Chin. J. Astron. Astrophys. Vol. 5 (2005), No. 3, 221–228
(*http://www.chjaa.org*)

Chinese Journal of
Astronomy and
Astrophysics

Relation between Starlight and Nebular Emission Lines of Star-Forming Galaxies *

Hong-Lin Lu[1], Hong-Yan Zhou[2], Ting-Gui Wang[2], Zhen-Quan Zhuang[1], Xiao-Bo Dong[2], Jun-Xian Wang[2] and Cheng Li[2]

[1] Department for Electronics of Science & technology, University of Science and Technology of China, Hefei 230026; *mtzhou@ustc.edu.cn*

[2] Center for Astrophysics, University of Science and Technology of China, Hefei 230026

Received 2004 November 9; accepted 2004 December 1

Abstract We present an exercise that intends to establish a relationship between the strength of nebular emission lines and optical stellar features in the spectrum of a galaxy. After accurately subtracting the stellar continuum and the underlying stellar absorption, we made reliable measurements of the emission lines of all the galaxies in the Sloan Digital Sky Survey Data Release 2 (SDSS DR2). More than 4000 star-forming galaxies with high S/N ratio of both the stellar spectrum and the emission lines are selected. These galaxy spectra are fitted with the 10 PCs of Yip et al., after all the emission line regions have been filtered out. We find that the flux of hydrogen Balmer emission lines, Hα and Hβ can be well recovered from the PCs, while the metal lines are not well reproduced. The fluxes of Hα and Hβ measured from the PC-reconstructed spectra and from the observed spectra agree well with an rms scatter of only ~ 0.1 dex. This result suggests that, with moderate spectral resolution and S/N ratio, the optical stellar spectrum of a galaxy can serve as an indicator of star formation rate.

Key words: galaxies: fundamental parameters — galaxies: starburst — radiation: lines, continuum

1 INTRODUCTION

The observed optical spectrum of a galaxy is a combination of stellar spectrum and nebular emission lines. A major problem in the study of the connection between star formation and nuclear activity of Type II AGN lies in the decoupling of nebular emission of star forming region from that of the AGN. Hydrogen Balmer emission line, Hα, is thought to be the most reliable indicator of star formation rate (SFR), because, in ionization-bounded nebulae of HII regions and HII galaxies, the Hα luminosity scales directly with the total ionizing flux of the embedded massive stars (c.f., Kewley, Geller & Jansen 2004).

SFR indicators developed three decades ago were based on stellar population synthesis models of galaxy colors (e.g., Tinsley 1968; Searle et al. 1973) and subsequently gave place to

* Supported by the National Natural Science Foundation of China.

more precise SFR indicators based on optical emission lines and UV, IR and radio luminosities (e.g., Hopkins et al. 2003; Hirashita, Buat & Inoue 2003). However, application of these more recent SFR indicators is limited to inactive galaxies, because it is difficult to estimate the relative contribution of the nuclei and the host galaxies. This difficulty has been in the way of our exploring the SFR of active galaxies. Starlight can leave numerous imprints in an AGN spectrum and this property has been used by many authors to address important issues, such as starburst-AGN connections. HeI and/or strong stellar absorption features in the high-order Balmer series have been detected in many bright Seyfert 2 galaxies, indicating the presence of young (a few Myrs) and intermediate-age (a few 100 Myrs) stellar populations (González Delgado, Heckman & Leitherer 2001). An examination of more than twenty thousand SDSS narrow emission line AGNs reveals that the hosts of high-luminosity AGN have much younger mean stellar ages (Kauffmann et al. 2003). However, no experiment has been attempted so far to directly measure SFR in the AGN host galaxies.

In order to estimate emission line fluxes from stellar features, galaxy spectra with moderately high resolution and S/N ratio are needed so that blend of Balmer emission/absorption lines can be disentangled. Fortunately, the data obtained by new large spectroscopic surveys, such as the SDSS (York et al. 2000), can allow us to do so. Yip et al. (2004) applied Principle Component Analysis (PCA, or Karhunen-Loeve transformation) to $\sim 170\,000$ galaxies and compressed their information to 10 Principle Components (PCs). Correlations between the emission and absorption lines as well as the SED can be identified in individual PC spectrum. This result indicates that tracing galaxy SFR based on optical stellar spectrum may be feasible. This potential SFR indicator can serve as a useful tool for the study of the physical link between starburst and AGN phenomena.

