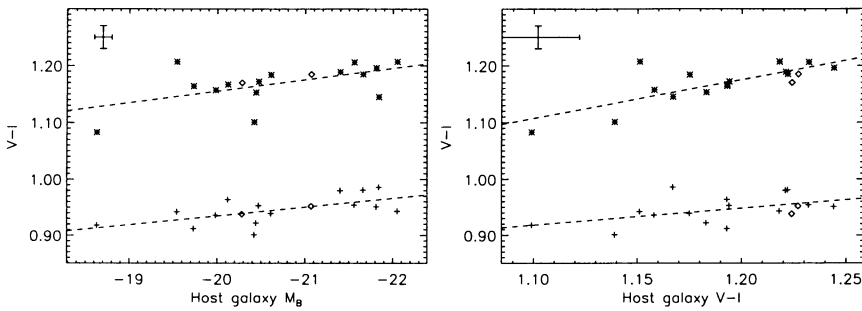
Figure 4.   GC peak colors versus host galaxy properties.

exists for $(V − I)_0$ vs. central velocity dispersion. The slope of the relations is steeper for the red GCs than for the blues. There is a 4 σ correlation between the peak colors of the red GCs and parent galaxy color[1] and a 2 σ correlation for the blues.

These results are somewhat different from previous findings (Forbes et al. 1997; Burgarella, Kissler-Patig & Buat 2001; Forbes & Forte 2001) in which the red GC colors were found to correlate strongly with parent galaxy properties while the blue ones were thought not to correlate significantly. This may be because of the heterogeneous nature of the data used in these studies which tends to mask more subtle correlations.

The implication of a correlation between the properties of the blue GCs and the parent galaxy is profound. At the time of formation, the blue GCs, or at least a significant fraction of them, must have known about the final galaxy to which they would belong. This may present problems for mergers and/or accretion as dominant mechanisms in the formation history of GC systems.

3.   More Distant Galaxies

Work is underway on the GCLFs and color distributions of a sample of more distant galaxies. These systems are too far away for accurate size estimates for individual clusters.

Table 1.   Color Distributions of Galaxies Beyond Virgo.

Galaxy	Location	Dist. Mod	$(V − I)_{0B}$	$(V − I)_{0R}$
NGC 3311	Hydra	33.5	0.91	1.09
IC 4051	Coma	35	0.95	1.15
NGC 4881	Coma	35	0.95	-
NGC 5846	Group	32.3	0.96	1.17
NGC 7562	GH166	33	0.97	-
NGC 7619	Peg I	33.5	0.99	1.24

---

[1]The lack of correlation reported previously was due to the use of inaccurate galaxy colors. New galaxy colors from Tonry et al. (2001) are used here and reveal the expected correlation.

In Table 1 we summarize our results for the peak colors of the blue and red subpopulations. Notice that these galaxies have peak colors that are typical of those found in the nearby sample. NGC 3311 was previously thought to have an almost exclusively metal-rich GC population (Secker et al. 1995) from ground-based photometry. This would present challenges to both the *in situ* and merger scenarios. However, Brodie, Larsen & Kissler-Patig (2000) recently showed that this galaxy has an entirely normal color distribution based on our deep HST data. NGC 4881 appears to be the analog of the nearby galaxy, NGC 4365, in having a single (broadened) metal poor peak, as noted by Baum et al. 1995. Too few clusters are assigned to the red peak in NGC 7562 to allow a meaningful estimate of $(V - I)_{0R}$.

4.   Formation Scenarios

As we have seen, any successful formation scenario needs to explain a very wide range of properties of GC color distributions:
• Bimodal with roughly equal N_{Blue} and N_{Red} (NGC 1404, 4649, 4472).
• Bimodal with much reduced N_{Red} (NGC 4406).
• Single (but broadened) metal poor peak (NGC 4365, 4881).
• Continuous color distribution spanning a range similar to true bimodals (NGC 524, 4552).

    In addition, a successful model will have to accommodate the increasing evidence that both subpopulations are old. That this is the case is borne out by our studies of the GCLFs of the blue and red subpopulations in that the turnover magnitudes differ only by the amount predicted by the SSP models under the assumption of similar old ages (see Puzia et al. 1999 for a detailed discussion). Evidence for old ages comes most convincingly from spectroscopy but few galaxies GC systems have so far been studied with large enough samples and adequate signal-to-noise to place interesting constraints on age differences. Only NGC 1399 (Kissler-Patig et al. 1998), NGC 4472 (Beasley et al. 2000), M87 (Cohen, Blakeslee & Ryzhov 1998), M81 (Schroder et al. 2001) and M31 (Huchra, Brodie & Kent 1991; Perrett et al. 2001) have good enough spectra for relative age constraints. It is likely to prove extremely diffiult to differentiate ages greater than ∼8–10 Gyr, even with superb signal-to-noise data because age contours are so closely spaced (they even cross in some models) in index-index planes generated from SSP models. Similar old ages would be at odds with the standard merger picture but would be compatible with mergers occurring at high redshift where the similarities between the models outweigh their differences.

