The Astrophysical Journal, 721:875–885, 2010 September 20
doi:10.1088/0004-637X/721/1/875

LoCuSS: CALIBRATING MASS–OBSERVABLE SCALING RELATIONS FOR CLUSTER COSMOLOGY WITH SUBARU WEAK-LENSING OBSERVATIONS*

N. Okabe[1,2], Y.-Y. Zhang[3], A. Finoguenov[4,5], M. Takada[6], G. P. Smith[7], K. Umetsu[2], and T. Futamase[1]
[1] Astronomical Institute, Tohoku University, Aramaki, Aoba-ku, Sendai 980-8578, Japan; okabe@asiaa.sinica.edu.tw
[2] Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), P.O. Box 23-141, Taipei 10617, Taiwan
[3] Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
[4] Max-Planck-Institut für extraterrestrische Physik, Giessenbachstraße, 85748 Garching, Germany
[5] University of Maryland, Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA
[6] Institute for the Physics and Mathematics of the Universe (IPMU), The University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa City, Chiba 277-8568, Japan
[7] School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
Received 2010 March 23; accepted 2010 July 26; published 2010 September 1

ABSTRACT

We present a joint weak-lensing/X-ray study of galaxy cluster mass–observable scaling relations motivated by the critical importance of accurate calibration of mass proxies for future X-ray missions, including *eROSITA*. We use a sample of 12 clusters at $z \simeq 0.2$ that we have observed with Subaru and *XMM-Newton* to construct relationships between the weak-lensing mass (M) and three X-ray observables, gas temperature (T), gas mass ($M_{\rm gas}$), and quasi-integrated gas pressure ($Y_{\rm X}$), at overdensities of $\Delta = 2500$, 1000, and 500 with respect to the critical density. We find that $M_{\rm gas}$ at $\Delta \leqslant 1000$ appears to be the most promising mass proxy of the three because it has the lowest intrinsic scatter in mass at a fixed observable, $\sigma_{\ln M} \simeq 0.1$, independent of the cluster dynamical state. The scatter in mass at fixed T and $Y_{\rm X}$ is a factor of ∼2–3 larger than at fixed $M_{\rm gas}$, which are indicative of the structural segregation that we find in the M–T and M–$Y_{\rm X}$ relationships. Undisturbed clusters are found to be ∼40% and ∼20% more massive than disturbed clusters at fixed T and $Y_{\rm X}$, respectively, at $\sim 2\sigma$ significance. In particular, A 1914—a well-known merging cluster—significantly increases the scatter and lowers the normalization of the relation for disturbed clusters. We also investigated the covariance between the intrinsic scatter in M–$M_{\rm gas}$ and M–T relations, finding that they are *positively* correlated. This contradicts the adaptive mesh refinement simulations that motivated the idea that $Y_{\rm X}$ may be a low-scatter mass proxy, and agrees with more recent smoothed particle hydrodynamic simulations based on the Millennium Simulation. We also propose a method to identify a robust mass proxy based on principal component analysis. The statistical precision of our results is limited by the small sample size and the presence of the extreme merging cluster in our sample. We therefore look forward to studying a larger, more complete sample in the future.

Key words: cosmology: observations – dark matter – galaxies: clusters: general – gravitational lensing: weak – X-rays: galaxies: clusters

Online-only material: color figures

1. INTRODUCTION

Galaxy clusters are the largest virialized objects in the universe; they formed from high amplitude peaks of the primordial density field. Clusters therefore occupy the high mass exponential tail of the dark matter halo mass function, which is sensitive to the matter density and expansion history of the universe and to modifications of the laws of gravity. Measurements of the evolution of the galaxy cluster mass function across a broad range of redshifts can thus provide a powerful tool for constraining the cosmological parameters (e.g., Vikhlinin et al. 2009a, 2009b). Numerous galaxy cluster surveys will soon begin delivering a huge amount of data at optical, X-ray, and millimeter wavelengths, e.g., from Subaru/Hyper-Suprime-Cam, *eROSITA*, *SPT*, and *ACT*. One of the main goals of these surveys is to measure the evolution of the galaxy cluster mass function, and thus to probe the expansion history of the universe. However, the mass of a galaxy cluster is not directly measurable. These surveys will therefore rely on "mass-like" observables (e.g., X-ray temperature—Evrard et al. 1996) and scaling relations between these observables and mass, to construct the all-important mass functions. Calibration of mass–observable scaling relations is therefore currently a high-priority observational goal.

Traditionally, observational studies of the mass–observable scaling relations have relied solely on X-ray observations, typically concentrating on the mass–temperature relation (e.g., Finoguenov et al. 2001; Sanderson et al. 2003; Ettori et al. 2004; Arnaud et al. 2005). X-ray-based mass measurements require hydrostatic equilibrium (H.E.) and spherical symmetry to be assumed, and either a measurement of the temperature profile, or an assumption of isothermality. Inclusion of X-ray temperature information in both axes of the mass–temperature relation may therefore induce intrinsic correlations into the measured relation. The validity of the underlying assumptions also warrants careful testing.

Gravitational lensing offers cluster mass measurements that are independent of X-ray observations and do not rely on assuming H.E. Joint lensing/X-ray studies (e.g., Okabe & Umetsu 2008; Kawaharada et al. 2010) are therefore a promising route for calibrating cluster mass–observable scaling relations. Indeed, early lensing/X-ray studies of cluster cores indicated that the scatter in cluster temperature may be as large as 40% at fixed mass and that the scatter is dominated by disturbed,

---

* This work is based in part on data collected at the Subaru Telescope and obtained from the SMOKA, which is operated by the Astronomy Data Center, National Astronomical Observatory of Japan. Based on observations made with the *XMM-Newton*, an ESA science mission with instruments and contributions directly funded by ESA member states and the USA (NASA).

merging clusters, in which H.E. may not hold (Smith et al. 2005). Subsequent work has concentrated on using weak-lensing data to extend this pioneering work beyond cluster cores to overdensities of $500 \lesssim \Delta \lesssim 2500$ with respect to the critical density (Bardeau et al. 2007; Hoekstra 2007; Pedersen & Dahle 2007; Zhang et al. 2007, Zhang et al. 2008). The main limiting factors in these weak-lensing/X-ray studies have been the limited statistical precision and heterogeneity of the available weak-lensing data, and also the small samples observed to date.

On the theoretical side, Kravtsov et al. (2006) proposed the so-called quasi-integrated pressure, $Y_X \equiv M_{\rm gas} \times T$, as a "new robust low-scatter X-ray mass indicator" or a mass-like observable. This was motivated by analysis of their hydrodynamic numerical simulations of clusters using an adaptive mesh refinement (AMR) code. They found that the temperature deviations from the M–T relation are anti-correlated with the gas-mass deviations from the M–$M_{\rm gas}$ relation. This anti-correlation found in their simulations acts to suppress the scatter in the M–Y_X relation, independent of the dynamical state of the clusters. This prediction has stimulated much observational effort within the X-ray community that has broadly supported the idea that Y_X is the optimal X-ray mass proxy (e.g., Maughan 2007; Arnaud et al. 2007; Vikhlinin et al. 2009a).

However, Stanek et al.'s (2010) smoothed particle hydrodynamic (SPH) Millennium Gas Simulations contradict Kravtsov et al.'s simulations. Stanek et al. predict that the temperature and gas-mass deviations are positively correlated; this result appears to be independent of the range of gas physics (gravity-only, cooling, preheating) implemented in the simulations. Juett et al. (2010) have also recently suggested that previous X-ray-only studies may have underestimated the scatter in mass–observable scaling relations by a factor of ∼2–3. In summary, a joint lensing/X-ray observational investigation of the relationships between mass and gas mass, temperature, and Y_X is urgently needed. Such joint studies also lend themselves well to the task of observationally testing various corrections that have been derived from numerical simulations to account for deviations from H.E. For example, numerous authors have pointed out that H.E. mass estimates may underestimate the cluster mass because of non-thermal pressure support due to turbulence caused by bulk motion of the cluster gas (e.g., Evrard 1990; Rasia et al. 2006; Nagai et al. 2007; Piffaretti & Valdarnini 2008; Fang et al. 2009), and Vikhlinin et al. (2009a) applied a 17% upward correction to X-ray masses of disturbed clusters, based on the results of simulations.

A key goal of the Local Cluster Substructure Survey (LoCuSS[8]) is to calibrate cluster mass–observable scaling relations for future cosmological experiments. LoCuSS is a multi-wavelength survey of galaxy clusters at $0.15 < z < 0.3$ selected from the *ROSAT* All-sky Survey catalogs (Ebeling et al. 1998, 2000; Böhringer et al. 2004). To date, we have published the first lensing/Sunyaev–Zeldovich effect comparison (Marrone et al. 2009), begun our lensing/X-ray scaling relation work with a pilot study (Zhang et al. 2008), and compared lensing-based masses with H.E. masses on both small (Richard et al. 2010) and large (Zhang et al. 2010) scales. This article is a continuation of our pilot study (Zhang et al. 2008), in which we combined weak-lensing mass measurements from the Canada–France–Hawaii Telescope (Bardeau et al. 2005, 2007) and from the Nordic Optical Telescope and UH 88in (Dahle 2006) with *XMM-Newton* observations to calibrate the mass–observable scaling relations.

As alluded to above, Zhang et al.'s results were limited by the quality of the weak-lensing mass measurements because the underlying data were heterogeneous in the observing facilities, fields of view, and filters used. In this article, we address these issues by using our own weak-lensing mass measurements based on uniform analysis of our Subaru/Suprime-Cam observations (Okabe & Umetsu 2008; Okabe et al. 2010). Nevertheless, our Subaru/*XMM-Newton* sample remains small, at just 12 clusters. As we discuss throughout this article, sample size therefore remains an issue, and we will address this in a future article.

The outline of this paper is as follows. In Section 2, we briefly describe the weak-lensing and X-ray analysis and measure the dynamical state of each cluster using *XMM-Newton* data. We present the main results on the mass–observable scaling relations in Section 3, discuss the results in Section 4, and summarize our work in Section 5. Throughout this paper, we assume $\Omega_{m,0} = 0.3$, $\Omega_\Lambda = 0.7$, and $h = H_0/100\,{\rm km\,s^{-1}\,Mpc^{-1}} = 0.7$.