In this paper, we try to validate this feasibility via estimating the nebular emission line fluxes using stellar spectrum. More than 4000 star-forming galaxies are used in this exercise. In Sect. 2 we carefully subtract the starlight from the SDSS galaxies and measure the emission line fluxes. In Sect. 3 we reconstruct the nebular emission lines of our star-forming galaxies by fitting the stellar component of the observed spectra using the 10 PCs as templates. The estimated emission line fluxes are measured from the reconstructed spectra. These estimated fluxes are compared with the observed values. Implications of the results are discussed in Sect. 4. We present our preliminary conclusion in Sect. 5, together with some future prospect.

2 THE SDSS DATA AND THE SAMPLE OF STAR-FORMING GALAXIES

The data analyzed in this work are drawn from the Sloan Digital Sky Survey Data Release 2 (DR2, Abazajian et al. 2004), which includes $\sim 2.6 \times 10^5$ galaxy spectra selected over 2627 deg^2 using well defined criteria. Detailed description of the data pipelines and products can be found in Stoughton et al. (2002). This large database of galaxy spectra with broad wavelength coverage, and generally high quality are extremely well suited for our present purpose of searching for possible relation between the nebular emission lines and spectral signatures of early-type stars, which are heavily blended with emission lines. All these galaxies are corrected for Galactic extinction using the extinction curve of Schlegel et al. (1998) transformed to their rest frame using the redshift provided by the SDSS spectroscopic pipeline.

2.1 Starlight Subtraction and Emission Line Measurement

In order to reliably measure the emission line parameters, stellar absorption features have to be carefully modelled and accurately subtracted. A two-step interactive procedure was used to model the stellar spectrum and to measure the emission lines.

In the first step, we fit emission line free regions in the galaxy spectra with

$$g^S(\lambda) = A(E_{B-V}, \lambda) \sum_{i=1}^{6} g_i IC^i(\lambda, \sigma_*), \quad (1)$$

where $g^S(\lambda)$ is the starlight component, $A(E_{B-V}, \lambda)$ is the starlight dust extinction assuming the Small Magellanic Cloud extinction curve of Pei (1992), and $IC^i(\lambda, \sigma_*)$ (i=1, 2, ..., 6) are six synthesized galaxy templates convolved with a Gaussian of σ_* width to match the stellar velocity dispersion of the galaxy under study. These templates are derived from the spectral library of SSPs of Bruzual & Charlot (2003) using our new method based on Ensemble Learning Independent Analysis (see Lu et al. 2004 and Li et al. 2005 for detailed description). The final fit is done through minimization of reduced χ^2 and E_{B-V}, g_i, and σ_* are free parameters.

We first filter out all the wavelength regions with width of 600 km s^{-1} centered at the emission lines that often appear in an HII galaxy. Once the modelled stellar spectrum is subtracted from the observed galaxy spectrum, we fit the line spectrum with Gaussian(s) in the second step. Then the measured wavelength ranges of emission lines are transformed to the first step, which is repeated a second time. The iteration continues until both the modelled stellar spectrum and the emission line parameters are acceptable (refer to Dong et al. 2005 for detail). Representative examples of the final filtered emission line regions and the EL-ICA model of starlight are shown in Figs. 1 and 2 (upper panels).

2.2 Selection of Star-Forming Galaxy Sample

Our sample of star-forming galaxies are selected according to the following criteria:

- confidence of redshift > 0.90: to ensure accurate redshift correction;
- median continuum S/N> 25: with such spectral quality, important parameters can be derived reliably from the stellar spectrum (Lu et al. 2004);
- S/N> 5 for [OIII]λ5007 and [NII]λ6583 and S/N> 10 for Hα and Hβ;
- spectroscopically classified as HII galaxies according to the diagnostic diagram of Baldwin, Phillips & Terlevich (1981, hereafter BPT).

We plot all the 8753 objects fulfilling the first three criteria in the [OIII]λ5007/Hβ against [NII]λ6583/Hα BPT diagram (Fig. 3). We adopt the criterion proposed by Kauffmannn et al. (2003) to demarcate between starburst galaxies and AGN: An object is defined as a bona fide star-forming galaxy if

$$\log([\text{OIII}]\lambda5007/\text{H}\beta) < \frac{0.61}{\log([\text{NII}]\lambda6583/\text{H}\alpha) - 0.05} + 1.3. \quad (2)$$

There are 4312 galaxies that fulfil this criterion. All of these galaxies are examined by eye and 13 objects are rejected either because of the presence of broad emission line(s), or because of heavy contamination by a companion galaxy. The remaining 4299 sources form our sample of star-forming galaxies.

