3B). To further characterize genic eQTLs, we compared genic subcategories (exonic excluding UTRs, intronic, 5′ UTR, 3′ UTR, upstream, and downstream). Upstream and 5′ UTR were distinctly important compared to other genic subcategories: The most intense enrichment was observed for the 5′ UTR (41.79 fold), followed by upstream (27.95 fold), and these two subcategories had largest mean $|\beta|$ and R^2 values (Table 2). Mean $|\beta|$ and R^2 values of seven categories (six genic subcategories and the intergenic category) were significantly different (ANOVA $P = 1.1E\text{-}05$ for $|\beta|$ and $P = 2.5E\text{-}08$ for R^2). Significantly different category pairs are shown in Table S2: The upstream had significantly larger $|\beta|$ value than intron, 3′ UTR, or downstream, and for R^2 values, no pairs of genic subcategories were significantly different. Consistently, *cis*-eQTLs with larger $|\beta|$ values were more common in 5′ UTR or upstream regions than in other regions (Figure 3C); additionally, R^2 value of each genic subcategories tended to be larger than the R^2 value of the intergenic category (Figure 3D). We did not observe statistically significant difference between non-synonymous and synonymous SNPs in enrichment or mean effect sizes (eQTL enrichment: Fisher's exact $P = 0.56$, $|\beta|$: $P = 0.257$ R^2: $P = 0.70$), and the distribution of $|\beta|$ values or R^2 values was similar (Figure 3E, 3F).

Intergenic eQTLs and RegulomeDB class. Next, we characterized the intergenic *cis*-eQTLs; again, we focused only on *cis*-eQTLs that affected mRNA transcripts. Although the mean effects of intergenic eQTLs were significantly smaller than those of genic eQTLs, intergenic eQTLs are still important because the 1,523 intergenic eQTLs constituted 50.85% of the *cis*-eQTLs (Table 2). Therefore, to further characterize this large number of intergenic eQTLs we analyzed each in terms of regulatory potential; this potential was predicted based on known epigenetic evidence. We classified each intergenic eQTL into one of seven numbered categories based on the RegulomeDB, which indicates how likely a variant is to disrupt transcription factor binding [19] (see Methods for the classification). We observed statistically significant trends in means of R^2 values ($P = 3e\text{-}05$) but not in means of $|\beta|$ ($P = 0.37$); the eQTLs classified into higher potential classes had stronger effects (Figure 4). Although the eQTLs in Category 1 had the largest mean R^2, the means of other categories were not apparently different.

Relationship between eQTL effects and distance. Next, we investigated whether and how distances between *cis*-eQTLs and their mRNA transcripts were related to the magnitudes of the effects. Distances and effects were strongly correlated with exponential decay in both the 5′ and 3′ directions (Figure 5A, 5D). eQTLs were concentrated in regions near genes; 73% of *cis*-eQTLs outside genes were located within 50 kb of their target genes. Promoters are usually located within 100 bp upstream of genes; nevertheless, eQTLs were not apparently enriched in promoter regions (Figure 5B, 5E). As distances increased, eQTLs of small effect became more common; this trend is evident in R^2 values for eQTLs>100 kb, but not in the $|\beta|$ values (Figure 5C, 5F).

Table 2. Counts and proportions of gene structure-based categories and protein consequences in local SNPs and *cis*-eQTLs.

					Mean effect	
	Categories	**Local SNPs (%)**	***cis*-eQTLs (%)**	**Enrich**	**β**	**R^2**
Intergenic		10,268,814 (93.11)	1,523 (50.85)	0.55	0.31	0.17
Genic		716,576 (6.50)	1,370 (45.74)	7.04	0.33	0.21
	Exonic	25,822 (0.23)	109 (3.64)	15.54	0.32	0.19
	Splicing	28 (0.00)	0 (0.00)	-	-	-
	Intronic	633,398 (5.74)	889 (29.68)	5.17	0.33	0.20
	3′ UTR	29,609 (0.27)	188 (6.28)	23.38	0.28	0.22
	5′ UTR	3,965 (0.04)	45 (1.50)	41.79	0.41	0.22
	Upstream	11,727 (0.11)	89 (2.97)	27.95	0.47	0.23
	Downstream	12,027 (0.11)	50 (1.67)	15.31	0.27	0.19
N.A.		42,870 (0.39)	102 (3.41)	8.76	0.38	0.21
Total		**11,028,260 (100.00)**	**2,995 (100.00)**	**1.00**	**0.32**	**0.19**
Exonic	nonsyn	11,739 (45.46)	52 (47.71)	1.05	0.35	0.20
	syn	13,662 (52.91)	53 (48.62)	0.92	0.29	0.18
	stopgain	52 (0.20)	2 (1.83)	9.11	0.45	0.24
	stoploss	10 (0.04)	0 (0.00)	-	-	-
	N.A.	359 (1.39)	2 (1.83)	1.32	0.32	0.19
Total		**25,822 (100.00)**	**109 (100.00)**	**1.00**	**0.32**	**0.19**

Local SNPs and *cis*-eQTLs that affect mRNA transcripts are counted within each gene-based functional category (upper panel) and for each protein consequence (lower panel).

Enrich: the fold change in proportion that each group constitutes among *cis*-eQTLs compared to among all local SNPs.

The category "Exonic" does not include 5′ and 3′ untranslated regions (UTRs); "Upstream" and "Downstream" each includes regions within 1 kb from transcription start or end sites of genes, respectively; "Splicing" includes SNPs 2 bp from exon-intron splicing junctions and within an intron; SNPs 2 bp from a splice junction and within an exon are designated "Exonic"); "Intronic" includes SNPs in introns, but not those 2 bp from exon-intron splicing junctions; "nonsyn" indicates a SNP in an Exonic that is non-synonymous; "syn" indicates an SNP in an Exonic that is synonymous; "stopgain" indicates an SNP in an Exonic and with a variant that causes the creation of stop codon; "stoploss" indicates an SNP in an Exonic and with a variant that eliminates a stop codon.

N.A. means "Not Available" and includes SNPs that were found in a gene, but that could not be assigned to a specific functional category.

Totals for gene-structure-based classification and protein consequences are shown in bold font.

doi:10.1371/journal.pone.0100924.t002

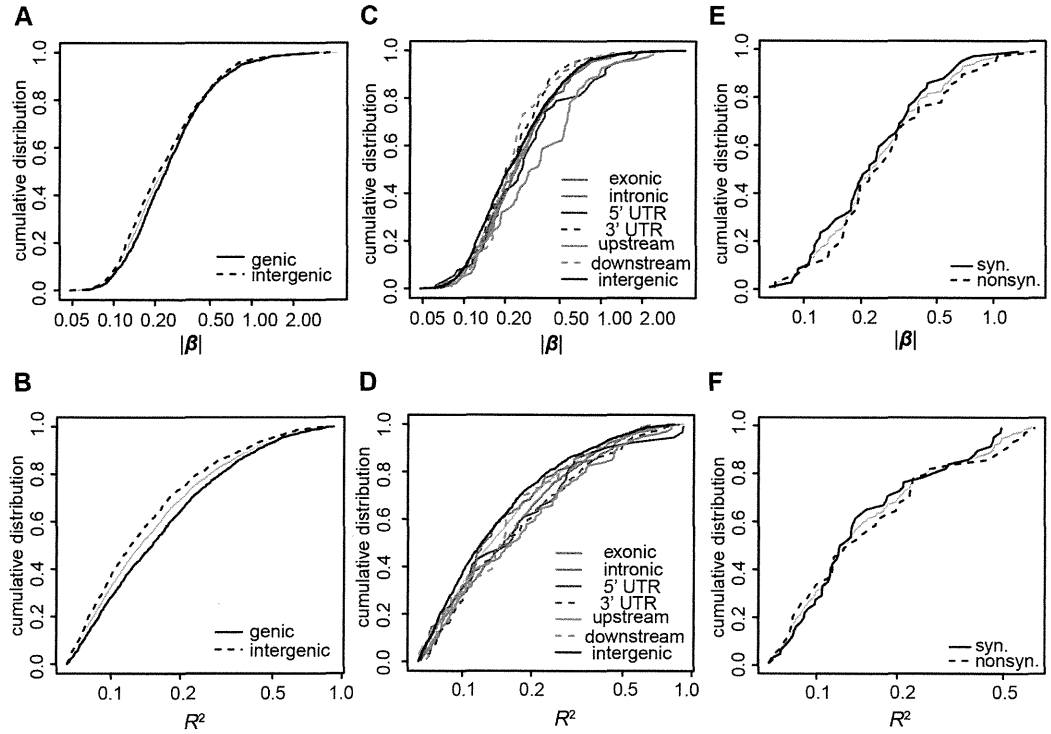


Figure 3. Cumulative curves of effect magnitudes of *cis*-eQTLs in gene-structure-based functional categories. Cumulative curves represent the distributions of |β| values or R^2 values of *cis*-eQTLs in each category. Cumulative distribution of all *cis*-eQTLs (A–D) or all exonic *cis*-eQTLs (E–F) are shown in grey. The X axis is a log scale. A, B) Distributions of genic and intergenic *cis*-eQTLs for |β| values (A) or for R^2 values (B). C, D) Distributions of genic subcategories and intergenics for |β| values (C) or for R^2 values (D). E, F) Distributions of nonsynonymous and synonymous eQTLs for |β| values (E) and for R^2 values (F).
doi:10.1371/journal.pone.0100924.g003

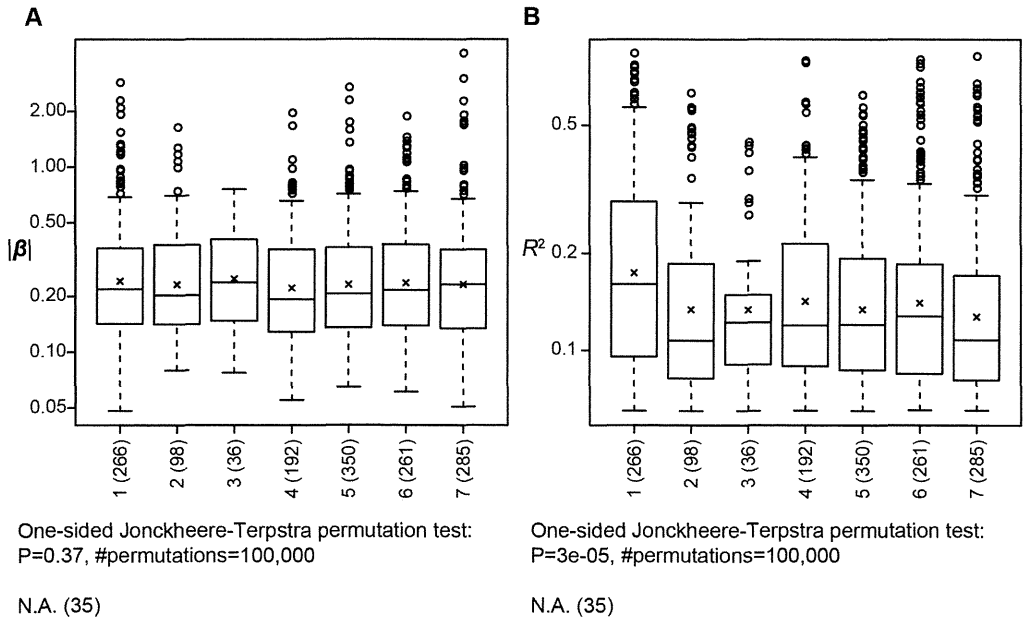


Figure 4. Trend in effects associated with regulatory classes of intergenic *cis*-eQTLs. The box-and-whisker plots show distributions of |β| values (A) or of R^2 values (B) of intergenic *cis*-eQTLs that affect mRNA transcripts for regulatory classes defined by the RegulomeDB. A cross indicates the mean effect of each class. The number of *cis*-eQTLs belonging to each class is shown in the parentheses following the class name. Jonckheere-Terpstra permutation test was used to test each trend, and the results are shown under the box-and-whisker plots. N.A.: not available.
doi:10.1371/journal.pone.0100924.g004