2. SAMPLE AND DATA ANALYSIS

2.1. Sample

For the purpose of this paper, we compiled a sample of 12 clusters—A 68, A 115, A 209, A 267, A 383, A 1835, A 1914, Z 7160, A 2261, RX J2129.6 + 0005, A 2390, and A 2631—that represents the overlap between the samples for which Subaru/Suprime-Cam and *XMM-Newton* data are available, and that we have previously published (Zhang et al. 2008; Okabe & Umetsu 2008; Okabe et al. 2010). The sample does not suffer, by design, any strong biases to extreme merging or extreme cool core clusters, and therefore can be regarded, qualitatively, as representative of massive, X-ray luminous clusters. However, given the small sample size, we refrain from attempting to quantify how these 12 might be biased with respect to the underlying cluster population in this article. Instead, this article presents some early results from our Subaru/*XMM-Newton* program that benefit from the use of our Subaru data, as opposed to the CFH12k/UH8k/NOT data that we used in Zhang et al. (2008). We defer a detailed discussion of sample definition and possible biases to future articles in this series that will address larger, more complete samples.

2.2. Weak-lensing Mass Measurements

The details of our weak-lensing analysis are described in by Okabe & Umetsu (2008) and Okabe et al. (2010); here, we provide a brief outline of some important aspects of our methods.

We selected background galaxies based on their location in the color-magnitude plane—typically $(V - i')/i'$—bluer or redder than the cluster red sequence by a minimum color offset (Umetsu & Broadhurst 2008; Umetsu et al. 2009; Okabe et al. 2010). As demonstrated by Okabe et al. (2010), contamination of the background galaxy catalogs by faint (unlensed) cluster members dilutes the weak-lensing signal. This effect is more pronounced at smaller clustercentric radii because the number density of cluster galaxies rises toward the cluster centers. In the absence of our color-selection techniques, weak-lensing M_{500} and M_{2500} measurements can be biased low by ∼20%–50%.

We used the COSMOS photometric redshift catalog (Ilbert et al. 2009) to estimate the redshift of the background galaxies. Specifically, we calculated the average lensing weight, $\langle D_{\rm LS}/D_{\rm OS}\rangle = \int_{z_d} dz dP_{\rm WL}/dz D_{\rm LS}/D_{\rm OS}$ (see also Equation (10) in Okabe et al. 2010), of each background galaxy catalog by selecting galaxies identical to both our catalogs and the COSMOS catalog. $D_{\rm OS}$ and $D_{\rm LS}$ are the angular diameter distances

---

[8] http://www.sr.bham.ac.uk/locuss

between the observer and source (background galaxy) and lens and source, respectively.

In cosmology the three-dimensional spherical mass, M_Δ, enclosed within a sphere of radius r_Δ for a given overdensity Δ is most relevant for the cluster mass function, where r_Δ is chosen such that the average density within the sphere is equal to the critical mass density at the cluster redshift, $\rho_{\rm cr}$, times the overdensity, Δ. We estimated M_Δ for each cluster by fitting the measured radial profile of lensing distortion signals to the Navarro–Frenk–White (NFW) model prediction parameterized by the mass M_Δ and c_Δ, where the NFW mass profile (Navarro et al. 1996, 1997) is given as $\rho \propto r^{-1}(1 + c_\Delta r/r_\Delta)^{-2}$ with c_Δ being the concentration parameter.

Describing cluster-scale dark matter halos as spherical objects may cause systematic errors in individual mass measurements because clusters are predicted to be triaxial in the collisionless CDM model (Jing & Suto 2002). For example, if the major axis of a triaxial halo is aligned with or perpendicular to the line of sight, a spherical model would overestimate or underestimate the mass, respectively, and also cause systematic errors in the measurement of the concentration parameter (Oguri et al. 2005; Gavazzi 2005; Corless et al. 2009). However, if the distribution of cluster orientations is random, then adopting spherical mass models should not introduce a significant bias into the properties of the sample. We therefore check that this is the case for our sample by comparing the spherical mass measurements from Okabe et al. (2010) that we use here with triaxial mass measurements of the same clusters using the same background galaxy catalogs from Oguri et al. (2010). On average, the spherical ($M_\Delta^{\rm (sph)}$) and triaxial ($M_\Delta^{\rm (tri)}$) masses agree well—$\langle M_\Delta^{\rm (tri)}/M_\Delta^{\rm (ave)}\rangle = 0.98 \pm 0.15, 0.90 \pm 0.17$, and 0.83 ± 0.21, for $\Delta = 500$, 1000, and 2500—confirming the expectation of negligible bias.

2.3. X-ray Observables

The observations and data reduction are described in detail by Zhang et al. (2007, 2008). In brief, the three mass proxies considered in this article are calculated as follows. The global temperature is a volume average of the spectrally measured, radial temperature profile limited to the radial range of $(0.2\text{–}0.5)r_{500}$. The gas mass $M_{\rm gas}(r)$ was obtained for each cluster by integrating a double-β model of the electron density that was fitted to the X-ray surface brightness profile. The quasi-integrated pressure is the product of the gas mass and the global temperature: $Y_{\rm X}(r) = M_{\rm gas}(r) \times T_{0.2\text{–}0.5r_{500}}$. Note that $M_{\rm gas}(r)$, $T_{0.2\text{–}0.5r_{500}}$, and $Y_{\rm X}(r)$ have all been calculated using radii obtained from the weak-lensing analysis and not using radii calculated from the X-ray analysis as in Zhang et al. (2008). This definition of radii introduces a subtle correlation with weak-lensing mass—we will explore this when estimating the intrinsic scatter in the mass–observable scaling relations in Section 3.6 and the Appendix. Finally, we adopted a self-consistent definition of the cluster centers based on the weak-lensing analysis. This caused us to change the centers of just two clusters—A 1914 and A 2631—from those used by Zhang et al. (2010).

2.4. X-ray Morphology and Dynamical State

Previous joint-lensing/X-ray studies have identified the dynamical state of clusters as a significant source of scatter in mass–observable scaling relations (Smith et al. 2005; Pedersen & Dahle 2007; Zhang et al. 2008, 2009). In this section, we therefore classify the clusters as either "disturbed" or "undisturbed," based on a new method patterned on those developed for the morphological classification of galaxies (e.g., Conselice 2003).

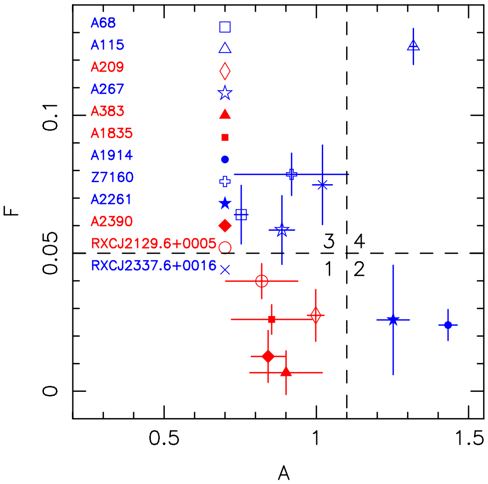

Figure 1. Asymmetry vs. fluctuation parameters using the $\leqslant r_{500}$ region. We define undisturbed clusters (red, low A and F parameters) and disturbed clusters (blue, high A parameter or high F one).

(A color version of this figure is available in the online journal.)

We calculate the asymmetry (A) and fluctuation (F) of the X-ray surface brightness distribution in the 0.7–2 keV band. Asymmetry is defined as $A = (\sum_{ij} |I_{ij} - R_{ij}|)/\sum_{ij} I_{ij}$, the normalized sum of the absolute value of the flux residuals, where I_{ij} is a matrix element of the combined MOS1+MOS2 *XMM-Newton* frame in the 0.7–2.0 keV band, flat-fielded, point-source subtracted, and refilled assuming a Poisson distribution, and R_{ij} is a matrix element obtained by rotating the above frame by 180°. The pixel size of both frames is $4'' \times 4''$. The fluctuation, F, measures deviations from a smooth flux distribution and is defined as $F = (\sum_{ij} I_{ij} - B_{ij})/\sum_{ij} I_{ij}$, where B_{ij} is an element in a frame smoothed on 2 arcmin scales, which corresponds to a physical scale of 400 kpc at $z = 0.2$. Such smoothing also suppresses the effect of the complex shape of the *XMM-Newton* point-spread function (Ghizzardi 2001). We estimate the statistical errors of A and F assuming Poisson noise computed within a radius of r_{500}, excluding CCD gaps and bad pixels. We also estimate the systematic error of A caused by uncertainties in the cluster centers by recalculating A, each time moving the cluster centers onto one of the neighboring pixels within the $r \leqslant 4''$ circle from the nominal cluster center.

The clusters span the ranges $A \sim 0.07\text{–}0.15$ and $F \sim 0\text{–}0.14$ (Figure 1). Dynamically disturbed clusters generally have an asymmetric X-ray morphology, with an offset between optical and X-ray centers, and are therefore expected to have larger A and F than undisturbed clusters. To separate the clusters into two subsamples that represent relatively disturbed and undisturbed systems, we subdivided the A–F plane into four quadrants: (1) $A < 1.1$ and $F < 0.05$—RX J2129, A 209, A 383, A 1835, and A 2390, (2) $A > 1.1$ and $F < 0.05$—A 2261 and A 1914, (3) $A < 1.1$ and $F > 0.05$ — A 68, A 2631, A 267, and Z 7160, and (4) $A > 1.1$ and $F > 0.05$ — A 115. We classify the five clusters in quadrant (1)—low A and low F—as undisturbed clusters, and the remaining seven as disturbed clusters. It is immediately obvious that this classification matches other possible classification schemes well. For example, four of the five undisturbed clusters host a cool core (e.g., Smith et al. 2003; Allen et al. 2001; Peterson et al. 2003), and the disturbed clusters have been discussed extensively as merging/cold-front clusters (e.g., Okabe & Umetsu 2008; Mazzotta & Giacintucci 2008; Gutierrez & Krawczynski 2005), in which complicated temperature/entropy distributions or large offsets between lensing/optical and X-ray centroids exist (e.g., Finoguenov et al. 2005; Smith et al. 2005; Sanderson et al. 2009b). In summary, all of the clusters identified as disturbed in the A–F plane are independently confirmed as disturbed by other methods in the literature. However, we stress again the relative nature of the disturbed/undisturbed classification and acknowledge that the disturbed clusters in particular likely comprise clusters in a wide variety of stages in their dynamical evolution. We will return to this issue later when we assess the impact of a single extreme merging cluster on our attempts to calibrate the mass–observable scaling relations.