3 RECONSTRUCTING EMISSION LINES BASED ON STELLAR SPECTRUM

The PC-reconstructed spectra of all the 4299 star-forming galaxies are obtained by linear combination of the 10 PCs provided by Yip et al. (2004),

$$g^R(\lambda) = \sum_{i=1}^{10} r_i PC^i(\lambda). \quad (3)$$

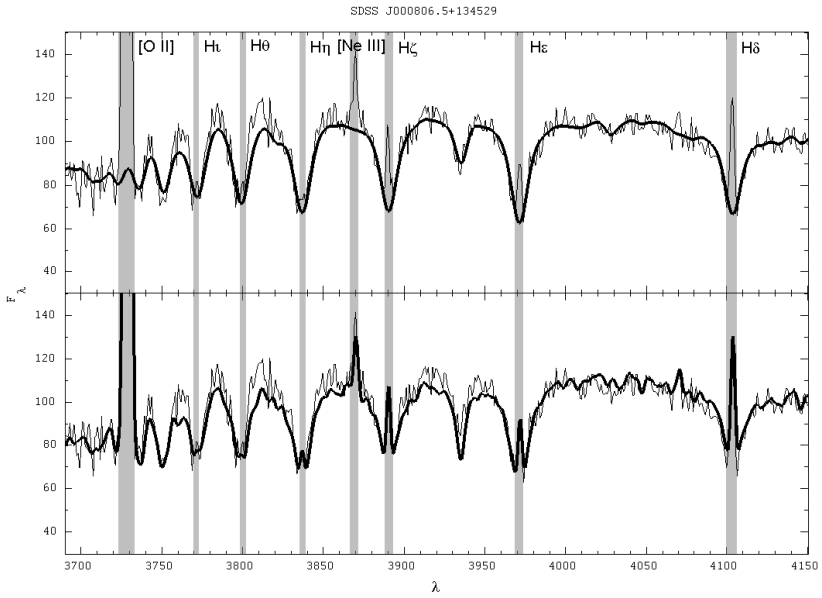

Fig. 1 An example of masking emission line regions in the procedure of modelling the stellar spectrum using the six ICs (upper panel) and reconstructing the observed spectrum using the 10 PCs (lower panel). Flux is in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1}. The thin lines are the observed spectrum of SDSS J000806.5+134529, and the thick lines are the best-fits. The miss-match continua of $\lambda \lesssim 3950$ Å between the modelled and the observed spectra are due to error in the flux calibration of the SDSS pipeline. The gray regions mark the filtered emission-line regions (labelled).

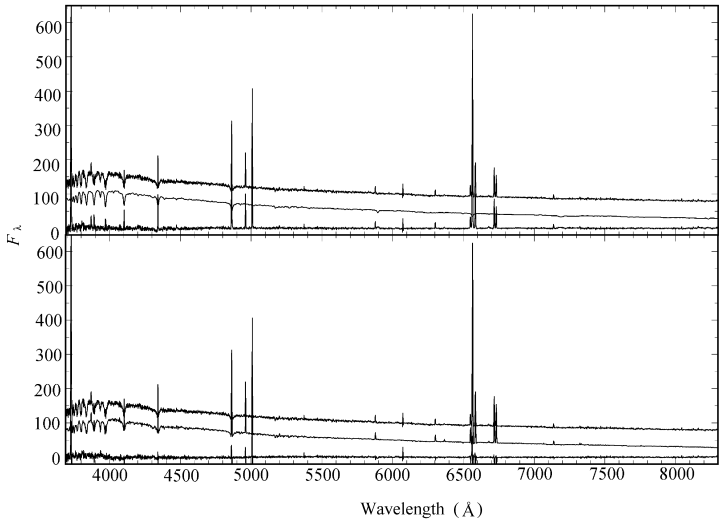

Fig. 2 A representative example of modelling the stellar spectrum using the six ICs (upper panel), and reconstructing the observed galaxy spectrum using the 10 PCs (lower panel, see the text for detail). Flux is in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1}. In each panel, the upper curve is the observed spectrum of SDSS J000806.5+134529, the middle one is the best-fitting spectrum, and the lower one is the residual (arbitrary constant is added to the observed spectrum for clarity).