    Particularly interesting constraints to emerge from this work are the correlations between GC colors and host galaxy luminosity (mass) and color for *both* red and blue GCs. If the correlation for the blue GCs is verified it will be difficult to explain under both the merger and accretion scenarios. Note too that we consistently find roughly equal numbers of red and blue GCs, or a preponderance of blues. This is also a challenge for these scenarios. In particular, in the merger picture, the lower specific frequency of spirals must be increased to the higher (by a factor of ∼3 for giant E galaxies) value in ellipticals by the production (in the merger) of numerous red clusters. It is hard to escape the conclusion that the resulting galaxy would have a preponderance of red clusters. The accretion

model is challenged only because such a large number of blue clusters must be acquired along with minimal amounts of starlight.

5. What's Next?

We do not yet understand the origin of the size difference between red and blue GCs, whether it is primordial or the result of dynamical evolution. To produce it by dynamical effects would require constancy in the ratio of the perigalactic distances of the blue and red subpopulations over a wide range of parent galaxy luminosities and environments. Were red GCs formed on predominantly radial orbits and blue GCs formed on predominantly circular orbits? It might be more natural to suppose that red GCs formed from denser protocluster clumps, perhaps under conditions influenced by their higher metallicities.

An answer to the size question will naturally require kinematic information for large samples of GCs. This information will also inevitably assist in differentiating between models for the origins of GC subpopulations but it is not yet clear how to interpret such data. We are accumulating relevant information at an increasing pace from spectroscopic programs on 8–10m class telescopes. What is largely lacking is a theoretical basis for comparison between the scenarios (but see Bridges article in these proceedings for a comparison of current data with merger ideas).

Spectroscopic data also offer the best chance of age discrimination and we can expect additional insights into the timescales for GC formation from studies of individual element abundance ratios, in particular the ratios of α elements to Fe which reflect the relative contributions of SN Types I and II.

Much is still to be learned from photometry, particularly with HST. Outstanding questions relevant to the work described here include: How common are the "faint fuzzies"? Are they always associated with the disks of lenticulars? What else will we find when we study the faint wings of GCLFs? Can we see the signatures of dynamical evolution by observing galaxies with a range of ages?

We are beginning to learn that there are many similarities in the characteristics of the subpopulations in spirals and ellipticals. Since it is not generally thought that spirals themselves form from *major* mergers of disk systems, this might present another challenge to the standard merger picture. However, more spiral GC systems (especially those of later type galaxies) need to be studied before we will know how widespread these similarities are and to what extent they will influence our ideas on GC and galaxy formation.

6. Conclusions

Overall then, our results seem to fit better with an *in situ* scenario in which both GC populations "knew" about the size of the final galaxy to which they would belong. This implies that the initial phase of GC formation in gE galaxies must have taken place after they assembled into individual entities, i.e., both GC populations formed within the potential well of the protogalaxy in multiple episodes of star formation. If the *in situ* idea is right, accurate age-dating of the red GCs will pinpoint the epoch at which the bulk of the galaxy was formed. Our best estimates suggest this occurred > 8–10 Gyr ago but it is of

critical importance to improve this constraint for comparison with models (such as hierarchical clustering) for the formation of structure in the early universe.

References

Ashman, K. M., Bird, C. M., and Zepf, S. E. 1994, AJ, 108, 2348

Ashman, K. M., Conti, A., and Zepf, S. E. 1995, AJ, 110, 1164

Ashman, K. M., and Zepf, S. E. 1992, ApJ, 384, 50

Beasley, M.A., Sharples, R.M., Bridges, T.J., et al. 2000, MNRAS, 318, 1249

Brodie, J.P. & Huchra, J.P. 1991, ApJ, 379, 157

Brodie, J.P., Larsen, S.S. & Kissler-Patig, M. 2000, ApJLetters, 543, L19

Burgarella, D., Kissler-Patig, M., Buat, V. 2001, AJ, 121, 2647

Cohen, J., Blakeslee, J., Ryzhov, A., 1998, ApJ, 496, 808

Côté, P., Marzke, R. O., and West, M. J. 1998, ApJ, 501, 554

Forbes, D. A., Brodie, J. P., and Grillmair, C. J. 1997, AJ, 113, 1652

Forbes, D. A., and Forte, J. C. 2000, MNRAS, 322, 257

Harris, G. L. H., Harris, W. E. and Poole, G. B. 1999, AJ, 117, 855

Hilker, M., Infante, L., and Richtler, T. 1999, A&AS, 138, 55

Huchra, J.P., Brodie, J.P. & Kent, S. 1991, ApJ, 370, 495

Kissler-Patig, M., Brodie, J. P., Schroder, L. et al., 1998, AJ, 115, 105

Kundu, A. 1999, PhD thesis, Univ. of Maryland

Kundu, A., and Whitmore, B. C. 1998, AJ, 116, 2841

Larsen, S. S., and Brodie, J. P., 2000, AJ, 120, 2938

Larsen, S. S., Brodie, J.P., Huchra, J.P. et al., 2001 AJ, 121, 2974

Larsen, S. S., Forbes, D. A., & Brodie, J. P., 2001, MNRAS, submitted

Perrett, K., et al. 2001 in preparation

Puzia, T. H., Kissler-Patig, M., Brodie, J. P., and Huchra, J. P. 1999, AJ, 118, 2734