3. RESULTS

In this section, we present the main empirical results of the slope, normalization, and intrinsic scatter in the mass–observable scaling relations and how these depend on the dynamical state of the clusters. We also discuss the correlation between gas mass and temperature deviations.

3.1. Scaling Relations and Fitting Methods

If gravitational heating is the dominant mechanism responsible for the X-ray properties of galaxy clusters, the following scaling relations are expected to hold:

$$ME(z) \propto (Y_X E(z))^{3/5}\, h^{1/2}, \quad (1)$$

$$ME(z) \propto M_{\rm gas} E(z)\, h^{3/2}, \quad (2)$$

$$ME(z) \propto T^{3/2}\, h^{-1}, \quad (3)$$

where M, $M_{\rm gas}$, and T are the total mass, gas mass, and temperature of a cluster, respectively, and $Y_X = M_{\rm gas} \times T$ is the quasi-integrated pressure. These relations, specifically the exponents of M, $M_{\rm gas}$, and T, are usually referred to as self-similar, following Kaiser (1986). Note that the term $E(z) = H(z)/H_0 = [\Omega_{m,0}(1+z)^3 + \Omega_\Lambda]^{1/2}$ accounts for the redshift evolution of the clusters in a flat universe.

In the following subsections, we therefore fit the functional form $M_z = M_0 X_z^\gamma$ to the data, where $M_z = ME(z)$, M_0 is the normalization, X_z is the X-ray observable (i.e., Y_X, T, or $M_{\rm gas}$) multiplied by $E(z)$ or not as appropriate based on Equations (1)–(3), and γ is the logarithmic slope. These fits are done at three overdensities with respect to the critical density: $\Delta = 2500, 1000$, and 500. The scaling relation slope and normalization measurements are based on orthogonal regression performed using the Orthogonal Distance Regression package (ODRPACK; e.g., Boggs et al. 1987) taking into account the measurement errors. In general, we ignore the subtle correlations introduced by measuring the X-ray observables within radii defined by the weak-lensing analysis, although we do take them into account in Section 3.3 when we measure the intrinsic scatter. To check for consistency with other work, we have also refitted the relations using the bisector modification of the BCES method (Akritas & Bershady 1996). The difference of the best-fit scaling relation parameters between the two fitting methods is a small fraction of statistical uncertainties. For example, the difference of the best-fit slopes and normalizations between the two methods is typically ∼30% and ∼6% of the statistical error, respectively. We also did the bootstrap resampling to estimate the sample variance on the slope parameter and found that it is $\lesssim$20% of the statistical errors.

Table 1
Slopes of the Mass–Observable Relations for the Full Sample

Relation	γ Using $M_{\rm WL}$		
	Δ:500	1000	2500
(1)	(2)	(3)	(4)
M_Δ–Y_X^γ	0.67 ± 0.14	0.59 ± 0.11	0.46 ± 0.11
M_Δ–$T^\gamma_{0.2\text{–}0.5r_{500}}$	1.49 ± 0.58	1.49 ± 0.46	1.26 ± 0.35
M_Δ–$M^\gamma_{\rm gas}$	0.98 ± 0.15	0.86 ± 0.14	0.68 ± 0.15

Notes. A single power-law form of slope γ is considered. The X-ray temperature is derived by the volume average of the deprojected radial temperature profile and using the radii within 0.2–0.5r_{500} ($T = T_{0.2\text{–}0.5r_{500}}$; see Zhang et al. 2008). Column 1: scaling relations. Columns 2–4: slopes of the mass–observable relations using weak-lensing masses at the overdensities of $\Delta = 500$, 1000, and 2500, respectively.

3.2. Slope and Normalization

We first fit the scaling relations to the full sample of 12 clusters with both slope γ and normalization M_0 as free parameters. At $\Delta = 500$, the best-fit slopes of all three relations agree well with the self-similar model (Table 1). At higher overdensities, the agreement deteriorates for all three relations, and indeed the slopes of the M_{2500}–$M_{\rm gas}$ and M_{2500}–Y_X relations are discrepant from self-similarity at ∼2σ–3σ at $\Delta = 2500$. This flattening in the scaling relations at higher Δ can also be seen graphically in Figures 2, 3, and 4, in which we show the M_Δ–Y_X, M_Δ–T_X, and M_Δ–$M_{\rm gas}$ relations, respectively.

To constrain the normalization parameter M_0, we fix the slope parameters to the self-similar values and repeat the fits. The measured normalizations are all consistent with those obtained by Zhang et al. (2008) using the same *XMM-Newton* data and independent weak-lensing data. The superior quality and uniformity of our Subaru data show differences between the normalizations for disturbed and undisturbed clusters. These differences are most pronounced at $\Delta = 500$ (see Table 2); specifically, at fixed Y_X undisturbed clusters are measured to be ∼22% more massive than disturbed clusters at ∼1.5σ significance. Similarly, at fixed T, undisturbed clusters are measured to be ∼43% more massive than disturbed clusters at ∼1.8σ significance. We confirm that our results are insensitive to whether or not the slopes are fixed to the self-similar value.

3.3. Scatter

We also measured the intrinsic scatter, $\sigma_{\ln M}$, for the logarithm of the Y-axis, $ME(z)$, for each mass–observable scaling relation using the Bayesian method described in the Appendix. Here we take into account the correlations caused by measuring X-ray observables within radii defined by the lensing analysis—see the Appendix. We also confirmed that the best-fit slopes and normalizations obtained using the ODR methods discussed above are consistent within errors with those obtained using the more sophisticated Bayesian method considered here. The intrinsic scatter in all three relations is well described by a lognormal distribution.

The M–T relation exhibits the largest intrinsic scatter (∼0.23–0.33; Table 3) among the three mass–observable

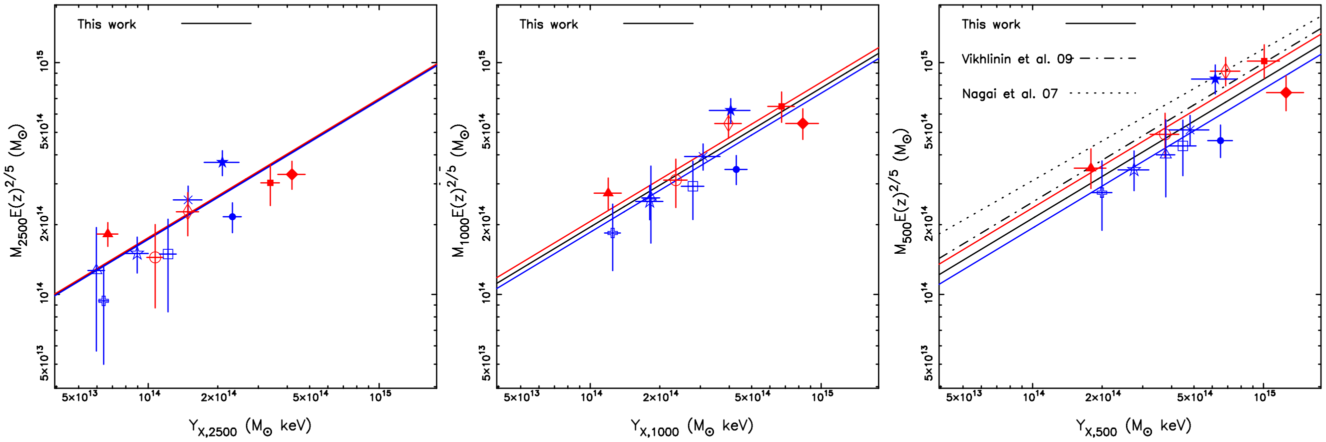

Figure 2. M–Y_X relation using weak-lensing masses and the quasi-integrated pressure at the overdensities of $\Delta = 2500$ (left panel), $\Delta = 1000$ (middle panel), and $\Delta = 500$ (right panel); see also Section 2.3. The solid black, red, and blue lines denote the best fits of the relations using weak-lensing masses for all, undisturbed, and disturbed clusters, respectively.

(A color version of this figure is available in the online journal.)

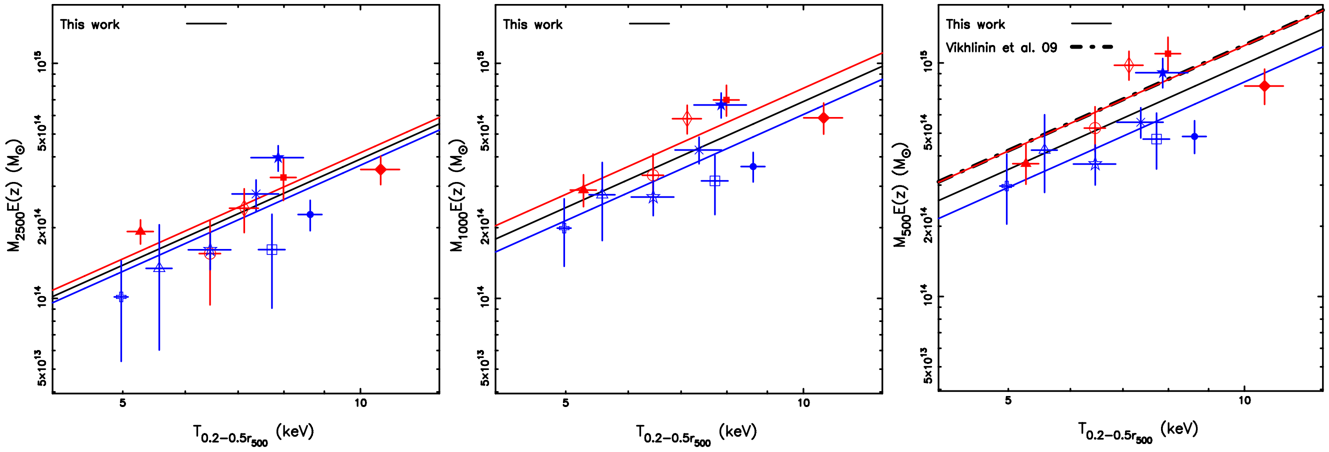

Figure 3. $M^{\rm WL}_\Delta$–T relation. The arrangement of panels with overdensity and the line coding are the same as in Figure 2.

(A color version of this figure is available in the online journal.)