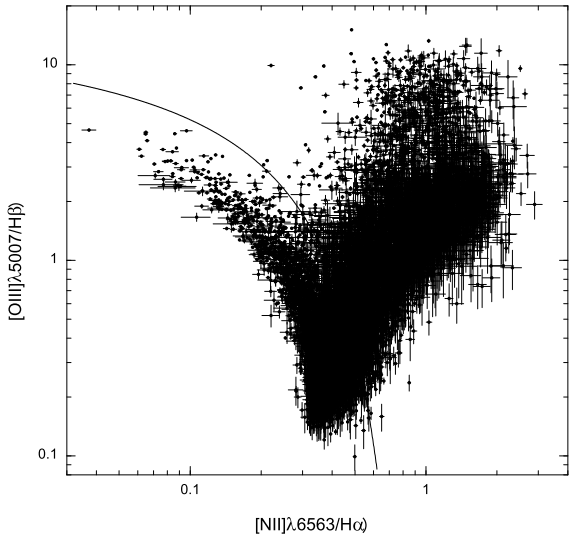

Fig. 3 Distribution of the 8753 high quality galaxies fulfilling the first three criteria of Sect. 2.3 in the BPT diagnostic diagram. The line ratio of [NII]λ6583/Hα is plotted against that of [OIII]λ5007/Hβ. Our sample of star-forming galaxies are distributed in the lower-left corner below the solid line of demarcation proposed by Kauffmann et al. (2003).

We perform the fits with all the regions of emission lines determined in Sect. 2.1 masked (Fig. 1 lower panel). Then the starlight in the PC-reconstructed spectra is subtracted (Fig. 2 lower panel) and the reconstructed emission lines are measured in the same way as described in Sect. 2.1.

We plot in Fig. 4 the fluxes of the Balmer emission lines, Hα and Hβ, measured from the PC-reconstructed spectra against that measured from the observed spectra. It can be seen that the reconstructed Hβ flux agrees quite well with the observed value. The $1\,\sigma$ rms scatter is ~ 0.09 dex. However, the reconstructed Hα flux is $\sim 85\%$ of the observed flux, with $1\ \sigma$ rms scatter of ~ 0.10 dex. The reconstructed fluxes of both Hβ and Hα are systematically smaller than the observed fluxes of weak emission line galaxies.

We find that forbidden lines, such as [OII]λ3727, [OIII]$\lambda\lambda$4959, 5007, [NII]$\lambda\lambda$6548, 6583, and [SII]$\lambda\lambda$6716, 6731, are reconstructed with much less accuracy.

A plot between the reconstructed and observed flux of [OIII]λ5007 is shown in Fig. 5. The differences here are much larger than for the hydrogen combination lines. The PCs over-predict the [OIII] flux except at the very low [OIII] flux end.

4 DISCUSSION

Nebular emission lines are produced in HII regions around young and massive (OB) stars. For a Simple Stellar Population (SSP[1]), the evolutionary timescale of O stars is $\sim 10^7$ yr. Hence the Hα flux from a galaxy is sensitive to stellar populations of ages $\lesssim 10^7$ yr. Other characteristic of early-type stars is strong hydrogen Balmer, neutral helium absorption lines and blue continuum. The strength of Balmer absorption lines, which are prominent in an SSP

[1] An SSP is defined as a stellar population whose star formation duration is short in comparison with the lifetime of its most massive stars.