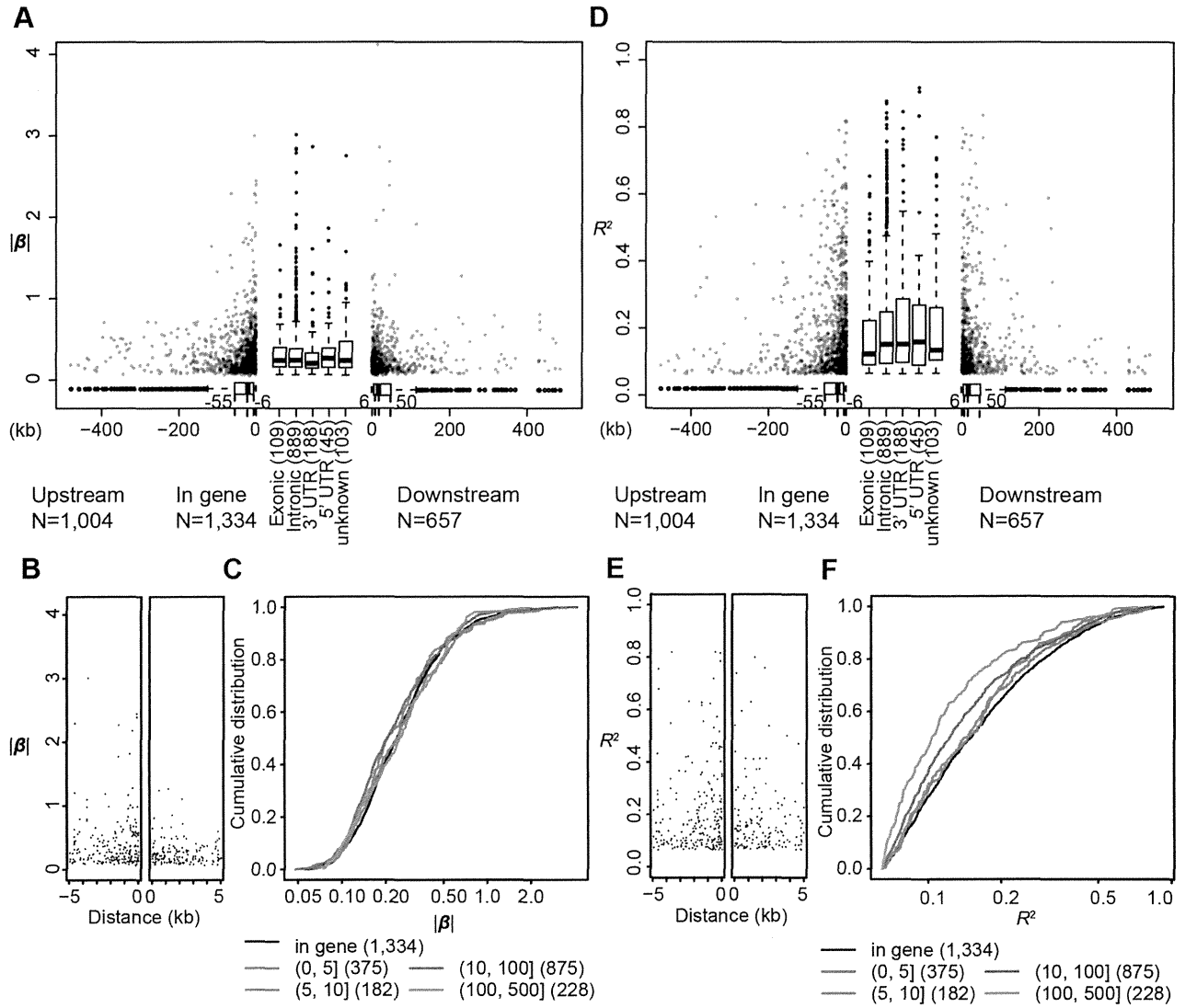


Figure 5. Relationships between effects of *cis*-eQTLs and distance from genes. |β| values (A) and R^2 values (D) of *cis*-eQTLs that affect mRNA transcripts are plotted against distances from the respective target genes by scatter (non-transcribed regions) and by box-and-whisker plots (transcribed regions). eQTLs in transcribed regions are shown for each gene-structure-based category. The number following each category name represents the number of *cis*-eQTLs classified into that category. Negative distance values indicate that the eQTL is upstream of the target gene, and positive values indicate that it is downstream, regarding transcriptional directions. Distributions of distances are represented by box-and-whisker plots below the scatters. Magnified view for <5 kb of genes is shown for |β| (B) and R^2 (E). Cumulative distribution of |β| (C) and R^2 (F) of eQTLs are shown for each division of eQTLs; each division represent a defined distance (kb) from the respective target gene. The number in the parentheses following each distance range in the legend is the number of *cis*-eQTLs identified in that range. The X-axis is a log scale. One eQTL located within a gene (*C16orf55*) that was assigned function of "downstream" is shown as "unknown"; therefore, the number of "In gene" eQTLs shown in (A) and (D) is the sum of the numbers of Exonic, Splicing, Intronic, 5′ UTR, 3′ UTR, and N.A. in Table 2 plus 1.
doi:10.1371/journal.pone.0100924.g005

Trans-eQTL analysis

P values of all tests for distant SNPs were distributed almost uniformly with a slight excess of small *P* values, suggesting that only a small fraction of distant SNPs affected transcriptional regulation (Figure S1B). With a stringent multiple-testing correction ($P<1.15E-12$) and excluding redundancy due to LD and excluding possible false positives because of cross-hybridization to local regions (see Methods), we identified 165 combinations of independent *trans*-eQTLs and transcripts that comprised 114 unique transcripts (74 genes) affected by 105 unique *trans*-eQTLs, and these *trans*-regulated transcripts represented 0.4% of all tested transcripts (Table 1). Large *trans*-effects were identified (Figure 2C, 2D and Table 1). The number of *trans*-eQTLs with |β| values larger than 0.3 was 118 for all transcripts, which covers 0.39% of all tested transcripts; additionally, each *trans*-eQTL had an R^2 value larger than 0.1 (Table 1). The ratio of the number of *trans*-eQTLs to the number of *cis*-eQTLs at the same cutoff values of |β| or R^2 tend to be smaller as the cutoff value became larger (Table 1), indicating strong effects are more abundant in *cis*-eQTLs. All the 165 *trans*-eQTL transcript pairs are provided in File S1.

We assigned RegulomeDB classes [19] to *trans*-eQTLs, and tested a trend in the same manner as *cis*-eQTLs. Unlike *cis*-eQTLs,

we did not observe statistically significant trends in means of R^2 values (P = 0.99) or in means of $|\beta|$ (P = 0.92).

Multi-regulatory eQTLs

A *cis*-eQTL that is associated with expression of multiple genes might indicate the existence of a long-range enhancer/repressor that influences the expression of a cluster of genes in a region. We identified 6 *cis*-eQTLs that were each associated with expression levels of three or more mRNA-coding genes (Table 3). These multi-regulatory *cis*-eQTLs were each associated with the regulated transcripts in the same direction (Figure S4A).

A *trans*-eQTL that is associated with the expression of multiple genes is a potential master regulator. Our *trans*-eQTL map indicates that there are some *trans*-eQTL hotspots that were involved in multiple genes across the genome (Figure S5). We identified 5 *trans*-eQTLs that were each associated with three or more mRNA-coding genes (Table 3). Rs7801498 was also identified as a *cis*-eQTL for two genes (*LRWD1* and *ORAI2*). Notably, again, these multi-regulatory *trans*-eQTLs were each associated with the regulated transcripts in the same direction (Figure S4B).

Replication analysis with independent studies

We compared our eQTLs to a meta-analysis of eQTL studies of whole blood samples conducted by Westra *et al.* [20]. They analyzed samples from 5,311 individuals from European populations. We focused on 15,733 genes that were commonly tested in both studies. At FDR<0.05, 10.6% of the genes were found *cis*-regulated in both studies; 60.9% of 2,750 *cis*-regulated genes in this study were replicated; and the concordance rate (*i.e.*, consistently *cis*-regulated or non-*cis*-regulated in both studies) was 68.8%. The concordance rate increased as FDR thresholds became more stringent up to 74.4% at FDR<1E-06. The replication rate of our *cis*-regulated genes was significantly associated with median non-adjusted expression levels (logistic regression P<2E-16, log OR = 0.14), but not with SD (P = 0.14). 45.2% of 3,106 pairs of our *cis*-eQTLs (including SNPs in r^2>0.8) and genes tested in the meta-analysis were replicated. For replication of *trans*-eQTLs we found 978 distant SNP-transcript pairs in our results that corresponded to *trans*-eQTL-gene pairs identified in the meta-analysis. Six pairs were significant at P<5.1E-05, which corresponds to Bonferroni-corrected P = 0.05 for 978 tests (Table S3). Particularly, *trans*-eQTL for *CALD1* was replicated at the original significance level (P = 5.30E-16). Regarding that only the limited number of SNP-gene pairs were tested in common with the meta-analysis for *trans*-eQTLs, we also compared our *trans*-eQTLs with those identified for whole blood samples obtained from 76 Japanese individuals [17]. Over 8.6 billion tests for SNP-gene pairs were performed in both studies. Of the common tests, 41 and 2 pairs were identified as *trans*-eQTLs in the current and previous studies, respectively. We identified 1 *trans*-eQTL-gene pair exactly consistent between the studies (rs4487686 for *POLR2J4*). Regarding the number of performed tests, identifying one consistent result by chance is extremely unlikely (Fisher's exact test P<5E-08).

Application of the eQTL map to interpretation of GWAS results

eQTL maps improve interpretation of GWAS results by linking SNPs and genes whose expressions are actually altered. We used previously published GWAS of Crohn's disease to comprehensively illustrate how our eQTL map improves interpretation of GWAS results. We identified 12 records for which our eQTL maps were informative for interpretation among all 220 records for Crohn's disease obtained from the NHRGI GWAS Catalog (http://www.genome.gov/gwastudies/) (Table 4). We define the following four informative cases for results of applying our eQTL map to GWAS results; a GWAS result is classified into Case 1 when the eQTL map may suggest different possible interpretation for GWAS, Case 2 when the eQTL map supports the interpretation provided by GWAS, Case 3 when the eQTL map helped to prioritize multiple genes inconclusively reported by the GWAS, or Case 4 when a *trans*-effect of GWAS-identified SNP was suggested (see supplementary note in File S3 for detailed definition).