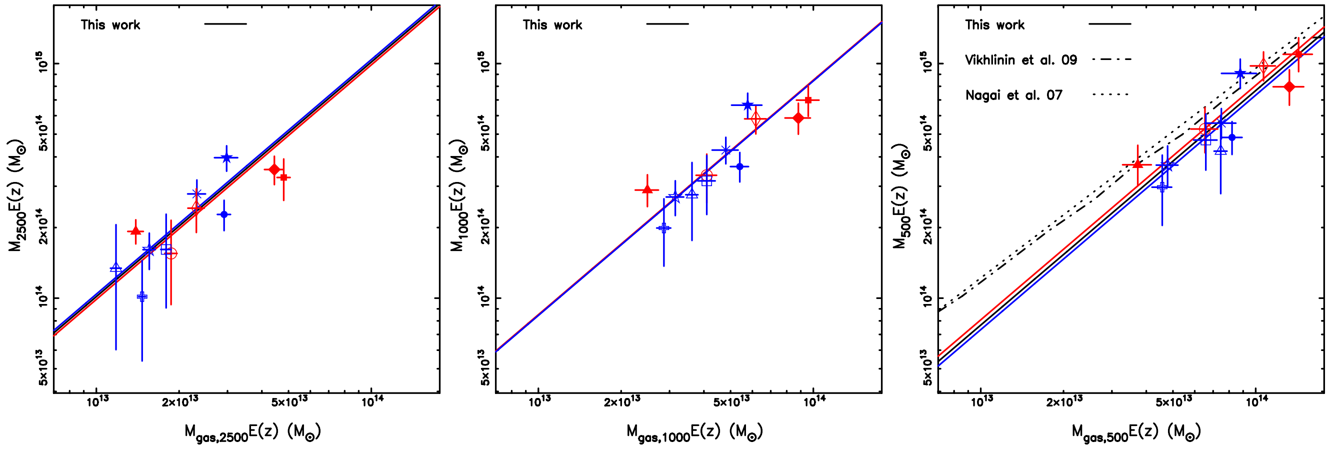

Figure 4. $M^{\rm WL}_\Delta$–$M_{\rm gas}$ relation. The arrangement of panels with overdensity and the line coding are the same as in Figure 2.

(A color version of this figure is available in the online journal.)

relations. We also observe an increase in the intrinsic scatter with increasing radius (i.e., decreasing the interior overdensity Δ). The same trend is found in undisturbed clusters, while the opposite trend is found in disturbed clusters. However, this trend is not a physical feature of the intracluster gas affected by gravitational heating because we used a fixed global temperature measurement in the radial range of 0.2–$0.5r_{500}$ for all the overdensities.

The M–$M_{\rm gas}$ relation is the tightest of the three, with an intrinsic scatter in mass of $\sigma_{\ln M} \sim 0.12$–0.16 at $\Delta = 500$ and 1000. At $\Delta = 2500$, the scatter is roughly double that at lower overdensity (Table 3), which may be due to different

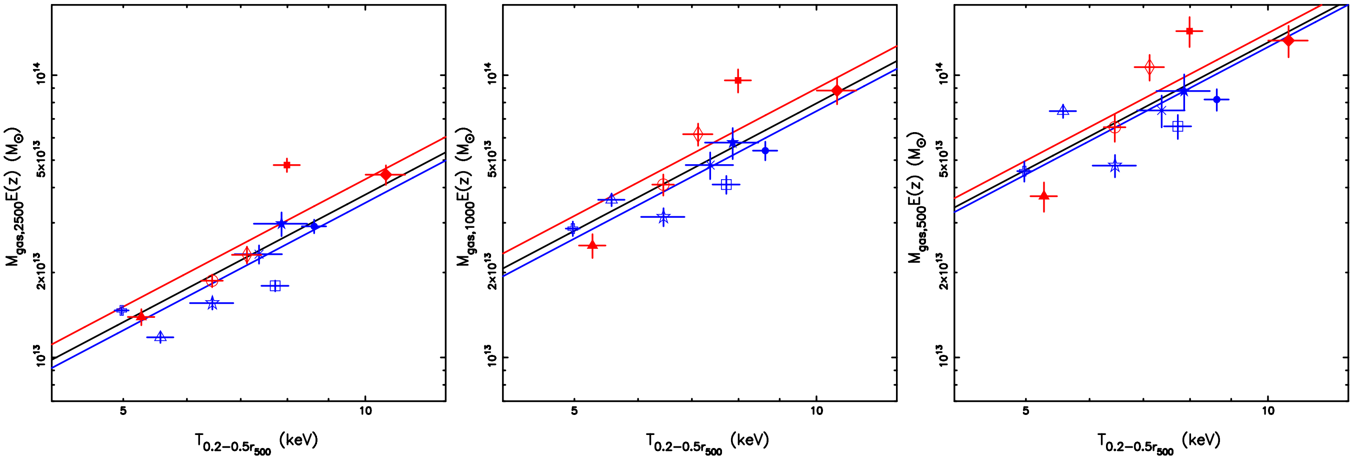

Figure 5. $M_{\rm gas,\Delta}$–T relation. The arrangement of panels with overdensity and the line coding are the same as in Figure 2.

(A color version of this figure is available in the online journal.)

Table 2

Normalization (M_0) and Morphological Dependence of the Mass–Observable Relations

Relation	All 12 Clusters			w/o A1914	
	All	Undisturbed	Disturbed	All	Disturbed
(1)	(2)	(3)	(4)	(5)	(6)
M_{500}–$Y_{\rm X}$	$4.92^{+0.35}_{-0.33}$	$5.46^{+0.57}_{-0.52}$	$4.47^{+0.41}_{-0.37}$	$5.16^{+0.34}_{-0.32}$	$4.83^{+0.41}_{-0.38}$
M_{1000}–$Y_{\rm X}$	$4.51^{+0.29}_{-0.27}$	$4.77^{+0.43}_{-0.40}$	$4.28^{+0.39}_{-0.36}$	$4.70^{+0.28}_{-0.26}$	$4.62^{+0.41}_{-0.37}$
M_{2500}–$Y_{\rm X}$	$4.02^{+0.29}_{-0.27}$	$4.06^{+0.49}_{-0.43}$	$3.99^{+0.41}_{-0.37}$	$4.20^{+0.30}_{-0.28}$	$4.38^{+0.43}_{-0.39}$
M_{500}–$T_{0.2\text{–}0.5r_{500}}$	$2.45^{+0.27}_{-0.24}$	$2.94^{+0.50}_{-0.42}$	$2.05^{+0.24}_{-0.22}$	$2.34^{+0.13}_{-0.12}$	$2.20^{+0.10}_{-0.10}$
M_{1000}–$T_{0.2\text{–}0.5r_{500}}$	$1.70^{+0.15}_{-0.14}$	$1.94^{+0.24}_{-0.21}$	$1.50^{+0.18}_{-0.16}$	$1.80^{+0.15}_{-0.14}$	$1.64^{+0.19}_{-0.17}$
M_{2500}–$T_{0.2\text{–}0.5r_{500}}$	$0.97^{+0.08}_{-0.07}$	$1.03^{+0.10}_{-0.10}$	$0.91^{+0.12}_{-0.10}$	$1.02^{+0.08}_{-0.07}$	$1.00^{+0.14}_{-0.12}$
M_{500}–$M_{\rm gas}$	$13.10^{+0.77}_{-0.73}$	$13.79^{+1.23}_{-1.13}$	$12.53^{+1.03}_{-0.96}$	$13.59^{+0.77}_{-0.73}$	$13.36^{+1.09}_{-1.01}$
M_{1000}–$M_{\rm gas}$	$14.46^{+0.86}_{-0.81}$	$14.52^{+1.52}_{-1.38}$	$14.41^{+1.13}_{-1.05}$	$14.91^{+0.89}_{-0.84}$	$15.36^{+1.14}_{-1.06}$
M_{2500}–$M_{\rm gas}$	$17.33^{+1.36}_{-1.26}$	$16.85^{+2.54}_{-2.21}$	$17.76^{+1.65}_{-1.51}$	$18.01^{+1.46}_{-1.35}$	$19.45^{+1.51}_{-1.40}$

Notes. The forms of M–$Y_{\rm X}$, M–$T_{\rm X}$, and M–$M_{\rm gas}$ relations are given by $M_\Delta E(z)^{2/5} = M_0(Y_{\rm x}/3 \times 10^{14} M_\odot\,{\rm keV})^{5/3} \times 10^{14}h^{1/2}M_\odot$, $M_\Delta E(z) = M_0(k_BT/5\,{\rm keV})^{3/2} \times 10^{14}h^{-1}M_\odot$, and $M_\Delta E(z) = M_0(M_{\rm gas}E(z))h^{3/2}M_\odot$, respectively. We fix the slopes to the self-similar values. The X-ray temperature is derived by the volume average of the radial temperature profile in the range of 0.2–$0.5r_{500}$ ($T = T_{0.2\text{–}0.5r_{500}}$; see Zhang et al. 2008). Column 1: scaling relations. Columns 2–4: normalization of the fit to the relation using weak-lensing masses. The results for the full sample, undisturbed, and disturbed clusters are presented. Columns 5 and 6: normalization of the mass–observable relations for the 11 clusters and 6 disturbed clusters, excluding A1914, respectively.