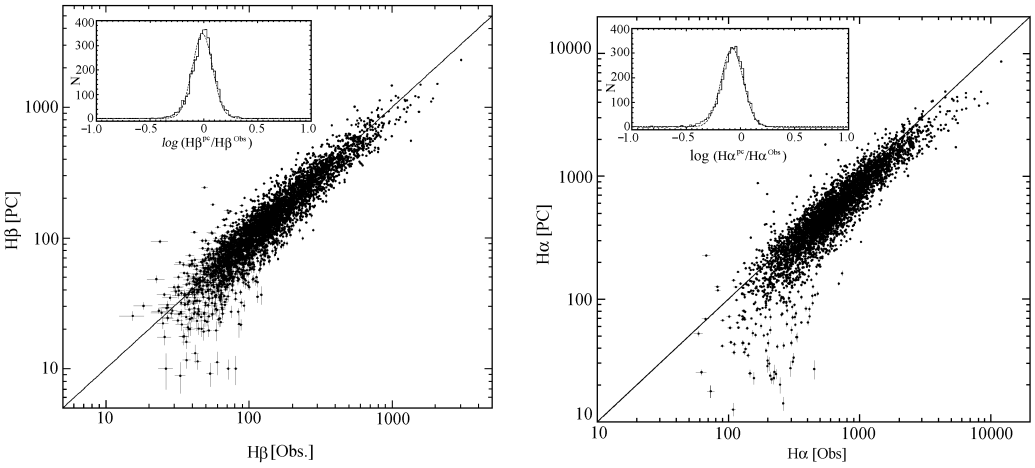

Fig. 4 Comparison between the PC-reconstructed and observed emission line flux of Hβ (left panel) and Hα (right panel) in units of 10^{-17} erg s^{-1} cm^{-2} Å^{-1}. The expected one-to-one line is plotted to help guide the eye. The inset plots show the deviation between the fluxes measured from the reconstructed and observed spectra. The dashed lines are Gaussians with $\sigma = 0.09$ and $\sigma = 0.10$ dex for Hβ and Hα, respectively. These tight correlations demonstrate that we can estimate hydrogen Balmer emission lines of star-forming galaxies from the optical stellar features.

spectrum with age $\gtrsim 3 \times 10^7$ yr, is strongly age dependent and continues to increase until $\sim$a few$\times 10^8$ yr. Therefore, for a given Initial Mass Function (IMF), if we accurately correct the dust extinction of the starlight and the emission lines, then we can use the absorption features and the continuum of the stellar component in a galaxy spectrum to estimate the flux of the Balmer emission lines, and hence the SFR of the galaxy.

Recalling the fact that the emission line regions are all filtered out in our reconstruction of galaxy spectra, the PC-reconstructed spectra of hydrogen Balmer emission lines agree well with the observed ones. This result strongly suggests that the stellar spectrum of a galaxy can be used to infer its current SFR, because in optical the emission line free regions of a galaxy spectrum are exclusively contributed by starlight. Using the Kennicutt (1998) relation between Hα luminosity and SFR,

$$\mathrm{SFR}(M_\odot \mathrm{yr}^{-1}) = 7.9 \times 10^{-42} L(\mathrm{H}\alpha) \ \ (\mathrm{erg\ s}^{-1}), \tag{4}$$

we calculate the SFRs for all the star-forming galaxies in our sample (a flat Λ-dominated cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1} and $\Omega_M = 0.3$ is assumed). The SFR in the SDSS 3″ fiber range of our sample spans four orders of magnitude, from $\sim 2 \times 10^{-2}$ to 2×10^2 $M_\odot$ yr^{-1} (Fig. 6). The specific SFRs (SFRs normalized by stellar mass M_*) are $\sim 10^{-11} - 10^{-7}$ using the stellar mass derived by Lu et al. (2004). This indicates that star formation histories can be traced using the optical stellar properties of galaxies with a wide SFR range. However, the present method tends to underestimate the SFR in weak emission-line galaxies. In these cases, the stellar properties of massive stars are seriously diluted by old stellar populations. Reddening correction is another problem. The reconstructed Balmer decrement is only $\sim 80\%$ of the observed value. This results from the fact that extinction is much larger for emission lines than for starlight. Lu et al. (2004) found the former is about 2.8 times larger than the latter.

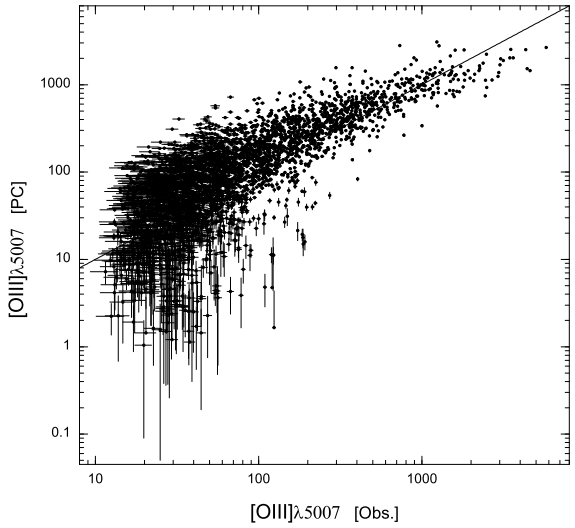

Fig. 5 Same as Fig. 4 but for [OIII]λ5007. Note the deviation between the reconstructed and observed flux of [OIII] is significantly larger than in the case of Balmer emission lines.