For an example of Case 1, an intergenic SNP, rs694739, was identified in a GWAS of Crohn's disease (record 3 in Table 4); the study reported *PRDX5* and *ESRRA* as putative causative genes [21]. The GWAS-identified SNP was found in LD (r^2 = 0.85) with a *cis*-eQTL (rs600377) for *CCDC88B* in our eQTL map (β = –0.26, P = 1.0E-06). A *cis*-eQTL was identified for *PRDX5*, but the *cis*-eQTLs for *PRDX5* and *CCDC88B* were not in LD (r^2 = 0.01); and after correcting for the genotypes of the GWAS-identified SNP, the *cis*-effect on *CCDC88B* expression was not significant (P_c = 0.51). Therefore, given the eQTL map, the most likely causative gene was *CCDC88B*. Based on our analyses, 6 of the 12 records were classified into Case 1, and thus, in each of these records, the eQTL-suggested gene should also be considered as another candidate gene.

Four of the 12 records were classified into Case 2. Three intergenic SNPs (rs7714584, rs11747270, rs13361189) were each reported in GWAS [21–23]; and in each study, *IRGM* was suggested as the candidate gene (records 8–10 in Table 4). All of these SNPs were each in perfect LD (r^2 = 1.00) with a *cis*-eQTL (rs1428554) that influenced expressions of *IRGM* (β = –0.40, P = 3.4E-13). None of the three SNPs were in LD (r^2>0.8) with any other *cis*-eQTLs that affected any other gene. Therefore, our eQTL analysis supported the conclusions of the GWAS.

As an example of Case 3, a GWAS (record 11 in Table 4) identified rs4656940 (in the intron of *CD244*) reporting two candidate causative genes (*CD244 and ITLN1*) [21]. The reported SNP was in perfect LD with a *cis*-eQTL (rs11265498) that influenced expressions of *ITLN1* (β = −0.67, P = 2.4E-17), where no *cis*-eQTL was identified for *CD244*. Therefore, our eQTL map indicated that *ITLN1* was the most likely causative gene. Our eQTL map helped to prioritize candidate genes for two of 12 records.

Any records for Crohn's disease were not classified into Case 4. In all GWAS records, we identified 13 Case-4 records (File S2). For instance, rs1354034 (in the intron of *ARHGEF3*, on chr3) was reportedly associated with platelet counts and mean platelet volume [24,25], and *ARHGEF3* was identified as a putative causative gene. In our eQTL map, the reported SNP was not associated with the expression of *ARHGEF3* or any other tested gene on the same chromosome, but with *CALD1* on a different chromosome, chr7 (β = –0.48, P = 5.3E-16). For another example, rs2517713 (intergenic, on chr6) was identified in a study of nasopharyngeal carcinoma [26] and *HLA-A* was reported as a putative causative gene. In our eQTL map, the reported SNP was not associated with the expression of *HLA-A* or any other tested gene on the same chromosome, but of *NRSN2* on a different chromosome, chr20 (β = –0.21, P = 2.2E-15). Notably, decreased expression of *NRSN2* was reported to be associated with hepatocellular carcinoma [27].

Similarly, we analyzed 10,076 (8,069) GWAS records (unique SNPs). We identified 386 cases in which *cis*- or *trans*-effects were identified for the reported SNPs, and classified each into one of the four cases; we found 191 (148) Case-1 records, 97 (80) Case-2 records, 85 (60) Case-3 records, and 13 (6) Case-4 records. We

Table 3. Multi-regulatory *cis*-eQTLs and *trans*-eQTLs.

					LD block			
eQTL	**Chr**	**Position**	**MAF**	**HWE-P**	**Start**	**End**	**Length**	**Gene Symbol**
cis								
rs7522860	1	156,275,281	0.49	0.644	156,208,230	156,314,627	106,398	*TMEM79;SMG5;C1orf85;PAQR6*
rs6464103	7	150,478,385	0.37	0.711	150,476,888	150,478,385	1,498	*TMEM176B;TMEM176A;ABP1*
rs4390300	10	60,144,207	0.47	0.817	60,144,207	60,168,003	23,797	*IPMK;UBE2D1;TFAM*
rs2416549	12	11,325,804	0.24	0.116	11,045,512	11,349,454	303,943	*TAS2R14;TAS2R30;PRB1*
rs35969491	12	11,339,020	0.24	0.084	11,045,512	11,349,454	303,943	*TAS2R10;PRR4;PRH2;PRB4*
rs7226263	17	44,814,884	0.32	0.111	44,788,310	44,853,872	65,563	*WNT3;ARL17B;ARL17A;NSF*
trans								
rs116711766	1	160,093,165	0.075	0.3909	160,093,165	160,093,165	1	*ITGA7;MC1R;FAM22G*
rs11718621	3	40,362,122	0.288	1.0000	40,362,122	40,463,063	100,942	*DIRC1;MAB21L2;PRSS36; HIST2H2BF;KRTAP19–2;FSD1;LRRD1*
rs6773917	3	40,469,254	0.492	0.4881	40,373,259	40,498,845	125,587	*DIRC1;MAB21L2;PRSS36; HIST2H2BF;NEURL;KRTAP19–2;FSD1;LRRD1*
rs7801498	7	102,089,595	0.368	0.8039	102,089,595	102,089,595	1	*MUC4;GFRA1;MIOX;GYPA*
rs10873415	14	92,558,171	0.380	0.0097	92,434,957	92,558,171	123,215	*GADD45GIP1;SOX13;TFEB;EIF2C1*

Chr, Position: chromosomal positions of eQTLs; MAF: minor allele frequency; HWE-P: Hardy-Weinberg Equilibrium test *P* value; LD block: range in which SNPs in LD (r^2>0.8) with the eQTLs exist.
doi:10.1371/journal.pone.0100924.t003

Table 4. Summary of GWAS records associated with Crohn's disease and eQTL mapping results.

		Suggested genes		SNPs			eQTL statistics			Top local SNP for GWAS gene			
Case	**Record**	**GWAS**	**eQTL**	**GWAS**	**eQTL**	r^2	β	*P*	P_c	**SNP**	β	*P*	r^2
Case 1	1[21]	*CCR6*[a]	*RNASET2*	rs415890	rs400837	0.99	−0.36	2.7E-39	0.87	Not tested			
	2[21]	*FADS1*	*FADS2*	rs102275	rs108499	0.97	0.16	3.2E-10	0.74	rs174570	0.17	6.2E-07	0.99
	3[21]	*PRDX5*	*CCDC88B*	rs694739	rs600377	0.85	−0.26	1.0E-06	0.51	rs2286614	0.42	4.5E-23	0.01
		ESRRA								rs641811	0.06	n.s.	0.01
	4[21]	*IKZF3*	*GSDMB*	rs2872507	rs1008723	0.98	−0.38	6.9E-38	0.81	rs56030650	0.05	n.s.	0.01
		ZPBP2								rs62065216	−0.09	n.s.	0.01
		ORMDL3								rs1054609	−0.18	3.6E-14	0.98
		GSMDL[a]								Not tested			
	5[22]	*ORMDL3*	*GSDMB*	rs2872507	rs1008723	0.98	−0.38	6.9E-38	0.81	rs1054609	−0.18	3.6E-14	0.98
	6[21]	*RTEL1*	*ZGPAT*	rs4809330	rs6011058	1.00	0.09	2.9E-07	1.00	rs2252258	−0.05	n.s.	0.002
		SLC2A4RG								rs310609	−0.07	n.s.	0.02
		TNFRS-F6B[a]								Not tested			
Case 2	7[21]	*PLCL1*	*PLCL1*	rs6738825	rs1866664	0.98	−0.25	3.0E-07	0.81	rs1866664	−0.25	3.0E-07	1
	8[21]	*IRGM*	*IRGM*	rs7714584	rs1428554	1.00	−0.40	3.4E-13	0.98	rs1428554	−0.40	3.4E-13	1
	9[22]	*IRGM*	*IRGM*	rs11747270	rs1428554	1.00	−0.40	3.4E-13	0.98	rs1428554	−0.40	3.4E-13	1
	10[23]	*IRGM*	*IRGM*	rs13361189	rs1428554	1.00	−0.40	3.4E-13	0.98	rs1428554	−0.40	3.4E-13	1
Case 3	11[21]	*ITLN1*[b]	*ITLN1*	rs4656940	rs11265498	1.00	−0.67	2.4E-17	1.00	rs11265498	−0.67	2.4E-17	1
		CD244								rs574610	−0.12	n.s.	0.16
	12[46]	*RNASET2*[b]	*RNASET2*	rs2149085	rs400837	0.99	−0.36	2.7E-39	0.87	rs400837	−0.36	2.7E-39	1
		FGFR1OP								rs73039162	0.68	7.5E-45	0.078
		CCR6[a]								Not tested			
		MIR3939[a]								Not tested			

[a]The GWAS-reported gene was not included in our study.
[b]GWAS-reported genes that match the eQTL-suggested genes in Case 3.
r^2: correlation of genotypes for linkage disequilibrium between the GWAS-identified SNP and *cis*-eQTL (in the "SNPs" column), or between the top local SNP for GWAS gene and *cis*-eQTL (in the "Top local SNP for GWAS Gene" column).
P_c: *P* value of a conditional regression on genotypes of GWAS-identified SNP.
Genes suggested by GWAS and our eQTL map are listed in the "Suggested genes" column; eQTL statistics are listed in the "eQTL statistics" column; most significant local SNP for the GWAS-reported gene is shown in the "Top local SNP for GWAS gene" column.
n.s: not significant.
doi:10.1371/journal.pone.0100924.t004

identified 6 lincRNAs in the Case-1 records that were most significantly associated with GWAS-reported SNPs. In summary, our eQTL map was informative for 3.8% of the GWAS records, each of which was classified into one of the four cases; 1.9% into Case 1, 1.0% into Case 2, 0.8% into Case 3, and 0.1% into Case 4. We provide the results of our application of our eQTL map to the GWAS records in File S2.

Discussion

This study identified the largest number of eQTLs for East Asian whole blood samples to our knowledge. We identified 3,804 *cis*-eQTLs and 165 *trans*-eQTLs. *Cis*-effects were previously found for 44% (6,418 genes) of tested genes [20] for Caucasian whole blood samples. In the current study, *cis*-effects were found for 16.9% of the tested genes, which is in line with estimated powers in a previous study [28].

We identified 74 genes with *trans*-effects, which constituted 0.4% of tested genes. We believe that we underestimated the proportion of true *trans*-effects because we used the most stringent corrections for multiple testing. In fact, the smallest R^2 for any of the *trans*-eQTLs ($R^2 = 0.16$) was 2.4-fold greater than the smallest R^2 for any identified *cis*-eQTLs ($R^2 = 0.065$).

We analyzed and characterized our eQTLs in various aspects; 1) *cis*-eQTLs in terms of gene structure, epigenetic factors, and distance from genes; 2) multi-regulatory eQTLs; 3) eQTLs for mRNA as compared to those for lincRNAs; 4) application of eQTL maps to GWAS results; and 5) replication with independent samples.