Table 3

Intrinsic Scatter in Mass–Observable Relations

Relation	All 12 Clusters			w/o A1914	
	All	Undisturbed	Disturbed	All	Disturbed
(1)	(2)	(3)	(4)	(5)	(6)
M_{500}–$Y_{\rm X}$	$0.203^{+0.066}_{-0.095}$	$\leqslant 0.283$	$0.216^{+0.098}_{-0.166}$	$0.154^{+0.071}_{-0.098}$	$\leqslant 0.225$
M_{1000}–$Y_{\rm X}$	$0.203^{+0.052}_{-0.078}$	$\leqslant 0.283$	$0.241^{+0.084}_{-0.166}$	$0.173^{+0.057}_{-0.081}$	$0.243^{+0.089}_{-0.204}$
M_{2500}–$Y_{\rm X}$	$0.245^{+0.053}_{-0.088}$	$\leqslant 0.353$	$0.306^{+0.091}_{-0.191}$	$0.245^{+0.058}_{-0.097}$	$0.329^{+0.113}_{-0.280}$
M_{500}–$T_{0.2\text{–}0.5r_{500}}$	$0.327^{+0.087}_{-0.136}$	$\leqslant 0.503$	$0.273^{+0.123}_{-0.215}$	$0.288^{+0.093}_{-0.133}$	$\leqslant 0.280$
M_{1000}–$T_{0.2\text{–}0.5r_{500}}$	$0.267^{+0.078}_{-0.113}$	$\leqslant 0.364$	$0.292^{+0.120}_{-0.217}$	$0.228^{+0.088}_{-0.125}$	$\leqslant 0.318$
M_{2500}–$T_{0.2\text{–}0.5r_{500}}$	$0.226^{+0.084}_{-0.117}$	$\leqslant 0.295$	$0.328^{+0.135}_{-0.244}$	$0.199^{+0.090}_{-0.132}$	$\leqslant 0.411$
M_{500}–$M_{\rm gas}$	$0.123^{+0.068}_{-0.102}$	$\leqslant 0.260$	$\leqslant 0.226$	$0.109^{+0.062}_{-0.106}$	$\leqslant 0.225$
M_{1000}–$M_{\rm gas}$	$0.160^{+0.070}_{-0.096}$	$\leqslant 0.330$	$0.207^{+0.103}_{-0.188}$	$0.147^{+0.072}_{-0.104}$	$\leqslant 0.225$
M_{2500}–$M_{\rm gas}$	$0.241^{+0.077}_{-0.114}$	$0.436^{+0.190}_{-0.410}$	$0.263^{+0.121}_{-0.230}$	$0.240^{+0.086}_{-0.127}$	$\leqslant 0.248$

Notes. Intrinsic scatter, $\sigma_{\ln M}$, in the mass–observable relations, using weak-lensing masses, for all 12 clusters and the 11 clusters excluding A1914 (Section 3.4), respectively. We refer to the mean of the posterior probability distribution of each parameter. Column 1: scaling relations; Columns 2–4: intrinsic scatter for all 12 clusters, undisturbed, and disturbed clusters, respectively. Columns 5 and 6: intrinsic scatter for the 11 clusters and 6 disturbed clusters, excluding A1914, respectively.

core properties of individual clusters. For example, cool-core clusters have denser, cuspier cores than non-cool core clusters (e.g., Croston et al. 2008; Sanderson et al. 2009a; McCarthy et al. 2008). Such differences between cluster cores have a much smaller effect on measurements at larger radii because the core regions make a small contribution to the total gas mass measured out to $\Delta = 1000$ and 500. However, note that the intrinsic scatter is not well constrained for M–$M_{\rm gas}$ because the scatter is dominated by statistical errors. A larger sample is clearly needed to improve the constraints on the intrinsic scatter in M–$M_{\rm gas}$, however, it is important to note that this is the only relation that appears to have ~10% intrinsic scatter.

The observed intrinsic scatter in the M–$Y_{\rm X}$ relation is intermediate between that of the M–$M_{\rm gas}$ and M–T relations, at $\sigma_{\ln M} \sim 0.20$–0.25 (Table 3). This is a factor of $\gtrsim 2$ greater than that originally predicted by Kravtsov et al. (2006) based on their AMR simulations.

3.4. *The Impact of an Outlier*

In this section, we highlight the impact of one cluster, A 1914, on our results. This cluster has previously been identified as a merging cluster with a complex X-ray morphology, radio halo, and weak-lensing-based dark matter distribution (Buote & Tsai 1996; Bacchi et al. 2003; Govoni et al. 2004; Okabe & Umetsu 2008). We have also identified it as having the most extreme X-ray/lensing mass discrepancy among the 12 clusters considered here (Zhang et al. 2010).

To assess the impact of such clusters on the measured reliability of X-ray observables as mass proxies, we repeated the calculations of normalization and scatter discussed in Sections 3.2 and 3.3 excluding A 1914 (Table 2). At $\Delta = 500$, the normalization of the M–$Y_{\rm X}$ relations for disturbed and undisturbed clusters are different at just $\sim 1.2\sigma$ significance when A 1914 is excluded from the disturbed sample, in contrast to the $\sim 1.8\sigma$ difference based on the full sample of 12 clusters. We also find that excluding A 1914 reduces the intrinsic scatter on all of the scaling relations. In particular, the intrinsic scatter on M–$Y_{\rm X}$ is reduced by ~25% from $\sigma_{\ln M} \sim 0.20$ to 0.15. Jackknife tests on samples of 11 clusters (i.e., removing each cluster in turn) also confirm that A 1914 is indeed the most significant outlier among our sample.

These results indicate that outliers in the cluster population require careful treatment in the construction and application of mass–observable scaling relations. In summary, reliable cluster selection functions are required to gain robust constraints. This will be especially true for future high-redshift surveys because the fraction of merging clusters is expected to increase with look-back time (Vikhlinin et al. 2009a).

3.5. *Covariance of Deviations*

We investigate the covariance of deviations from the best-fit M–$M_{\rm gas}$ and M–T relations following recent numerical simulation studies (e.g., Kravtsov et al. 2006; Stanek et al. 2010). For a given *mean* scaling relation $Y = f(X)$, the deviations of each cluster from the mean relation are quantified as $\delta Y \equiv [Y - f(X)]$ and $\delta X \equiv [X - f^{-1}(Y)]$. We use the mean normalizations for a full sample of 12 clusters (Figure 5), however, we found that the following results do not change significantly when the best-fit normalizations of undisturbed and disturbed clusters are used instead.

The temperature and $M_{\rm gas}$ deviations, $\delta T/T(M_\Delta)$ and $\delta M_{\rm gas}/M_{\rm gas}(M_\Delta)$, appear to be positively correlated (Figure 6). We test this quantitatively using Spearman's rank correlation coefficient test, obtaining $r_s = 0.531 \pm 0.009$. The probability of obtaining a value of r_s greater than or equal to the measured value is low: $\mathcal{P} = 0.075 \pm 0.006$. This test therefore indicates that the positive correlation is significant. However, the apparent positive correlation between the temperature and gas mass deviations does not show the correlation between intrinsic scatter, but between total scatter, which is a convolution of measurement errors and intrinsic scatter, because we did not take into account the measurement uncertainties here. When dealing with observational constraints on scaling relations, it is therefore essential to include both the covariance of intrinsic scatter and the measurement errors with which the scatter is convolved in robust calculations.

3.6. *Covariance of Intrinsic Scatter*

We simultaneously fit M–T and M–$M_{\rm gas}$ relations and measure the covariance of the intrinsic scatter using a multi-dimensional fitting method described in the Appendix. This method considers not only the matrix of the observational errors for individual clusters, $\mathbf{\Sigma}_{{\rm obs},i}$, but also the covariance matrix of the intrinsic scatter, $\mathbf{\Sigma}_{\rm int}$. The covariance of the intrinsic scatter is given by

$$\mathbf{\Sigma}_{\rm int} = \begin{pmatrix} \sigma_t^2 & \sigma_{tg} \\ \sigma_{tg} & \sigma_g^2 \end{pmatrix}, \tag{4}$$

where σ_t^2 and σ_g^2 are variances for the logarithm of temperature and gas mass, respectively, and $\sigma_{tg} = r\sigma_t\sigma_g$ is a covariance with a coefficient r. Here, we do not need to take into account the intrinsic scatter on mass, because the gas properties, under a cluster mass given by the cosmology, only have intrinsic scatter physically due to the gas evolution. When we estimate cluster masses from X-ray observations via scaling relations, there is intrinsic scatter on mass due to the propagation from the intrinsic scatter of gas. As shown in the Appendix, the observational error matrix for individual clusters is given by

$$\mathbf{\Sigma}_{\rm obs} = \begin{pmatrix} \frac{4}{9}e_m^2 + e_t^2 & \frac{2}{3}\left(1 - \frac{\partial \ln M_{\rm gas}}{\partial \ln M}\right)e_m^2 \\ \frac{2}{3}\left(1 - \frac{\partial \ln M_{\rm gas}}{\partial \ln M}\right)e_m^2 & \left(1 - \frac{\partial \ln M_{\rm gas}}{\partial \ln M}\right)^2 e_m^2 + e_g^2 \end{pmatrix},$$

where e_m, e_t, and e_g are observational errors for the logarithm of the mass, temperature, and gas mass, respectively. The coefficients of e_m are the slopes of the mass–observable relations $T \propto M^{2/3}$ and $M_{\rm gas} \propto M$, and a term to account for the propagation of the error on r_Δ (derived from the lensing analysis) to the error on the X-ray observables, $\partial \ln M_{\rm gas}/\partial \ln M$ (see the Appendix). Since the correlation between the observational errors of temperature and gas mass is negligible, we set them to zero. We measure the intrinsic scatter for the M–$M_{\rm gas}$ and M–T relations (σ_g, σ_t, and r) with their slopes set to the self-similar prediction. The likelihood (Equation (A2)) is given in the Appendix. We explore parameter space using the Metropolis–Hastings algorithm, as described in the Appendix, restricting ranges of values explored to $0 \leqslant \sigma_t \leqslant 1$, $0 \leqslant \sigma_g \leqslant 1$, and $|r| \leqslant 1$.

We first impose a flat prior on σ_t, σ_g, and r within the limits referred to above, and obtain measurements of intrinsic scatter that are in good agreement with those obtained by fitting the M–$M_{\rm gas}$ and M–T relations independently: $\sigma_g = 0.132^{+0.066}_{-0.086}$ and $\sigma_t = 0.213^{+0.055}_{-0.080}$ (Table 4). However, the coefficient r is not well determined due to the tail of the posterior probability

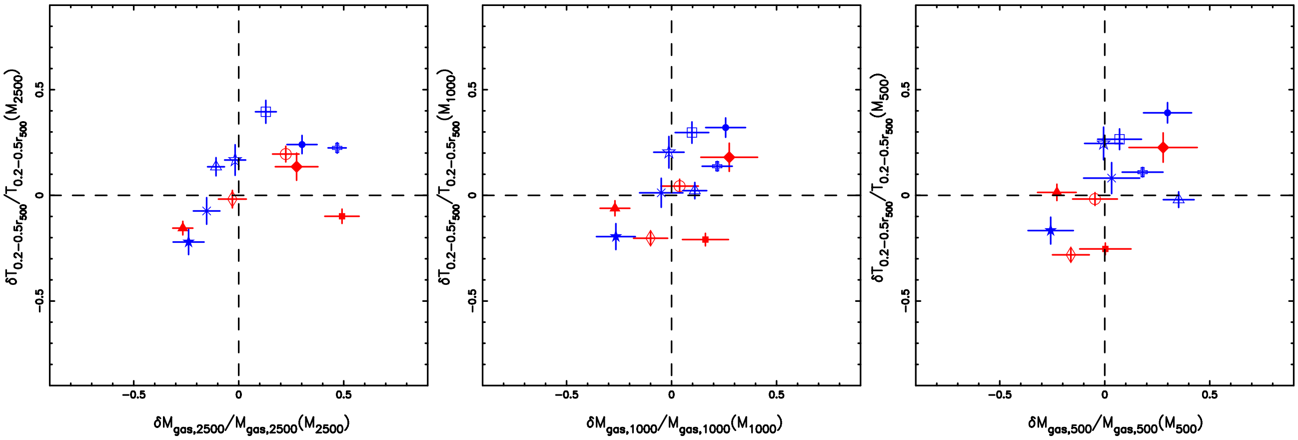

Figure 6. Normalized temperature deviations from M–T vs. normalized gas mass deviations from M–$M_{\rm gas}$, from the best-fit scaling relations for the full sample at the radii with $\Delta = 2500$ (left), $\Delta = 1000$ (middle), and $\Delta = 500$ (right), respectively. The colors and symbols have the same meaning as in Figure 2.