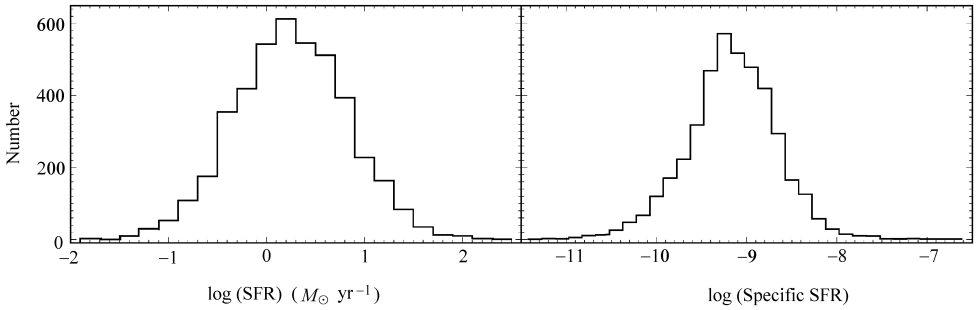

Fig. 6 Distribution of the SFR (left panel) and the specific SFR (right panel) of the sample of the SDSS star-forming galaxies.

It can also be easily understood that metal emission lines cannot be well reconstructed based on the stellar spectrum. Over-prediction of [OIII] may be a consequence of inclusion of AGN in the construction of the PCs. The main cause, perhaps, is that the chemical abundance is in general quite different in stars and nebulae, and the discrepancy, furthermore, differs from galaxy to galaxy.

5 CONCLUSIONS AND FUTURE PROSPECTS

We have used more than 4000 star-forming galaxies to probe the feasibility of reconstructing nebular emission lines using the stellar spectrum of the galaxy. We demonstrate that hydrogen Balmer emission lines of a galaxy can be well reconstructed based on its stellar contents. This result implies that stellar spectrum of a galaxy may serve as an SFR indicator, which is very useful when tracing the history of star formation of the host galaxy of the AGN.

However, the 10 PCs of Yip et al. (2004) are derived from observed galaxy spectra. Reddening of nebular emission lines and starlight cannot be properly taken into account using these PCs as templates. Neither can stellar populations be derived using the PCA-based approaches. Another problem is that the galaxy sample used to derive the PCs by Yip et al. include non-negligible fraction of narrow-emission line AGN, which may degrade the correlations between the nebular emission lines and the stellar features.

ICA is a new statistical technique for revealing hidden factors that underlie observational data sets. It can be taken as an extension of PCA but much more powerful. Using EL ICA, we can obtain reliably and simultaneously important spectral parameters of the galaxy, such as starlight reddening, stellar velocity dispersion, and stellar contents (Lu et al. 2004). It can be anticipated that the relation between stellar contents and nebular emission lines of galaxies can be adequately explored using this new technique. After properly taking account of dust reddening and carefully calibrating with other SFR indicators, starlight may serve as a reliable SFR indicator, which can be used to analyze the star formation history of the AGN host galaxy. We expect future studies along these lines.

Acknowledgements We thank the referee for many useful comments that lead to a significant improvement of the paper. This work was supported by the National Natural Science Foundation of China under No. 10233030, the Hundred Talents Program by CAS, and a key program of Chinese Science and Technology Ministry. We have made use of the data from the SDSS. Funding for the creation and the distribution of the SDSS Archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U. S. Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Princeton University, the United States Naval Observatory, and the University of Washington.

References

Abazajian K., et al., 2004, AJ, 128, 502
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Dong X., Zhou H., Wang T., Li C., Zhou Y., 2005, ApJ, 620, 629
González Delgado R. M., Heckman T., Leitherer C., 2001, ApJ, 546, 845
Kauffmann G., et al., 2003, MNRAS, 346, 1055
Kewley L. J., Geller M. J., Jansen R. A., 2004, AJ, 127, 2002
Hirashita H., Buat V., Inoue A. K., 2003, A&A, 410, 83
Hopkins A. M., et al., 2003, ApJ, 599, 971
Kennicutt R. C., 1998, ARA&A, 36, 189
Li C., Wang T., Zhou H., Dong X., Cheng F., 2005, AJ, 129, 669
Lu H., Zhou H., Wang T. et al., 2004, AJ, submitted
Pei Y. C., 1992, ApJ, 395, 130
Searle L., Sargent W. L. W., Bagnuolo W. G., 1973, ApJ, 179, 427
Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525
Stoughton C., et al., 2002, AJ, 123, 485
Tinsley B. M., 1968, ApJ, 151, 547
Yip C. W., et al., 2004, AJ, 128, 585
York D. G., et al., 2000, AJ, 120, 1579