1) *Cis*-eQTL analyses

The comparison between the genic and intergenic *cis*-eQTLs suggested that factors involved in expression levels are more enriched and stronger in genic regions (those located within a gene or within 1 kb of a gene) than intergenic regions (>1 kb from genes). All genic subcategories were each overrepresented compared to the intergenic regions (Table 2). We also showed that upstream and 5′-UTR regions particularly had strong effects compared to other genic regions. It would be reasonable to consider that upstream regions are important because transcription factor binding sites and transcription regulatory modules are enriched in 5′ flanking regions of genes. Strong effects in 5′ UTRs would imply that post-transcriptional regulation via 5′ UTRs has a particularly strong impact on expression levels. The significant association between R^2 of *cis*-eQTLs and epigenetic classification indicated that epigenetic factors (e.g., transcription regulatory modules) have influences on transcription that depend upon nucleotide sequences. Interestingly, the trend was not observed for $|\beta|$.

92% of *cis*-eQTLs were within their target genes or in 100 kb flanking regions, which is consistent with previous studies [3,7]; and it was also consistent that most of large-effect eQTLs were located within 20 kb [29].

2) Multi-regulatory eQTLs

We identified 6 and 5 multi-regulatory *cis*- and *trans*-eQTLs, respectively. We note that a pair of multi-regulatory *cis*-eQTLs on chr12, rs2416549 and rs35969491, and another pair of multi-regulatory *trans*-eQTLs on chr3, rs11718621 and rs6773917, each are likely to indicate the same locus because they were each close ($r^2 = 0.99$ and 0.39, respectively) and the regulated gene sets are similar. Multi-regulatory eQTLs may comprise two types of eQTLs; some may be true master regulators, while others may each comprise a group of eQTLs in strong LD, each of which regulates one gene. Further studies are needed to identify more multi-regulatory eQTLs so that they would be further analyzed in terms of LD structure and effect sizes comparing with eQTLs regulating one gene. Presence of *trans*-acting master regulators has been increasingly suggested[30–32]. However, it is very challenging to identify master regulators because statistical power to detect *trans*-eQTLs is low because of multiple testing corrections. Interestingly, with the stringent threshold of this study, *trans*-regulated genes were often associated with multi-regulatory *trans*-eQTLs (Figure S5), which may suggest multi-regulatory *trans*-eQTLs tend to have large effects.

3) mRNA and lincRNA transcripts

The importance of lincRNAs to phenotypic variation is increasingly recognized; nevertheless, previous eQTL studies focused only on coding genes, and did not include analyses of lincRNA transcripts. Here, we examined the genetic causes of variation in expression of coding genes and of lincRNAs. Coding genes and lincRNAs exhibited different characteristics; for example, the proportion of *cis*-regulated transcripts was 3 times larger for mRNAs (15.1% vs. 4.8%, Table 1); sequence variations influence coding gens more than lincRNAs. Nevertheless, eQTLs for lincRNAs should not be ignored because still 5.3% of lincRNAs were regulated by either *cis*- or *trans*-eQTLs, and the mean R^2 values of *cis*- or *trans*-eQTLs regulating lincRNAs were as large as those regulating mRNAs (Table 1, Wilcoxon's rank-sum $P = 0.094$), and $|\beta|$ values were even larger (Table 1, Wilcoxon's rank-sum $P = 3.2\text{E-}14$), which might indicate that lincRNAs are more variable than mRNAs, while the eQTL effects were similar in terms of R^2. These differences and similarities between coding transcripts and lincRNAs may indicate interesting mechanisms underlying the expressional regulations.

4) Application to GWAS results

The rationales behind utilizing eQTL mapping to interpret GWAS are that evidence from GWAS supports that transcriptional alterations contribute to risks of complex diseases; 1) a substantial fraction of GWAS-identified SNPs fell intergenic regions; and 2) eQTLs identified in previous study are enriched in GWAS-reported SNPs. Indeed, our eQTLs were also enriched in GWAS-reported SNPs: 1.7-fold for *cis*-eQTLs (one-sample proportion test $P < 2.2\text{E-}16$) and 3.7-fold for *trans*-eQTLs (one sample proportion test $P = 3.5\text{E-}15$). Interestingly, *trans*-eQTLs were more enriched than *cis*-eQTLs. We identified 386 records for which our eQTL map may provide another evidence to interpret GWAS results. We emphasize that our results of applying our eQTL map to GWAS interpretation can only suggest another possibilities for candidate causative genes based on expressional variations and that the significant association with expression does not necessarily indicate the gene is causative (an example was shown for *RPS26* and type I diabetes [33]). Thorough and close assessment is required for each case to conclude what gene is truly causative. Still, reviewing previous GWAS results while referring to eQTL maps, not only regarding *cis*-eQTLs but also *trans*-eQTLs, would be worthwhile, and eQTL maps will provide useful information for interpreting and understanding future GWAS results as well.

5) Replication

Cis-regulated genes identified in our study were in a good concordance with those identified by Westra *et al.* [20]: 60.9% of our *cis*-regulated genes were replicated. The 60% replication rate seems reasonable for whole blood samples because, in the current study, we replicated 56% of 112 *cis*-regulated genes identified in a

previous study [17] for whole blood samples from 76 Japanese individuals. On the other hand, replication of *trans*-eQTLs was challenging; only <1% of *trans*-eQTLs identified by Westra et al. [20] were replicated in the current study. Variation between different populations might be important for *trans*-eQTLs because we could replicate one of two *trans*-eQTLs in the previous study for the Japanese population [17]. We speculate the reason of low replication for *trans*-eQTLs as follows: Mechanisms of *trans*-effects of many sequence variations are considered as that a variant induces transcriptional alteration in a *cis* manner or functional change by substituting amino acids of proteins that involve in transcriptional regulation of other genes, and then, the locally induced change causes changes in expression levels of other genes [34]. Although *trans*-regulatory mechanisms are largely unknown, such a regulatory system may depend on a network of genes in which the genes interactively and cooperatively work in the same biological process; consequently, individual out-put gene expression levels are a cumulative result of a net effect of the whole network which could involve complex feedback mechanisms. The state of such a network should change dynamically with cell types, environmental conditions, and time. This is one of the reasons for the low reproducibility of *trans*-eQTLs. It should be noted that our *trans*-eQTLs were identified under just one set of conditions; therefore, the validity of applying our results to situations that represent different conditions needs to be carefully evaluated. However, we believe that our *trans*-eQTL analysis provides general insights into *trans*-effects, such as how effect magnitudes, β or R^2, are distributed.

Methods

Subjects and ethics statement

The study subjects were 301 apparently healthy individuals residing in Nagahama City, Japan. All participants provided written informed consent. The study protocol was approved by the Ethics Committee of Kyoto University Graduate School and Faculty of Medicine.

SNP genotyping and quality control

We extracted DNA from leukocytes and carried out genome-wide SNP genotyping with the Infinum HumanOmni5Exome BeadChip (Illumina, Inc., San Diego, CA, USA). We excluded any SNP with a missing rate >1%, Hardy-Weinberg equilibrium test P value<1E-07, minor allele frequency <5%, or that mapped to a sex chromosome. Ultimately, we examined a final set of 1,425,832 autosomal SNPs in the analysis. We excluded three samples from the analysis; one was excluded because of unsuccessful DNA extraction, and two others were excluded because of kinship with other sample. The snpStats package (http://www.bioconductor.org/packages/release/bioc/html/snpStats.html) in Bioconductor [35] was used to conduct the principal component analysis, and no subjects were identified as outliers relative to the HapMap JPT (Figure S2).

Gene expression profiles

Whole blood was collected from each participant when in a non-stimulated state; PAXgene Blood RNA Kits (QIAGEN, Hilden, Germany) were then used to collect samples of total RNA. For each participant, we used the Low Input Quick Amp Labeling Kit (Agilent Technologies, Inc., Santa Clara CA, USA) according to the manufacturer's protocol and 100 ng of total RNA to synthesize each labeled cRNA sample. We used Gene Expression Hybridization kits (Agilent Technologies, Inc.) to hybridize labeled cRNA to arrays from SurePrint G3 Human Gene Expression 8×60 K Microarray Kits (Agilent Technologies, Inc., design ID: 028004); Gene Expression Wash Packs (Agilent Technologies, Inc.) were then used according to the manufacturer's protocols to wash each microarray. Each microarray was scanned with a DNA Microarray Scanner (Agilent Technologies, Inc.), and Feature Extraction Ver.9.5.3 (Agilent Technologies, Inc.) was used to measure signal intensity.

Normalization and exclusion of expression data

The data were processed using the GeneSpringGX11 as follows. For each set of duplicated probes, the mean signal intensity was calculated. Signal intensities less than 1 were each set to 1, and each signal intensity value was transformed by taking the binary logarithm. Normalization was carried out by a 75th percentile shift; this normalization procedure was recommended by Agilent. After this normalization, the 75th percentile signal intensity of each chip was set to 0, at which point the signal values ranged from −7.3 to 12.3 with a median (mean) of −2.6 (−2.1).

We excluded 5,550 probes for which we were not able to obtain specific positions on the chromosomes of their target genes and 1,488 probes that were mapped on the sex chromosomes. We did not filter any probes based on expression abundance because the information that the transcript is not expressed might be of biological importance. However, signal values for low or non-expressed genes are often unreliable; therefore, we show median expression values for our eQTL-regulated transcripts provided in File S1; and to interpret the expression values Figure S3 shows how expression values were distributed for expressed or non-expressed transcripts.

Annotation of expression microarray probes

Annotation for gene expression probes of our chip (Agilent Technologies, Inc., design ID: 028004) was obtained from eArray (release date: 2012/04/11, build version: hg19:GRCh37:Feb2009, available online https://earray.chem.agilent.com/earray/). We defined three groups of probes: probes for *mRNA* transcripts, probes for *lincRNA* transcripts, and probes for *other* transcripts. Probes were classified into the *mRNA* group if they had assigned RefSeq NM accession numbers. *lincRNA* probes were indicated as such in Agilent's annotation. All the other probes were classified into the *other* group. The transcription start and end sites of genes represented by the probes that were classified into *mRNA* or *other* were obtained from a seq_gene.md file downloaded from the NCBI website (http://www.ncbi.nlm.nih.gov/accessed on 2013/02/20); and those represented by *lincRNA* probes were obtained from either Agilent's annotation or lincRNAsTranscripts table downloaded from the UCSC Genome Browser (http://genome.ucsc.edu/accessed on 2013/04/09).