(A color version of this figure is available in the online journal.)

Table 4
Multi-dimensional Fitting for Covariance of Intrinsic Scatter of Gas Mass and Temperature at $\Delta = 500$

12 Clusters (1)	Flat Prior (2)	Gaussian Prior (3)
σ_g	$0.132^{+0.066}_{-0.086}$	$0.121^{+0.072}_{-0.061}$
σ_t	$0.213^{+0.055}_{-0.080}$	$0.206^{+0.043}_{-0.051}$
r	$\geqslant 0.185$	$\geqslant 0.170$
$\mathcal{P}(\lvert r'\rvert \geqslant \lvert r\rvert)$	$\leqslant 0.565$	$\leqslant 0.597$

Notes. The model parameters in the covariance of intrinsic scatter for gas mass and temperature derived from the multi-dimensional fitting for the full sample. We refer to the mean of the posterior distribution of each parameter. The lower limit is at a 68.3% confidence level. Column 1: model parameters. Column 2: results with the flat prior, $p_{\rm prior} = 1$. Column 3: results with the Gaussian prior with the best-fit intrinsic scatter derived from the independent measurement for the mass–observable relation. $\mathcal{P}(|r'| \geqslant |r|)$ denotes the probability that the correlation coefficient of the two random variables for 12 pair realizations is higher than the observed one.

distribution extending to negative values. Nevertheless, we derive a 68.3% confidence lower limit of $r \gtrsim 0.185$. We also measure the mode of the marginalized posterior probability distribution to be $r = 0.575$.

We then repeat the fit, this time using Gaussian priors centered on the best-fit measurements of σ_t and σ_g. As shown in the second column in Table 4, all resulting parameters are consistent with those in the flat prior. This time the mode of the posterior probability distribution is $r = 0.570$, and we find that the lower limit on r is again positive—i.e., it is positive independent of the prior. The positive coefficient leads to a large intrinsic scatter of $Y_{\rm X} = M_{\rm gas}T$ because the last term in $\sigma_{Y_{\rm X}}^2 = \sigma_g^2 + \sigma_t^2 + 2r\sigma_t\sigma_g$ becomes positive. The positive coefficient indicates that deviations in gas mass and temperature are partially correlated.

However, we stress that the small size of our observed sample severely limits the statistical power of our results—specifically, we cannot rule out the possibility that the gas mass and temperature deviations are two random variables that are correlated by accident. To quantify this, we calculated the probability, $\mathcal{P}(|r'| \geqslant |r|)$, that the correlation coefficient of the two random variables in a sample of 12 drawings, r', is higher than the observed value (Pugh & Winslow 1966). Since the probability is non-negligible ($\mathcal{P}(|r'| \geqslant |r|) < 0.565$ and 0.597; Table 4), we cannot yet completely exclude this possibility. Therefore, we need to increase the sample size before we can make definitive statements on the correlation of intrinsic scatter.

4. DISCUSSION

4.1. *Comparison with Previous Observations*

First we compare our results with those of our pilot study (Zhang et al. 2008). The main difference between Zhang et al.'s analysis and that presented here is that Zhang et al.'s weak-lensing mass measurements were drawn from the literature, and thus suffered heterogeneous image quality, observed depth, and systematic uncertainties relating to background galaxy selection and faint galaxy shape measurement. Despite these differences, the overall normalization of our mass–observable scaling relations agree within the uncertainties with those of Zhang et al. (2008). However, despite our sample being roughly a factor of two smaller than that of Zhang et al., we detect structural segregation in the M–$Y_{\rm X}$ and M–T relations at $\sim 2\sigma$ significance. Our ability to make this detection is likely due to the factor of $\gtrsim 2$ smaller statistical errors on weak-lensing mass measurements of individual clusters, thanks to the superb quality of our Subaru data.

This is the first time that structural segregation has been found in the M–$Y_{\rm X}$ relation; however, it has been detected in the M–T relation at a similar amplitude, and level of significance in previous joint lensing/X-ray studies (Smith et al. 2005; Pedersen & Dahle 2007). Theoretical studies (e.g., Randall et al. 2002) suggest that this segregation may be caused by cluster–cluster mergers boosting the temperature of disturbed merging clusters. However, orientation effects may also contribute to systematic errors in weak-lensing mass measurements that exaggerate the segregation (Meneghetti et al. 2010). We therefore defer physical interpretation of the observed segregation to a future careful investigation of the degeneracy between cluster orientation and residuals on mass–observable scaling relations.

We also compare the normalizations of our M–$Y_{\rm X}$ and M–T relations at $\Delta = 500$ with the same from Vikhlinin et al.'s (2009a) X-ray-only study of "relaxed" clusters with *Chandra*. Note that, as discussed in Section 3.1, our results are insensitive to whether we use orthogonal regression or bisector fitting techniques; Vikhlinin et al. used the bisector method. The normalizations of our M–$Y_{\rm X}$ and M–T relations for undisturbed clusters agree within the uncertainties with Vikhlinin et al.'s (2009a) relaxed clusters (Figures 2 and 3). This suggests that orientation effects may not be a major influence on the normalization of our undisturbed cluster scaling relations. On the other hand, it is thus clear that the normalizations of our disturbed cluster scaling relations differ from Vikhlinin et al.'s results at $\sim 2\sigma$ significance. It will therefore be important in the future to study the mass–observable scaling relations with large samples comprising clusters that span a wide range of dynamical states. This need will become more acute at high redshift because the fraction of disturbed clusters likely increases toward higher redshifts.

4.2. Comparison with Simulations

The normalizations of our M–M_{gas} and M–Y_{X} relations are lower in mass at fixed X-ray observables (Figures 2 and 3) than the predictions from Nagai et al.'s (2007) simulations. This result is consistent with our pilot study (Zhang et al. 2008), in which we compared our joint lensing/X-ray observational results with a wider range of simulations including those of Borgani et al. (2004) and showed that the normalizations disagree with the simulations at $>2\sigma$ for M–Y_{X} and $>1\sigma$ for M–T and M–M_{gas} relations, respectively.

The structural segregation found in our M–T and M–Y_{X} relations—i.e., undisturbed clusters are ∼40% and ∼20% more massive than disturbed clusters at fixed temperature and fixed Y_{X}, respectively, at $\sim 2\sigma$ significance—disagrees with Kravtsov et al.'s (2006) simulations, upon which they based their proposal of Y_{X} as a low-scatter mass proxy. In summary, they found no difference in the normalization of M–Y_{X} between clusters that they classified as "relaxed" and "unrelaxed" (roughly equivalent to our undisturbed/disturbed classification), and that their relaxed clusters are less massive at fixed temperatures than unrelaxed clusters—i.e., opposite to our observational result.

The intrinsic scatter in the M–T and M–M_{gas} relations at $\Delta = 500$ (∼20%–30% and ∼10%, respectively) is comparable to Kravtsov et al. (2006). However, the anti-correlation between temperature and gas mass deviations from the mean M–T and M–M_{gas} relations predicted by Kravtsov et al. that has motivated much attention on Y_{X} as a low-scatter mass proxy is not supported by our data, although we need to investigate this for a larger sample. In contrast with Kravtsov et al. (2006), we find that the lower limit of the positive coefficient for intrinsic scatter is in better agreement with Stanek et al.'s SPH Millennium Gas Simulations showing that the coefficients between the spectroscopic-like temperature and gas fraction for $z = 0$ and $\Delta = 200$ are positive irrespective of the process of preheating and cooling.

Correlated gas mass and temperature deviations imply a possibility that adiabatic compression/expansion of the intracluster gas is important in cluster evolution. These adiabatic fluctuations propagate much faster than cooling losses, and thus help to explain why Stanek et al.'s simple hydrodynamical simulation matches the observations well. Umetsu et al. (2010) have also suggested that, in the adiabatic expansion phase of a post-merger, both temperature and (encompassed) gas mass decrease compared to those before the merger. Another possibility can be due to the departure from the spherical symmetry. As long as the intracluster gas is in H.E., even in the elongated gravitational potential well, the gas mass distribution is supposed to be more round than the distribution of dark matter. In our modeling, we assume spherical symmetry of the gas and mass profiles, in which deviations from spherical geometry are transferred into systematic measurement and may in part lead to a positive correlation between the gas mass and mass. Furthermore, many possible deviations in mass–observable relations could be relevant to the interpretation of these results. For example, star formation efficiency affects the total gas mass and the epoch for cluster formation affects the cluster temperature. Some of these effects imply other correlations that can be explored in the future with a larger sample, e.g., correlations between temperature deviations and dark matter profile shapes, or between deviations in the gas mass and total stellar fraction.