Schroder, L., Brodie, J.P., Huchra, J.P., et al. 2001, AJ, in press

Secker, J., Geisler, D., McLaughlin, D.E. & Harris, W.E. 1995, AJ, 109, 1033

Schweizer, F. 1987, in Nearly Normal Galaxies, ed. S. M. Faber (New York: Springer), 18

Schweizer, F. 1997, in The Nature of Elliptical Galaxies, ASP Conf. Ser., Vol. 116, eds. M. Arnaboldi, G.S. Da Costa, P. Saha, 447

Tonry, J. L., Dressler, A., Blakeslee, J. P., et al. 2001, ApJ, 546, 681

Zinn, R., 1985, ApJ, 293, 424

Discussion

P. Goudfrooij : When you state that the color of GCs does not correlate with galaxy color, are you referring to the <u>Global</u> galaxy color as listed in the major galaxy catalogs, or to the <u>Local</u> galaxy color at the galactocentric radius of the clusters?

J. Brodie : As to your suggestion that the stars in the host galaxies may continue to be formed at some level and cause some chemical enrichment of the stellar population: The amount of such chemical enrichment is hampered by the observation that the stellar population in ellipticals shows highly supersolar alpha-element abundances which require very rapid star formation ($\tau_{SF} < 1 Gyr$).

P. Goudfrooij : The colors of the galaxies come from the literature so were not measured at the same radius as the clusters?

J. Brodie : Yes, it would be interesting to see if the rapid star formation timescales can accomodate enough further enrichment of the field stars to match their observed colors relative to the colors of the red globular clusters.

F. Schweizer : From your work and the published literature, does any E galaxy exist with only red, metal-rich globular clusters?

J. Brodie : No, so far we have not found any. Some which were previously thougth to be exclusively metal-rich have since been shown to be normal bi-modal systems.

F. Schweizer : Have you looked for correlations between the color difference $\Delta(V - I)_0$ between the blue- and red-cluster peaks and the spectroscopic galaxy ages of, e.g., Duncan Forbes? Merger models would predict that such correlations should exist?

J. Brodie : Yes, we have looked and no, we do not see such a correlation. $\Delta(V - I)_0$ is actually fairly constant.

M. Kissler-Patig : The correlations between GC proterties and galaxy properties does not imply that all GC formed in the final MW halo, but only that <u>some</u> GCs did. "Faint Fuzzies" were found in 2 out of 4 galaxies. Is there a correlation with galaxy type? Any analogy with MW fuzzies? Are they expected to survive in the center of galaxies (would this explain their absence?).

J. Brodie : The two galaxies in which the faint fuzzies are found are both lenticular. Of the two where they are not found, one is a lenticular, NGC 3115 (where their presence can be excluded with high confidence) and the other is an elliptical. There are no analogs to these objects in the MW. The spatially extended clusters in the MW are blue (metal poor) not red and they are located only far out in the MW halo. We are attempting to get spectra for ghese objects to see

what their kinematics can tell us about their origins and their ability to survive.

M. Fall : Could you comment on a puzzle indicated by your observationary relations? You find a correlations between cluster color and galaxy luminosity but not one between cluster color and galaxy color. Yet there is a well-known correlation between galaxy color and galaxy luminosity (bright galaxies are redder than faint galaxies on average). How do you reconcile all these?

J. Brodie : We were also surprised by this result for that very reason. The answer may be in the scatter of the galaxy L vs. color relation. There is a wide scatter in color at a given L. The galaxy relation is defined over a very wide range of luminosities. Over the relatively narrow range of L occupied by our sample it is difficult to see such a relation.

W. Harris : A nice discussion of a comprehensive data set. In your list of more distant ellipticals, you did not list NGC 4874, the supergiant elliptical at the dynamical center of the Coma cluster. The study we published last year (Harris et al. 2000, AJ) showed that the large number of clusters in it all are blue, i.e. a narrow metal-poor distribution with virtually no metal-rich component at all. This is extremely puzzling because it's sitting at the center of the deepest potencial well in the local universe, and if there are metal-rich clusters anywhere, they ought to be there. No scenarios fit this very well. I believe cases like this show us that we haven't seen everyting yet!

J. Brodie : NGC 4874 was not in my list because we ourselves had not examined it. That list represented only those galaxies forming part of our homogeneous data set and was not meant to be a complete summary of work on more distant galaxies. I agree, though, that your result is very interesting and puzzling.