Annotation of SNPs

BLAST was used to map probes from the SNP genotyping array into GRCh37; a rsID was assigned to each SNP based on its mapped chromosomal position on GRCh37. We defined a distance between a SNP and a gene as base pairs between the chromosomal position of the SNP and the position of the nearest transcription start/end site of the gene. If the SNP was located within the gene, then the distance was set to 0. Directions of genes were considered, and the sign associated with each distance indicated that the SNP was located upstream (negative) or downstream (positive) of the gene. ANNOVAR version 2013-05-09 [36] (http://www.openbioinformatics.org/annovar/) was used to annotate SNPs for classification into gene-structure-based categories; the RefSeq Gene (build version 19) was used as the reference. We annotated SNPs with ANNOVAR's default

definitions and precedence of SNP functional categories if a SNP was located within its target gene or within 1 kb-flanking regions of its target gene, and the gene name in the ANOVVAR annotation matched the target gene (if the gene name did not match, no specific functions were assigned); and otherwise, the SNP was categorized into *intergenic* (see supplementary note in File S3 for details). Using this method, we would classify an eQTL as intergenic if it was located outside its target gene even though it was located within another gene; in a different example, an eQTL in an intron of its target gene was classified as intronic even though it was located in any other category of another gene.

We classified each intergenic SNP into one of the regulatory potential classes as defined based on epigenetic information available in public databases by RegulomeDB [19] for dbSNP132 (downloaded from http://regulome.stanford.edu/on 2013/07/24). We were able to assign a regulatory classification to each of 1,396,242 SNPs (97.9% of the tested SNPs). We considered seven categories (Category 1–7) of regulatory classes as defined by the RegulomeDB, but we did not use the 15 subcategories (1a–f, 2a–c, 3a–b, 4–7). Briefly, lower scores indicated more evidence for the SNP being located in a regulatory region. Each known eQTL with known additional epigenetic functional annotation was assigned to Category 1. Category 2 requires direct evidence of binding through ChIP-seq and DNase. Category 3 requires a less complete set of evidence of binding. Categories 4–6 each comprised SNPs with minimal evidence of effects on transcription factor binding; Category 4 SNPs had DNase and ChIP-seq evidence; Category 5 SNPs had DNase or ChIP-seq evidence; and Category 6 had any single annotation not categorized above. Finally, Category 7 SNPs had no known evidence of TF binding.

eQTL mapping

We performed surrogate variable analysis [37] to identify unmodeled latent factors that cause heterogeneity in expression data. We identified two significant surrogate variables with age and gender used as known covariates using sva package (http://bioconductor.org/packages/release/bioc/html/sva.html) in Bioconductor [35,38]. We corrected expressions of each transcript for age, gender, and the two surrogate variables by fitting a multiple linear model in R version 3.0.2 (http://www.R-project.org/). We further excluded 4,972 probes that were mapped to regions with SNPs that was found polymorphic in the HapMap JPT samples or our study subjects because polymorphisms in such regions can alter hybridization efficiency; consequently, signal intensities may not reflect the actual amount of RNA [39–42]. The remaining 30,395 probes were included in the analysis. We assumed an additive model for all SNPs, and we coded each SNP genotypes as 0, 1, or 2, to represent the number of minor alleles in each individual. PLINK v1.07 [43] (http://pngu.mgh.harvard.edu/purcell/plink/) was used to perform the association analysis between each adjusted transcriptional phenotype and each of 1,425,832 autosomal SNPs with 298 individuals.

We define a *local* SNP as a SNP located on the same chromosome and within 500 kb from the nearest transcription start/end site of the gene that encodes the transcript, and a *distant* SNP, otherwise. We defined a *cis*-eQTL as a local SNP that significantly affects expression of a gene; similarly we defined a *trans*-eQTL as a distant SNP that significantly affects expression of a gene. We examined 16,986,695 local SNP-transcript pairs (11,028,260 for mRNAs, 3,485,407 for lincRNAs, and 2,473,028 for other transcripts). The mean number of local SNPs per probe was 560 (minimum 1, maximum 4,630). We examined about 43 billion distant SNP-transcript pairs. To identify *cis*-eQTLs, we estimated FDR with the permutation approach as described by Westra *et al.* [20]. Briefly, sample identifiers were permuted for 10 times, and only the local SNP with the smallest P value for each transcript was used to simulate the null distribution. With this approach we estimated FDR only for the SNP with the smallest P value for each transcript, and local SNPs with the FDR smaller than 5% were identified as *cis*-eQTLs. Therefore, no more than one *cis*-eQTL was identified for each transcript. If multiple SNPs in perfect LD ($r^2 = 1$) were the most significant with the same P value, the middle SNP was used to represent the eQTL. To exclude possible false discoveries caused by outliers or violation of normality assumptions, we performed Kruskal-Wallis test [44], a non-parametric test, and excluded *cis*-eQTL-transcript pairs with P value$>$0.00015 (see supplementary note in File S3).

To identify *trans*-eQTLs, we used the Bonferroni correction for multiple comparisons among the approximately 43 billion tests; only distant SNPs with nominal P values smaller than 1.15E-12, which corresponds to a family-wise error rate of 5%, were considered significant. We applied intensive exclusion criteria to obtain reliable *trans*-eQTLs. First, we excluded *trans*-eQTLs that may only capture *cis*-effects because of LD by a conditional regression on *cis*-eQTL genotypes (i.e., excluded when residuals of fitting *cis*-eQTL genotypes were not significantly associated with *trans*-eQTL genotypes by $P<0.05$). This analysis was performed when a *trans*-eQTL and its target transcript were located on the same chromosome, and a *cis*-eQTL was also identified for the transcript (*cis*-eQTLs excluded by Kruskal-Wallis tests were also considered). Second, we excluded redundant *trans*-eQTLs because of LD with other *trans*-eQTLs by sequential conditional regressions. For each transcript, *trans*-eQTLs on the same chromosome were iteratively tested starting from the *trans*-eQTL of the smallest P value for the transcript. If significant ($P<0.05$), the *trans*-eQTL is kept and residuals were used for the next iteration. If not, the *trans*-eQTL was excluded as redundant, and residuals were not taken for the next iteration. After this procedure, we tested *trans*-eQTLs that were found significant in the sequential conditional regression all together with a multiple linear regression, and non-significant *trans*-eQTLs ($P>0.05$) were further removed. Third, in order to confirm that the *trans*-eQTLs were not false positives because of cross-hybridization of probes to unexpected transcripts near the *trans*-eQTLs, we mapped the probe sequence to the flanking region (±500 kb) of its *trans*-eQTL by SHRiMP v.2.2.3 [45] for each probe-*trans*-eQTL combination. The human reference DNA sequence (GRCh37.p5) was downloaded from the NCBI (http://www.ncbi.nlm.nih.gov/). We used the same relaxed settings as Westra *et al.* [20] (match score of 10, mismatch score of 0, gap open penalty of −250, gap extension penalty of −100, and minimal Smith-Waterman score of 30%); -m 10 -i 0 -q -250 -f -100 -h 30%. We excluded a *trans*-eQTL if its associated probe was mapped to its flanking region. Fourth, we excluded low expression transcripts whose median expression levels were lower than −4.5 because we observed deviation from the distribution of median expression levels of *cis*-regulated transcripts (Figure S6). The cutoff was defined as the 5th percentile of the median expression levels of the *cis*-regulated transcripts. Kruskal-Wallis tests for the remaining SNP-transcript pairs were all significant ($P<0.00015$). We used the remaining *trans*-eQTLs in the further analyses.

The approach we used to correct for multiple testing with local SNPs differed from that used with distant SNPs because the high peak at low P values observed with local SNPs indicated that a substantial fraction of local SNPs were truly associated with the expression phenotype of one or more transcripts, whereas the uniform distribution of P values observed with distant SNPs indicated that the null hypothesis was true for most of the tests (Figure S1).

Identifying multi-regulatory eQTLs

We defined a multi-regulatory *cis*-eQTL as a *cis*-eQTL that is associated with expression levels of at least three different local protein-coding genes (assigned RefSeq NM accessions). For this, we did not count probes that cross-hybridize to other local genes associated with the same *cis*-eQTL by mapping probe sequences to the exon sequences with SHRiMP v.2.2.3 [45] using the same set of options used for detecting cross-hybridization for *trans*-eQTLs above.

Similarly, we defined a multi-regulatory *trans*-eQTL as a *trans*-eQTL that is associated with expression levels of at least 3 different distant protein-coding genes, after excluding cross-hybridized probes in the same procedure as used for multi-regulatory *cis*-eQTLs.

Statistical analysis for eQTLs

For comparison of mean effects of gene-based functional categories, we excluded SNPs that we were not able to assign to a specific category; we also excluded categories that comprised fewer than 5 eQTLs. Values of $|\beta|$ and R^2 were log-transformed and then subjected to the ANOVA; the ANOVA was followed by Tukey's HSD test (which performs all pairwise comparisons between two subcategories for multiple testing correction). The trend of log-transformed $|\beta|$ and R^2 values with seven RegulomeDB classes (Classes 1a–f, 2a–c and 3a–b were grouped as 1, 2 and 3, respectively) was tested with Jonckheere-Terpstra permutation test (one-sided, 100,000 permutations) provided in clinfun package (http://cran.r-project.org/web/packages/clinfun/index.html) in R version 3.0.2 (http://www.R-project.org/). r^2 of LD between SNPs were computed with PLINK v1.07 [43].

Replication analysis

We downloaded the eQTL map by Westra *et al.* [20] from their browser (http://genenetwork.nl/bloodeqtlbrowser/), and annotation files for HT12v3 and Agilent Human Genome 4×44 K array from the GEO (http://www.ncbi.nlm.nih.gov/geo/). We matched Entrez GeneIDs to compare with the replication studies. We referred to GWAS catalog to obtain SNPs tested for *trans*-eQTL in [20] (SNPs reported by 16, July, 2011). We found 978 distant SNP-transcript pairs in our study that corresponded to rs IDs and Entrez Gene IDs tested in [20].

Matching eQTLs with GWAS-identified SNPs

We downloaded 16,541 public GWAS records from the NHRGI GWAS Catalog (http://www.genome.gov/gwastudies/ accessed on 2014/04/23). We excluded 323 records with reported *P* values that were not significant (reported as NS or Pending); we excluded another 6,142 records because the reported SNPs were not included in our tested SNPs. Ultimately, we examined 10,076 records for 8,069 unique SNPs reported by 1,436 GWAS. We matched GWAS-reported SNPs to our eQTLs when they exactly matched or were in LD ($r^2>0.8$). We excluded records if conditional regression on genotypes of a GWAS-identified SNP was significant $P<0.05$ (File S2) because they might be false discoveries of trait-eQTL association where eQTL and GWAS-identified SNP are two different genetic factors [33]. When matching gene symbols, we also searched their aliases downloaded from the HGNC BioMart version 0.7 (http://www.genenames.org/biomart/accessed on 2013/09/28).

Accession numbers

Our expression microarray data are available at the NCBI's Gene Expression Omnibus under accession number GSE53351.