4.3. Principal Component Analysis of Observational Data

Finally, we propose a method, based on principal component analysis, for constructing and calibrating a low-scatter mass proxy using solely observational data. Thanks to multi-dimensional fitting, since we obtained both the intrinsic covariance and the normalization, we do not need to take into account measurement errors here. By solving the eigenvalue equation $(\boldsymbol{\Sigma} - \sigma^2 \boldsymbol{I})\boldsymbol{y} = 0$, where $\boldsymbol{I}$ is the identity matrix, we can obtain its eigenvalues $\sigma_\pm^2$ and eigenvectors $y_\pm = \ln Y_\pm$, as follows:

$$\sigma_\pm^2 = \frac{1}{2}\left[\sigma_t^2 + \sigma_g^2 \pm \left(\left(\sigma_t^2 - \sigma_g^2\right)^2 + 4r^2\sigma_t^2\sigma_g^2\right)^{1/2}\right], \tag{5}$$

$$Y_\pm = M_{\mathrm{gas}} T^{p_\pm}, \tag{6}$$

$$p_\pm = \frac{\sigma_t^2 - \sigma_g^2 \pm \left(\left(\sigma_t^2 - \sigma_g^2\right)^2 + 4r^2\sigma_t^2\sigma_g^2\right)^{1/2}}{2r\sigma_t\sigma_g}, \tag{7}$$

where Y_- is the mass proxy with the smaller scatter σ_-. Note that if the coefficient r is negative, as predicted by Kravtsov et al. (2006), then the temperature exponent, p_-, is always positive. In this framework, the scaling relation can be written as

$$ME(z) \propto (Y_\pm E(z))^{3/(3+2p_\pm)} h^{(9-4p_\pm)/(3+2p_\pm)/2} \tag{8}$$

and the intrinsic scatter on mass when one estimates from X-ray observables via a new scaling relation is given by

$$\sigma_{\ln M} = \frac{3}{|3 + 2p_\pm|}\sigma_\pm. \tag{9}$$

Basically, a combination of highly correlated/anti-correlated observables gives a new mass proxy with smaller scatter. Unfortunately, our sample is too small to constrain the covariance of the intrinsic scatter well, and to perform this principal component analysis. This exercise awaits the enlargement of our sample. In principle, a combination of the principal component analysis and the method for measuring the covariance of intrinsic scatter (see the Appendix) could be applied to multi-dimensional (i.e., weak-lensing, X-ray, SZ, and optical observable) data sets both within LoCuSS and in large forthcoming surveys.

5. SUMMARY

We have presented a joint weak-lensing and X-ray analysis of 12 clusters based on Subaru and *XMM-Newton* observations as part of the LoCuSS Survey. The main goal is to calibrate the scaling relations between cluster mass obtained from weak-lensing observations (M) and X-ray observables, specifically the gas temperature, gas mass, and quasi-integrated pressure (T, M_{gas}, and $Y_{\mathrm{X}} = M_{\mathrm{gas}} \times T$). An accurate understanding of these relations will be essential to the success of future attempts to constrain dark energy with clusters via growth of structure experiments. Our main results are summarized below.

1. The dynamical state of clusters can be diagnosed empirically via morphological classification of clusters based on asymmetry (A) and fluctuation (F) parameters derived from imaging data from *XMM-Newton*. Undisturbed clusters are identified as those with relatively symmetric and smooth X-ray morphology—$A < 1.1$ and $F < 0.05$; disturbed clusters satisfy either or both of these criteria. Five clusters are classified as undisturbed and seven as disturbed. This classification matches those based on alternative measures such as the presence/absence of cool cores, cold fronts, and substructures in lensing mass maps.
2. We detected structural segregation in the M–T and M–$Y_{\rm X}$ relations at $\Delta = 500$, in the sense that undisturbed clusters are ∼40% and ∼20% more massive than disturbed clusters at fixed T and $Y_{\rm X}$, respectively, at $\sim 2\sigma$ significance. Segregation in the M–T plane is qualitatively in agreement with some of the previous observational results (Smith et al. 2005; Pedersen & Dahle 2007); as far as we know these are the first joint lensing/X-ray results on the M–$Y_{\rm X}$ relation. These results contradict Kravtsov et al.'s predictions upon which they based their proposal that $Y_{\rm X}$ may be a useful low-scatter mass proxy.
3. The intrinsic scatter in the observed M–T, M–$Y_{\rm X}$, and M–$M_{\rm gas}$ relations is measured to be $\sigma_{\ln M} \sim 0.3$, ~ 0.2, and ~ 0.1, respectively, at $\Delta = 500$. $M_{\rm gas}$ therefore appears to be the most promising mass proxy of these three observables, especially because the scatter in M–$M_{\rm gas}$ appears to be independent of the cluster dynamical state.
4. The best-fit mass–observable scaling relations are sensitive to the inclusion/exclusion of one cluster in our sample, namely, A 1914, a well-known merging cluster. If this cluster is excluded from our analysis then the scatter is greatly reduced—most notably, the intrinsic scatter in the M–$Y_{\rm X}$ relation is reduced by ∼25%. We conclude that a larger, more complete sample of clusters is required to reliably calibrate the scaling relations and to robustly measure how the most extreme merging clusters influence the relations. This will be particularly valuable as scaling relation studies proceed to higher redshifts at which merging clusters are expected to become more prevalent.
5. Temperature deviations from the best-fit M–T relation and gas-mass deviations from the M–$M_{\rm gas}$ relation are positively correlated. The coefficient between the gas mass and temperature deviations is positive, independent of our analysis methods, and is found to be $r \geqslant 0.185$. This result, in particular the lower limit on r, agrees well with predictions based on Millennium Gas Simulations (Stanek et al. 2010) and disagrees with predictions based on Kravtsov et al.'s (2006) simulations. However, we caution that the chance probability, $\mathcal{P}(|r'| \geqslant |r|)$, that the correlation coefficient of the two random variables in a sample of 12 drawings, r', is higher than the observed one, is not small. We therefore cannot exclude the possibility that the measured correlation between gas mass and temperature residuals is an accident. A larger sample of clusters is needed to achieve definitive results.
6. Finally, we outlined a new method for constructing a robust low-scatter mass proxy, $Y_{-} = M_{\rm gas} T^{p_{-}}$, calibrated solely by observational data, based on a principal component analysis. This is a generalization of the quasi-integrated pressure $Y_{\rm X} = M_{\rm gas} T$, proposed by Kravtsov et al. (2006). In principle, a combination of the principal component analysis and the method for measuring the covariance of intrinsic scatter could be applied to multidimensional data sets in order to construct a robust new mass proxy.

Our future program will concentrate on expanding the sample of clusters for which high-quality Subaru and *XMM-Newton* data are available in order to achieve definitive results on the issues raised in this article. Key issues will include improvement of the statistical uncertainties on the scaling relation fits, characterizing more fully the influence of extreme merging clusters on the scaling relations, and exploring the balance between physical and orientation effects in causing the observed structural segregation in the M–T and M–$Y_{\rm X}$ relations.

We are very grateful to the members of the LoCuSS collaboration, in particular, Arif Babul, for invaluable discussions and comments. N.O. acknowledges Yuji Chinone for helpful comments on MCMC. N.O., M.T., and T.F. are supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (N.O.: 20740099; M.T.: 20740119; T.F.: 20540245). Y.-Y.Z. and N.O. acknowledge support by the DFG through the Emmy Noether Research Grant RE 1462/2, through Schwerpunkt Program 1177, and through project B6 "Gravitational Lensing and X-ray Emission by Non-Linear Structures" of Transregional Collaborative Research Centre TRR 33 "The Dark Universe," and support by the German BMBF through the Verbundforschung under grants 50 OR 0601 and 50 OR 1005. A.F. acknowledges support from BMBF/DLR under grant 50 OR 0207 and MPG. A.F. was partially supported by NASA grant NNX08AX46G to UMBC. K.U. is partially supported by the National Science Council of Taiwan under the grant NSC95-2112-M-001-074-MY2. G.P.S. acknowledges support from the Royal Society. This work is supported by a Grant-in-Aid for the COE Program "Exploring New Science by Bridging Particle–Matter Hierarchy" and the G-COE Program "Weaving Science Web beyond Particle–Matter Hierarchy" in Tohoku University, funded by the Ministry of Education, Science, Sports and Culture of Japan. This work is in part supported by a Grant-in-Aid for Science Research in a Priority Area "Probing the Dark Energy through an Extremely Wide and Deep Survey with Subaru Telescope" (18072001) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan. Y.-Y.Z. and A.F. acknowledge the hospitality of the Tohoku University during their frequent visits. This work is supported in part by the World Premier International Research Center Initiative (WPI Initiative), MEXT, Japan.

APPENDIX

MULTI-DIMENSIONAL FITTING OF DATA WITH COVARIANCE OF INTRINSIC SCATTER

The measurement of a covariance of intrinsic scatter beyond observational errors is of prime importance in order to understand their intrinsic characteristics. We derive the multidimensional fitting with the covariance of intrinsic scatter, in the context of the Markov Chain Monte Carlo (MCMC) method with standard Metropolis–Hastings sampling, taking into account observational errors.

Here, we suppose a data set ($\boldsymbol{x} = \{x_i, y_{1i}, y_{2i}, \ldots, y_{pi}\}_{i=0}^{n}$) of n sampling numbers and $p+1$ variables. A linear regression equation defined by $a_p + b_p x_i$ and an intrinsic covariance matrix $\boldsymbol{C}_{\rm int}$ ($p \times p$) in the y-coordinates are applied to modeling the relationship between x and y_p in the data set. The diagonal elements in the covariance of intrinsic scatter, $\boldsymbol{C}_{\rm int}$, are intrinsic scatter, σ_p^2, for the y_p observables. The off-diagonal elements describe the intrinsic covariance between y_p variables.