Supporting Information

Figure S1 Histogram of *P* values of all association tests. A) Histogram of *P* values obtained from the 16,986,695 association tests between all autosomal transcripts and local SNPs. The excess of smaller *P* values indicates that a substantial fraction of associations are truly positive. B) Histogram of *P* values obtained from about 43 billion association tests between all autosomal transcripts and distant SNPs. The almost uniformly distributed *P* values suggests that most of distant SNPs have no effects on transcriptional regulation, though a slight increase at the low *P* values in frequency indicates a tiny fraction of distant SNPs are truly positive. Also see File S3 for a comment about influence of surrogate variable analysis on the distribution.
(TIF)

Figure S2 Principal component analysis of study population in comparison with HapMap samples. The first and second principal components are shown. CEU: Utah residents with Northern and Western European ancestry from the CEPH collection; YRI: Yoruba in Ibadan, Nigeria; JPT: Japanese in Tokyo; CHB: Han Chinese in Beijing, China; Sample: samples of the current study.
(TIF)

Figure S3 Distribution of normalized expression data. Distribution of normalized expression data for all 42,405 probes and 298 samples are shown. "A" (absent) if a foreground signal is <2.6 SD of background signal; "M" (marginal) if it was saturated, not uniform in a spot, or not uniform among replicated probes, or "P" (present) otherwise. The number following each class name is the number of data classified into the class.
(TIF)

Figure S4 Regression coefficients of multi-regulatory eQTLs. Regression coefficients, β, of each multi-regulatory *cis*-eQTLs (A) or *trans*-eQTLs (B) are shown. Directions of effects of each multi-regulatory eQTL are consistent.
(TIF)

Figure S5 *Trans*-eQTL map. A) Chromosomal positions of *trans*-eQTLs are plotted against chromosomal positions of associated transcripts. B) $-\log_{10} P$ values of *trans*-eQTLs are plotted against the respective chromosomal positions. (C) $-\log_{10} P$ values of *trans*-eQTLs are plotted against the chromosomal positions of associated transcripts. The horizontal and vertical dashed lines separate chromosomes; the diagonal dashed line indicates that the *trans*-eQTL is located at the same chromosomal positions as transcripts. mRNA transcripts are shown in red; lincRNA transcripts are shown in green; and other transcripts are shown in black. $-\log_{10} P$ values are truncated at 50, and a triangle indicate truncation.
(TIF)

Figure S6 Median expression levels of *cis*-regulated or *trans*-regulated genes.
(TIF)

Table S1 Demographic characteristics of study subjects.
(DOCX)

Table S2 *P* values of Tukey's HSD test.
(DOCX)

Table S3 Replicated *trans*-eQTLs identified by Westra *et al.* [20].
(XLSX)

File S1 Annotations and statistics of *cis*-eQTLs and *trans*-eQTLs.
(XLS)

File S2 Results of our application of eQTL map to GWAS records. Sheet "Case1–3" shows records classified into Case 1, 2, or 3; sheet "Case 4" shows records classified into Case 4; sheet "Excluded" shows excluded records because GWAS SNP and eQTL are not likely to colocalize.
(XLSX)

File S3 Supplementary notes.
(PDF)

References

1. Göring HHH, Curran JE, Johnson MP, Dyer TD, Charlesworth J, et al. (2007) Discovery of expression QTLs using large-scale transcriptional profiling in human lymphocytes. Nat Genet 39: 1208–1216.
2. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, et al. (2004) Genetic analysis of genome-wide variation in human gene expression. Nature 430: 743–747.
3. Dixon AL, Liang L, Moffatt MF, Chen W, Heath S, et al. (2007) A genome-wide association study of global gene expression. Nat Genet 39: 1202–1207.
4. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, et al. (2005) Mapping determinants of human gene expression by regional and genome-wide association. Nature 437: 1365–1369.
5. Stranger BE, Forrest MS, Clark AG, Minichiello MJ, Deutsch S, et al. (2005) Genome-wide associations of gene expression variation in humans. PLoS Genet 1: e78.
6. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science 315: 848–853.
7. Stranger BE, Nica AC, Forrest MS, Dimas A, Bird CP, et al. (2007) Population genomics of human gene expression. Nat Genet 39: 1217–1224.
8. Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, et al. (2007) A survey of genetic human cortical gene expression. Nat Genet 39: 1494–1499.
9. Mehta D, Heim K, Herder C, Carstensen M, Eckstein G, et al. (2013) Impact of common regulatory single-nucleotide variants on gene expression profiles in whole blood. Eur J Hum Genet 21: 48–54.
10. Spizzo R, Almeida MI, Colombatti A, Calin GA (2012) Long non-coding RNAs and cancer: a new frontier of translational research? Oncogene 31: 4577–4587.
11. Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. Annu Rev Biochem 81: 145–166.
12. Risch N, Merikangas K (1996) The Future of Genetic Studies of Complex Human Diseases. Science 273: 1516–1517.
13. Spielman RS, Bastone LA, Burdick JT, Morley M, Ewens WJ, et al. (2007) Common genetic variants account for differences in gene expression among ethnic groups. Nat Genet 39: 226–231.
14. Zhang W, Duan S, Kistner EO, Bleibel WK, Huang RS, et al. (2008) Evaluation of genetic variation contributing to differences in gene expression between populations. Am J Hum Genet 82: 631–640.
15. Bushel PR, McGovern R, Liu L, Hofmann O, Huda A, et al. (2012) Population differences in transcript-regulator expression quantitative trait loci. PLoS One 7: e34286.
16. Duan S, Huang RS, Zhang W, Bleibel WK, Roe CA, et al. (2008) Genetic architecture of transcript-level variation in humans. Am J Hum Genet 82: 1101–1113.
17. Sasayama D, Hori H, Nakamura S, Miyata R, Teraishi T, et al. (2013) Identification of single nucleotide polymorphisms regulating peripheral blood mRNA expression with genome-wide significance: an eQTL study in the Japanese population. PLoS One 8: e54967.
18. Cookson W, Liang L, Abecasis G, Moffatt M, Lathrop M (2009) Mapping complex disease traits with global gene expression. Nat Rev Genet 10: 184–194.
19. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, et al. (2012) Annotation of functional variation in personal genomes using RegulomeDB. Genome Res 22: 1790–1797.
20. Westra H-J, Peters MJ, Esko T, Yaghootkar H, Schurmann C, et al. (2013) Systematic identification of trans eQTLs as putative drivers of known disease associations. Nat Genet 45: 1238–1243.
21. Franke A, McGovern DPB, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118–1125.
22. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, et al. (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40: 955–962.
23. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, et al. (2007) Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. Nat Genet 39: 830–832.
24. Li J, Glessner JT, Zhang H, Hou C, Wei Z, et al. (2013) GWAS of blood cell traits identifies novel associated loci and epistatic interactions in Caucasian and African-American children. Hum Mol Genet 22: 1457–1464.
25. Gieger C, Radhakrishnan A, Cvejic A, Tang W, Porcu E, et al. (2011) New gene functions in megakaryopoiesis and platelet formation. Nature 480: 201–208.
26. Tse K-P, Su W-H, Chang K-P, Tsang N-M, Yu C-J, et al. (2009) Genome-wide association study reveals multiple nasopharyngeal carcinoma-associated loci within the HLA region at chromosome 6p21.3. Am J Hum Genet 85: 194–203.
27. Ma H-Q, Liang X-T, Zhao J-J, Wang H, Sun J-C, et al. (2009) Decreased expression of Neurensin-2 correlates with poor prognosis in hepatocellular carcinoma. World J Gastroenterol 15: 4844–4848.
28. Brown CD, Mangravite LM, Engelhardt BE (2013) Integrative modeling of eQTLs and cis-regulatory elements suggests mechanisms underlying cell type specificity of eQTLs. PLoS Genet 9: e1003649.
29. Veyrieras J-B, Kudaravalli S, Kim SY, Dermitzakis ET, Gilad Y, et al. (2008) High-resolution mapping of expression-QTLs yields insight into human gene regulation. PLoS Genet 4: e1000214.
30. Fairfax BP, Makino S, Radhakrishnan J, Plant K, Leslie S, et al. (2012) Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. Nat Genet 44: 502–510.
31. Small KS, Hedman AK, Grundberg E, Nica AC, Thorleifsson G, et al. (2011) Identification of an imprinted master trans regulator at the KLF14 locus related to multiple metabolic phenotypes. Nat Genet 43: 561–564.
32. Fehrmann RSN, Jansen RC, Veldink JH, Westra H-J, Arends D, et al. (2011) Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. PLoS Genet 7: e1002197.
33. Plagnol V, Smyth DJ, Todd JA, Clayton DG (2009) Statistical independence of the colocalized association signals for type 1 diabetes and RPS26 gene expression on chromosome 12q13. Biostatistics 10: 327–334.
34. Cheung VG, Nayak RR, Wang IX, Elwyn S, Cousins SM, et al. (2010) Polymorphic cis- and trans-regulation of human gene expression. PLoS Biol 8: e1000480.
35. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.
36. Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38: e164.
37. Leek JT, Storey JD (2007) Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet 3: 1724–1735.
38. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28: 882–883.
39. Alberts R, Terpstra P, Li Y, Breitling R, Nap J-P, et al. (2007) Sequence polymorphisms cause many false cis eQTLs. PLoS One 2: e622.
40. Benovoy D, Kwan T, Majewski J (2008) Effect of polymorphisms within probe-target sequences on oligonucleotide microarray experiments. Nucleic Acids Res 36: 4417–4423.
41. Walter NAR, McWeeney SK, Peters ST, Belknap JK, Hitzemann R, et al. (2007) SNPs matter: impact on detection of differential expression. Nat Methods 4: 679–680.
42. Sliwerska E, Meng F, Speed TP, Jones EG, Bunney WE, et al. (2007) SNPs on chips: the hidden genetic code in expression arrays. Biol Psychiatry 61: 13–16.
43. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559–575.
44. Kruskal WH, Wallis WA (1952) Use of Ranks in One-Criterion Variance Analysis. J Am Stat Assoc 47: 583–621.
45. David M, Dzamba M, Lister D, Ilie L, Brudno M (2011) SHRiMP2: sensitive yet practical SHort Read Mapping. Bioinformatics 27: 1011–1012.
46. Yang S-K, Hong M, Zhao W, Jung Y, Baek J, et al. (2014) Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations. Gut 63: 80–87.

Acknowledgments

We would like to offer our special thanks to Non-profit Organization "Zero-ji Club for Health Promotion", the staff and associates of the Nagahama City Office, Kohoku Medical Association, Nagahama City Hospital, Nagahama Red Cross Hospital, Nagahama Kohoku City Hospital, and the participants of the Nagahama Study for data collection.

Author Contributions

Conceived and designed the experiments: FM YT. Performed the experiments: SN MI KM. Analyzed the data: MN RY. Contributed reagents/materials/analysis tools: YT TK. Wrote the paper: MN RY. Developed the Human Genome Variation Browser: KH.

INTERNAL MEDICINE

□ CASE REPORT □

Successful Use of Intensive Immunosuppressive Therapy for Treating Simultaneously Occurring Cerebral Lesions and Pulmonary Arterial Hypertension in a Patient with Systemic Lupus Erythematosus

Ryu Watanabe[1], Hiroshi Fujii[1], Tsuyoshi Shirai[1], Shinichiro Saito[1], Akira Hatakeyama[2], Koichiro Sugimura[3], Yoshihiro Fukumoto[3], Tomonori Ishii[1] and Hideo Harigae[1]

Abstract

A 59-year-old woman who had been diagnosed with systemic lupus erythematosus (SLE) was admitted to our hospital due to paralysis in all of her limbs. The patient presented with dysarthria, cerebellar ataxia and hypoxia. Magnetic resonance imaging (MRI) revealed vasogenic edema in the brain stem and the cerebellum. She was diagnosed with neuropsychiatric lupus syndrome (NPSLE) and pulmonary arterial hypertension (PAH), and was successfully treated using immunosuppressive therapy. To our knowledge, this is the first reported case of simultaneously developing NPSLE and PAH.

Key words: cerebral lesion, pulmonary arterial hypertension, systemic lupus erythematosus

(Intern Med 53: 627-631, 2014)
(DOI: 10.2169/internalmedicine.53.0514)

Introduction

Systemic lupus erythematosus (SLE) is an autoimmune disease that is characterized by the production of pathogenic autoantibodies which results in damage to multiple organs (1). Central nervous system (CNS) involvement is one of the major manifestations of SLE and occurs in approximately 15% to 75% of lupus patients (2). In 1999, a multidisciplinary committee of the American College of Rheumatology published the nomenclature for neuropsychiatric lupus syndromes (NPSLE). The neuropsychiatric syndromes were divided into 19 different conditions which included the neurologic disorders of the central, peripheral and autonomic nervous system as well as the psychiatric syndromes (3). Although technological advances in neuroimaging have proved useful in monitoring brain damage, the diagnosis of NPSLE is difficult and requires a careful assessment. NPSLE still accounts for 4% to 16% of the deaths of lupus patients (4). Pulmonary arterial hypertension (PAH) is sometimes associated with connective tissue diseases (CTD) such as systemic sclerosis (SSc), mixed connective tissue diseases (MCTD) and SLE. The prevalence of PAH is estimated to be 0.5% to 17.5% in SLE patients (5, 6). PAH is also associated with a poor prognosis, and the three-year survival rate of SLE-PAH patients has only been reported as 75% (7). We herein describe a case of SLE that was simultaneously diagnosed with NPSLE and PAH. Although each of these manifestations may not be rare in SLE, this is the first reported case to have concurrently developed both complicating conditions. The patient was successfully treated with intensive immunosuppressive therapy.

Case Report

A 59-year-old woman was admitted to our hospital due to paralysis in all of her limbs. She had previously been diagnosed with SLE based on polyarthralgia, facial rash and serological tests showing positivity for anti-nuclear antibody (ANA, ×160, speckled pattern) and anti-Smith antibody. Her

[1]Department of Hematology and Rheumatology, Tohoku University Graduate School of Medicine, Japan, [2]Division of Arthritis and Connective Tissue Diseases, Tohoku Rosai Hospital, Japan and [3]Department of Cardiovascular Medicine, Tohoku University Graduate School of Medicine, Japan
Received for publication March 4, 2013; Accepted for publication May 19, 2013
Correspondence to Dr. Hiroshi Fujii, hfujii@med.tohoku.ac.jp

Table 1. **Laboratory Findings of the Patient**

Complete blood cell counts			Urinalysis					
WBC	5,500	/μL	protein	(1+)		Na	133	mEq/L
Seg	90	%		0.8	g/g·cr	K	3.0	mEq/L
Lym	7	%	occult.blood.	(2+)		Cl	97	mEq/L
Mon	3	%	<sediment>			BUN	19	mg/dL
Eos	0	%	RBC	10-29	/HPF	Cr	0.6	mg/dL
Bas	0	%	Cast	10-29	/LPF	UA	4.8	mg/dL
RBC	454×10^4	/μL	Biochemistry			Ferritin	264	ng/mL
Hb	14.6	g/dL	T.Bil	1.4	mg/dL	CRP	1.1	mg/dL
MCV	94.8	fl	ALP	132	IU/L	C3	47	mg/dL
Hct	43	%	γ-GTP	30	IU/L	C4	5.8	mg/dL
Ret	0.6	%	AST	25	IU/L	CH50	25.4	U/mL
Plt	11.1×10^4	/μL	ALT	15	IU/L	ANA	80	fold
Coagulation			LDH	254	IU/L	dsDNA	6.1	IU/mL
PT.INR	0.96		TP	7.2	g/dL	Sm	133.1	index
APTT	30.6	sec	Alb	3.5	g/dL	RNP	179.3	index
Fbg	347	mg/dL	Haptoglobin	22.2	mg/dL	SS-A	97.1	index
D-Dimer	1.5	μg/mL	KL-6	604	U/mL	SS-B	9.0	index
LAC	1.2		BNP	162	pg/mL	β2GP1CL	< 1.3	U/mL
※LAC: lupus anticoagulant			HbA1c	6.1	%	Cardiolipin	6.0	U/mL

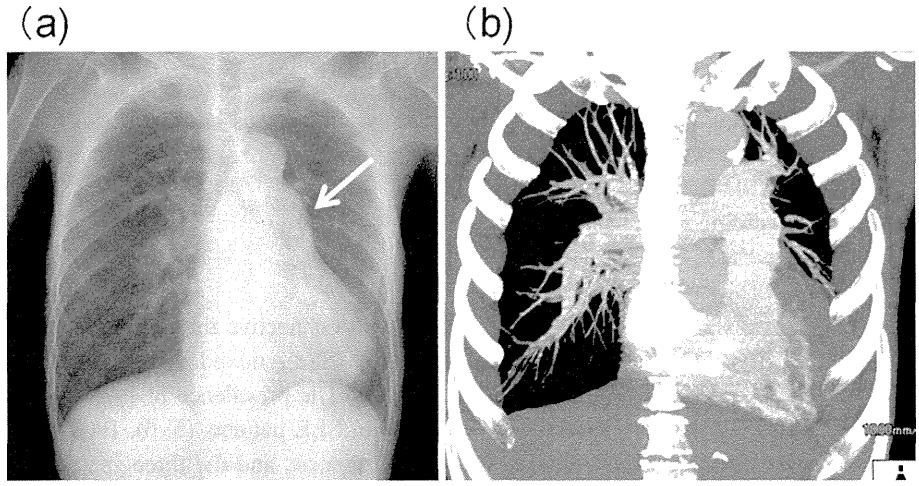


Figure 1. **The chest X-ray and CT findings upon admission. (a) The chest X-ray showed a protrusion of the left second arch (arrow). (b) Enhanced CT showed no signs of pulmonary embolism.**

symptoms had been well controlled with low dose prednisolone (PSL) for the five years leading up to the current admission. Upon admission, her blood pressure was 178/108 mmHg, her body temperature was 37.8℃, her heart rate was 126 beats/min and her oxygen saturation level (SpO_2) was 93%. A physical examination showed facial and palmar erythema and pretibial edema. A neurological examination revealed that her consciousness was slightly altered and her Glasgow Coma Scale (GCS) score was 14/15. In addition, dysarthria and cerebellar ataxia were also observed. Bilateral manual muscle testing (MMT) produced an upper limb score of 4/5 and lower limb score of 3/5. Laboratory tests demonstrated positive results for anti-RNP, anti-Smith and anti-SS-A antibodies as well as hypocomplementemia, an elevated brain natriuretic peptide (BNP) level and abnormal urinary findings (Table 1). Chest X-rays showed a protrusion of the left second arch of the cardiac silhouette and a cardiothoracic ratio of 59.6% (Fig. 1a). Enhanced computed tomography (CT) revealed no evidence of either pulmonary embolism or interstitial pneumonia, but right ventricular hypertrophy and a small amount of pericardial effusion were observed (Fig. 1b). Echocardiography showed the ejection fraction to be normal (79.6%), but the tricuspid pressure gradient (TRPG) was elevated (50 mmHg). Pulmonary scintigraphy showed no signs of any blood flow defects.

Magnetic resonance imaging (MRI) of the brain revealed multiple high intensity signals in the brain stem and the bilateral cerebellum on a T2-weighted image (T2WI) and a fluid-attenuated inversion recovery (FLAIR) image (Fig. 2b, c). The T1-weighted images (T1WI) and diffusion-weighted images (DWI) of these lesions were almost normal, thus suggesting vasogenic edema (Fig. 2a, c). Cerebral blood flow scintigraphy showed a significant decrease in the flow to the bilateral cerebellum and the right frontal and

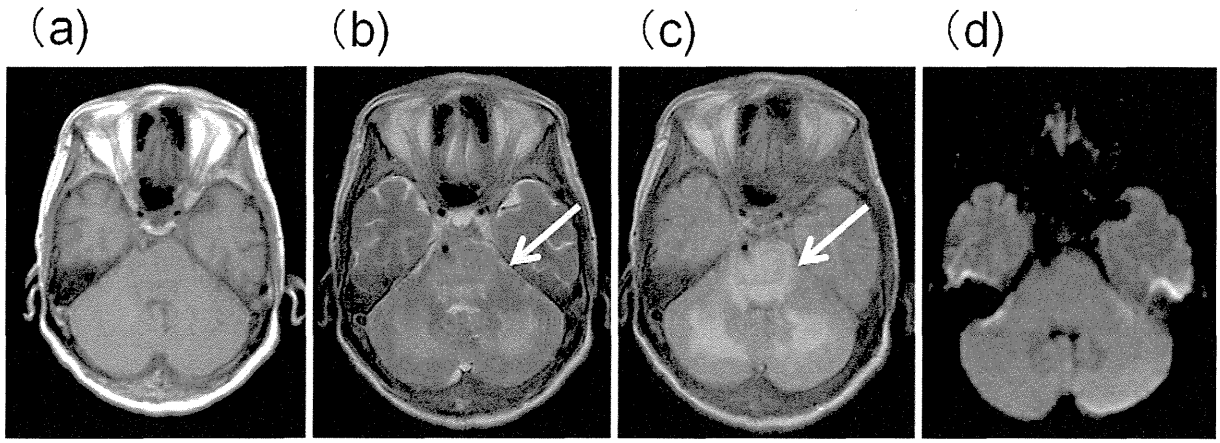


Figure 2. **MRI findings of the brain stem and cerebellum opon admission. (a) T1WI. (b, c) T2WI and FLAIR. High intensity signals (arrow) were observed. (d) DWI.**