A fitting parameter $\boldsymbol{\theta}$ is composed of the linear regression parameters, $\boldsymbol{a}, \boldsymbol{b}$, and the elements of the intrinsic covariance matrix. From Bayesian statistics, the posterior probability of the parameter vector $\boldsymbol{\theta}$ is proportional to the conditional probability of $\boldsymbol{x}$, $p(\boldsymbol{x}|\boldsymbol{\theta})$, and a prior probability function, $p_{\rm prior}(\boldsymbol{\theta})$,

$$p(\boldsymbol{\theta}|\boldsymbol{x}) \propto p(\boldsymbol{x}|\boldsymbol{\theta})\, p_{\rm prior}(\boldsymbol{\theta}). \tag{A1}$$

The conditional probability of $\boldsymbol{x}$, given parameters $\boldsymbol{\theta}$, is the likelihood described by

$$\begin{aligned} p(\boldsymbol{x}|\boldsymbol{\theta}) = \prod_{i=0}^{n} \frac{1}{(2\pi)^{p/2}|\det(\boldsymbol{C_i})|^{1/2}} \exp\big[-(\boldsymbol{y_i} - (\boldsymbol{a} + \boldsymbol{b}x))^T \boldsymbol{C_i}^{-1} \\ \times (\boldsymbol{y_i} - (\boldsymbol{a} + \boldsymbol{b}x))/2\big], \end{aligned} \tag{A2}$$

where the covariance matrix $\boldsymbol{C_i} = \boldsymbol{C}_{i,\rm obs} + \boldsymbol{C}_{\rm int}$ with the observational error covariance matrix, $\boldsymbol{C}_{\rm obs}$, and intrinsic covariance. The log likelihood of Equation (A2) is given by

$$\begin{aligned} -2L = \sum_i \log(\det(\boldsymbol{C_i})) + \sum_i (\boldsymbol{y_i} - (\boldsymbol{a} + \boldsymbol{b}x))^T \boldsymbol{C_i}^{-1} \\ \times (\boldsymbol{y_i} - (\boldsymbol{a} + \boldsymbol{b}x)). \end{aligned} \tag{A3}$$

The second term on the right-hand side of Equation (A3) is referred to as the χ^2 in some papers (e.g., Akritas & Bershady 1996; Tremaine et al. 2002; Pizagno et al. 2005; Novak et al. 2006; Weiner et al. 2006), and the intrinsic scatter is estimated by requiring the reduced χ^2 to be unity, in the framework of the χ^2 minimization fitting. It is, however, an inadequate fitting procedure because the first term, depending on the parameters, cannot be ignored (see also D'Agostini 2005). We therefore employ the likelihood function for a calculation of the covariance of intrinsic scatter. The diagonal elements in the observed error matrix are given by $(b_p - \partial y_{p,i}/\partial x_i)^2 e_{x,i}^2 + e_{y_p,i}^2$ with the observed variances e_x^2 and $e_{y_p}^2$ in x and $y_p(x)$ variables. Here, $\partial y_i/\partial x_i$ represents the error propagation between two variables. The off-diagonal elements, $C_{pp',i}$, for sample i are $(b_p - \partial y_{p,i}/\partial x_i)(b_{p'} - \partial y_{p',i}/\partial x_i) e_{x,i}^2 + \langle e_{y_{p,i}} e_{y_{p',i}} \rangle$. It describes the correlation between observational errors: the first term is an error correlation via the same x values, and the second one is the correlation between observational errors of y_p and $y_{p'}$. In this study, we measure the covariance of intrinsic scatter in the parameter plane of $x = \ln(M)$, $\boldsymbol{y} = (\ln(T), \ln(M_{\rm gas}))$.

REFERENCES

Akritas, M. G., & Bershady, M. A. 1996, ApJ, 470, 706
Allen, S. W., Schmidt, R. W., & Fabian, A. C. 2001, MNRAS, 328, L37
Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2005, A&A, 441, 893
Arnaud, M., Pointecouteau, E., & Pratt, G. W. 2007, A&A, 474, L37
Bacchi, M., Feretti, L., Giovannini, G., & Govoni, F. 2003, A&A, 400, 465
Bardeau, S., Kneib, J.-P., Czoske, O., Soucail, G., Smail, I., Ebeling, H., & Smith, G. P. 2005, A&A, 434, 433
Bardeau, S., Soucail, G., Kneib, J. P., Czoske, O., Ebeling, H., Hudelot, P., Smail, I., & Smith, G. P. 2007, A&A, 470, 449
Boggs, P. T., Byrd, R. H., & Schnabel, R. B. 1987, SIAM J. Sci. Stat. Comput., 8, 1052
Böhringer, H., et al. 2004, A&A, 425, 367
Borgani, S., et al. 2004, MNRAS, 348, 1078
Buote, D. A., & Tsai, J. C. 1996, ApJ, 458, 27
Conselice, C. J. 2003, ApJS, 147, 1
Corless, V. L., King, L. J., & Clowe, D. 2009, MNRAS, 393, 1235
Croston, J. H., et al. 2008, A&A, 487, 431
D'Agostini, G. 2005, arXiv:astro-ph/0511182
Dahle, H. 2006, ApJ, 653, 954
Ebeling, H., Edge, A. C., Allen, S. W., Crawford, C. S., & Fabian, A. C. 2000, MNRAS, 318, 333
Ebeling, H., Edge, A. C., Boehringer, H., Allen, S. W., Crawford, C. S., Fabian, A. C., Voges, W., & Huchra, J. P. 1998, MNRAS, 301, 881
Ettori, S., Tozzi, P., Borgani, S., & Rosati, P. 2004, A&A, 417, 13
Evrard, A. E. 1990, ApJ, 363, 349
Evrard, A. E., Metzler, C. A., & Navarro, J. F. 1996, ApJ, 469, 494
Fahlman, G., Kaiser, N., Squires, G., & Woods, D. 1994, ApJ, 437, 56
Fang, T., Humphrey, P., & Buote, D. 2009, ApJ, 691, 1648
Finoguenov, A., Böhringer, H., & Zhang, Y.-Y. 2005, A&A, 442, 827
Finoguenov, A., Reiprich, T. H., & Böhringer, H. 2001, A&A, 368, 749
Gavazzi, R. 2005, A&A, 443, 793
Ghizzardi, S. 2001, In-Flight Calibration of the PSF for the MOS1 and MOS2 Cameras, EPIC-MCT-TN-011 (Madrid: ESA)
Govoni, F., Markevitch, M., Vikhlinin, A., VanSpeybroeck, L., Feretti, L., & Giovannini, G. 2004, ApJ, 605, 695
Gutierrez, K., & Krawczynski, H. 2005, ApJ, 619, 161
Hoekstra, H. 2007, MNRAS, 379, 317
Ilbert, O., et al. 2009, ApJ, 690, 1236
Jing, Y. P., & Suto, Y. 2002, ApJ, 574, 538
Juett, A. M., Davis, D. S., & Mushotzky, R. 2010, ApJ, 709, L103
Kaiser, N. 1986, MNRAS, 222, 323
Kawaharada, M., et al. 2010, ApJ, 714, 423
Komatsu, E., et al. 2009, ApJS, 180, 330
Kravtsov, A. V., Vikhlinin, A., & Nagai, D. 2006, ApJ, 650, 128
Marrone, D. P., et al. 2009, ApJ, 701, L114
Maughan, B. J. 2007, ApJ, 668, 772
Mazzotta, P., & Giacintucci, S. 2008, ApJ, 675, L9
McCarthy, I. G., Babul, A., Bower, R. G., & Balogh, M. L. 2008, MNRAS, 386, 1309
Meneghetti, M., Rasia, E., Merten, J., Bellagamba, F., Ettori, S., Mazzotta, P., Dolag, K., & Marri, S. 2010, A&A, 514, A93
Nagai, D., Vikhlinin, A., & Kravtsov, A. V. 2007, ApJ, 655, 98
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Novak, G. S., Faber, S. M., & Dekel, A. 2006, ApJ, 637, 96
Oguri, M., Takada, M., Okabe, N., & Smith, G. P. 2010, MNRAS, 405, 2215
Oguri, M., Takada, M., Umetsu, K., & Broadhurst, T. 2005, ApJ, 632, 841
Okabe, N., Takada, M., Umetsu, K., Futamase, T., & Smith, G. P. 2010, PASJ, 62, 811
Okabe, N., & Umetsu, K. 2008, PASJ, 60, 345
Pedersen, K., & Dahle, H. 2007, ApJ, 667, 26
Peterson, J. R., Kahn, S. M., Paerels, F. B. S., Kaastra, J. S., Tamura, T., Bleeker, J. A. M., Ferrigno, C., & Jernigan, J. G. 2003, ApJ, 590, 207
Piffaretti, R., & Valdarnini, R. 2008, A&A, 491, 71
Pizagno, J., et al. 2005, ApJ, 633, 844
Pugh, E. M., & Winslow, G. H. 1966, The Analysis of Physical Measurements (Reading, MA: Addison-Wesley), Section 12-5
Randall, S. W., Sarazin, C. L., & Ricker, P. M. 2002, ApJ, 577, 579
Rasia, E., et al. 2006, MNRAS, 369, 2013
Richard, J., et al. 2010, MNRAS, 404, 325
Sanderson, A. J. R., Edge, A. C., & Smith, G. P. 2009a, MNRAS, 395, 764
Sanderson, A. J. R., O'Sullivan, E., & Ponman, T. J. 2009b, MNRAS, 395, 764
Sanderson, A. J. R., Ponman, T. J., Finoguenov, A., Lloyd-Davies, E. J., & Markevitch, M. 2003, MNRAS, 340, 989
Smith, G. P., Edge, A. C., Eke, V. R., Nichol, R. C., Smail, I., & Kneib, J.-P. 2003, ApJ, 590, L79
Smith, G. P., Kneib, J.-P., Smail, I., Mazzotta, P., Ebeling, H., & Czoske, O. 2005, MNRAS, 359, 417
Stanek, R., Rasia, E., Evrard, A. E., Pearce, F., & Gazzola, L. 2010, ApJ, 715, 1508
Tremaine, S., et al. 2002, ApJ, 574, 740
Umetsu, K., & Broadhurst, T. 2008, ApJ, 684, 177
Umetsu, K., et al. 2009, ApJ, 694, 1643
Umetsu, K., et al. 2010, ApJ, 714, 1470
Vikhlinin, A., Burenin, R., Forman, W. R., Jones, C., Hornstrup, A., Murray, S. S., & Quintana, H. 2007, in Heating Versus Cooling in Galaxies and Clusters of Galaxies, ed. H. Böhringer, G. W. Pratt, A. Finoguenov, & P. Schuecker (Berlin: Springer), 48
Vikhlinin, A., et al. 2006, ApJ, 640, 691
Vikhlinin, A., et al. 2009a, ApJ, 692, 1033
Vikhlinin, A., et al. 2009b, ApJ, 692, 1060
Weiner, B. J., et al. 2006, ApJ, 653, 1049
Zhang, Y.-Y., Reiprich, T. H., Finoguenov, A., Hudson, D. S., & Sarazin, C. L. 2009, ApJ, 699, 1178
Zhang, Y.-Y., et al. 2007, A&A, 467, 437
Zhang, Y.-Y., et al. 2008, A&A, 482, 451
Zhang, Y.-Y., et al. 2010, ApJ, 711, 1033