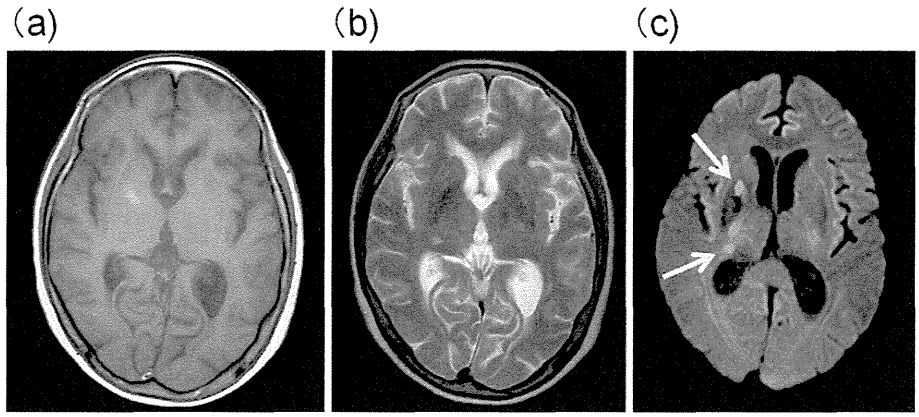


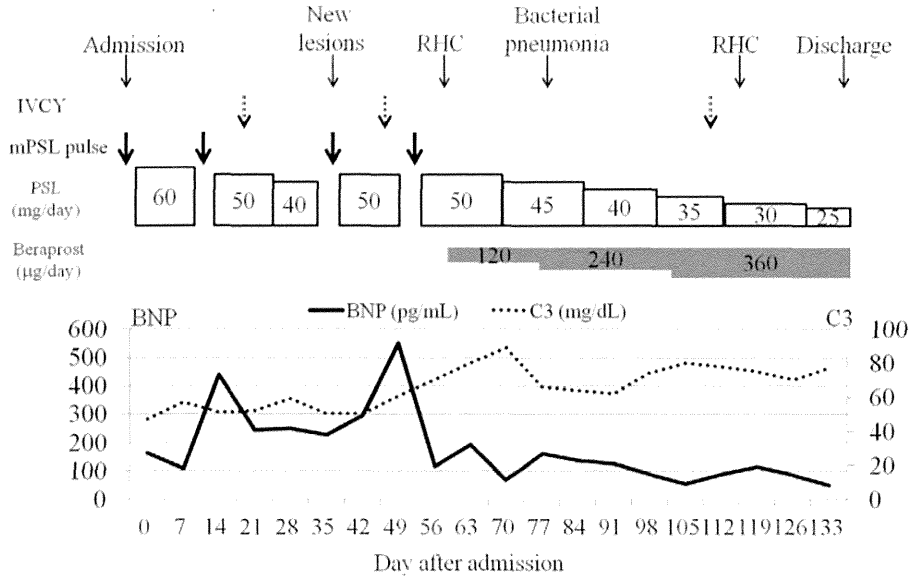
Figure 3. **MRI findings of the right thalamus and caudate nucleus one month after admission. (a) T1WI. (b) T2WI. (c) DWI. High intensity signals (arrow) were observed.**

temporal lobes. An electroencephalogram (EEG) showed diffuse slow waves (8-10 Hz) with multiple bursts. An examination of the cerebrospinal fluid revealed normocytosis and a slightly increased concentration of total protein (50 mg/dL: normal, 10-40 mg/dL), but the IgG index was within normal limits (0.76). The SLE disease activity index (SLEDAI) (8) was 35. From these findings, we concluded that SLE-associated NPSLE and PAH had developed simultaneously in this patient.

Intravenous steroid pulse therapy followed by high dose PSL therapy (60 mg/day), intravenous cyclophosphamide pulse therapy (IVCY, 500 mg/day) and anti-coagulant therapy dramatically improved the dysarthria and ataxia, and the cerebellar and brain stem lesions disappeared rapidly after two weeks. However, new high intensity signals on DWI that indicated cerebral infarction were observed in the right thalamus and caudate nucleus at one month after her admission (Fig. 3). Magnetic resonance angiography (MRA) revealed no vascular stenosis, embolism or aneurysm. A repeated course of intravenous steroid pulse therapy led to improvement after two weeks. The complete clinical course is shown in Fig. 4. A right heart catheterization (RHC) was performed two months after admission, and revealed a mean pulmonary arterial pressure (mPAP) of 36 mmHg, a pulmonary capillary wedge pressure (PCWP) of 3 mmHg, a cardiac index (CI) of 3.52 L/min and a pulmonary vascular resistance (PVR) of 526 dyne·sec/cm^5, thus leading to a diagnosis of PAH and the administration of beraprost. Since then, no new cerebral lesions have been observed even though the patient suffered from bacterial pneumonia during her hospital stay. She was discharged four months after admission. RHC performed prior to discharge showed that her mPAP and PVR had decreased remarkably (Table 2). Ten rounds of IVCY were sufficient to maintain both the NPSLE and PAH in remission for 3 years with maintenance low dose PSL therapy (10 mg/day) and no required additional vasodilative therapy. Her current SLEDAI score is 2.

Discussion

Neuropsychiatric manifestations are well known to be a serious complication associated with SLE. Previous reports have suggested that pathogenic autoantibodies such as the anti-phospholipid antibody, the anti-ribosomal P antibody and the anti-N-methyl-D-aspartate (NMDA) antibody, as well as inflammatory cytokines such as interleukin (IL)-2,



IVCY: intravenous cyclophosphamide, PSL: prednisolone, RHC: right heart catheterization

Figure 4. The clinical course of the patient.

IL-6, IL-8, IL-10, tumor necrosis factor (TNF)-α and interferon (IFN)-α are key players in the pathogenesis of NPSLE (9). Intensive immunosuppressive therapy such as high dose PSL and cyclophosphamide is required for the treatment of NPSLE (10).

MRI is one of the most common methods used in clinical practice to evaluate CNS involvement in lupus patients. It allows for a very sensitive detection of infarctions, hemorrhages and acute myelitis, and it can be used to monitor the response to therapy (11). DWI measures the diffusivity of water protons and has been increasingly used to distinguish cytotoxic edema in acute infarction from vasogenic edema and chronic infarction (2, 12). In our patient, MRI of the brain revealed two different cerebral lesions. First, high intensity signals on T2WI and FLAIR, which showed isointensity signals on T1WI and DWI, were observed in the brain stem and the bilateral cerebellum indicating vasogenic edema (Fig. 2a-d). This is commonly seen in the bilateral parieto-occipital subcortical white matter, and the condition is known as reversible posterior encephalopathy syndrome (RPLS) (12). However, these lesions can also occur in the frontal lobe, basal ganglia, thalamus, cerebellum, and brain stem (12). Most of the cases of RPLS that are observed in lupus patients are associated with triggers such as hypertension, preeclampsia, or with the administration of immunosuppressive agents. However, RPLS can also occur as a neurological manifestation of active lupus and sometimes requires intensive immunosuppressive therapy (13). Recently, RPLS has been increasingly considered to be one of the neuropsychiatric syndromes of active lupus (14). In this patient, an anti-hypertensive agent was not administered immediately following her admission because of the significantly decreased cerebral blood flow that was observed in the bilateral cerebellum and the right frontal and temporal lobes. The rapid response to the administered immunosuppressive therapy without irreversible changes suggested that these lesions were a vasogenic edema that associated with active lupus. Second, high intensity signals on DWI observed in the right thalamus and caudate nucleus indicated cerebral infarction (Fig. 3). These lesions were probably caused by a decreased cerebral blood flow. As expected, they rapidly improved after treatment. Therefore, the two cerebral lesions in this patient were both radiographically and mechanistically different.

Table 2. Hemodynamics of the Patient

Duration after admission (months)	2	4	6
PAP (mmHg)	56/22 (36)	42/15 (25)	32/14 (18)
CO (CI) (L/min)	4.26 (3.52)	5.5 (4.23)	3.47 (2.67)
PVR (dyne・sec・cm^{-5})	526	291	346
BNP (pg/mL)	293.5	74.5	52.5

PAP: pulmonary arterial pressure, CO (CI): cardiac output (index), PVR: pulmonary vascular resistance, BNP: brain natriuretic peptide

PAH is defined by an mPAP of greater than 25 mmHg at rest and a PCWP of less than 15 mmHg. It has been increasingly recognized that inflammatory mechanisms could play an important role in the PAH pathogenesis and progression, particularly in patients with CTD (15). PAH associated with CTD, but not systemic sclerosis, responds well to intensive immunosuppressive therapy (16-18). In this patient, immunosuppressive therapy dramatically improved her pulmonary hemodynamics. A follow-up RHC that was performed six months after the start of her admission revealed that her mPAP had completely normalized (Table 2). No recurrence of PAH was observed for 3 years. These findings showed that PAH associated with active lupus could respond

to intensive immunosuppressive therapy.

NPSLE and PAH are sometimes observed in lupus patients; however, the simultaneous occurrence of both manifestations is very rare. Hardie et al. reported a 28-year-old woman who was diagnosed with SLE and then tetraplegia developed PAH 6 years after her original diagnosis (19). Funauchi et al. reported that 6 out of 306 lupus patients (1.9%) had both NPSLE and PAH (20). Cefle et al. also reported that 4 out 107 patients (3.7%) with SLE had both conditions (21). These reports suggested that the complication of these manifestations does occur in lupus patients; however, there are currently no case reports detailing the simultaneous development of NPSLE and PAH. Therefore, to the best of our knowledge, this is the first case report of a lupus patient who was concurrently diagnosed with both conditions.

The mechanism that caused both manifestations has not yet been elucidated. This patient may have several different pathogenic autoantibodies or an atypical autoantibody that caused both conditions. Vascular endothelial cell injuries may have been involved in the pathogenesis. Anti-endothelial cell antibodies (AECAs) are often detected in lupus patients and are considered to play important roles in the development of nephritis and atherosclerotic lesions related to vascular endothelial injuries (22). AECA was detected in the sera of this patient when we measured the binding activity of IgG to human umbilical vein endothelial cells (HUVECs) using flow cytometry (data not shown) (22). This AECA activity may have the potential to cause both manifestations.

In conclusion, we herein presented a case of SLE in which CNS involvement and PAH developed concurrently. Intensive immunosuppressive therapy was very effective for treating both conditions, thus indicating that both of these manifestations were mediated by autoimmune mechanisms. This case report may provide some useful insights concerning the pathogenesis of NPSLE and PAH.

The authors state that they have no Conflict of Interest (COI).

References

1. Liu Z, Davidson A. Taming lupus-a new understanding of pathogenesis is leading to clinical advances. Nat Med **18**: 871-882, 2012.
2. Castellino G, Govoni M, Giacuzzo S, Trotta F. Optimizing clinical monitoring of central nervous system involvement in SLE. Autoimmun Rev **7**: 297-304, 2008.
3. ACR Ad Hoc Committee on Neuropsychiatric lupus nomenclature. The American College of Rheumatology nomenclature and case definitions for neuropsychiatric lupus syndromes.. Arthritis Rheum **42**: 599-608, 1999.
4. Borchers AT, Aoki CA, Naguwa SM, Keen CL, Shoenfeld Y, Gershwin ME. Neuropsychiatric features of systemic lupus erythematosus. Autoimmun Rev **4**: 329-344, 2005.
5. Dhala A. Pulmonary arterial hypertension in systemic lupus erythematosus: current status and future direction. Clin Dev Immunol **2012**: 854941, 2012.
6. Johnson SR, Granton JT. Pulmonary hypertension in systemic sclerosis and systemic lupus erythematosus. Eur Respir Rev **20**: 277-286, 2011.
7. Condliffe R, Kiely DG, Peacock AJ, et al. Connective tissue disease-associated pulmonary arterial hypertension in the modern treatment era. Am J Respir Crit Care Med **179**: 151-157, 2009.
8. Bombardier C, Gladman DD, Urowitz MB, Caron D, Chang CH. Derivation of the SLEDAI. A disease activity index for lupus patients. The Committee on Prognosis Studies in SLE. Arthritis Rheum **35**: 630-640, 1992.
9. Efthimiou P, Blanco M. Pathogenesis of neuropsychiatric systemic lupus erythematosus and potential biomarkers. Mod Rheumatol **19**: 457-468, 2009.
10. Bertsias GK, Ioannidis JPA, Aringer M, et al. EULAR recommendations for the management of systemic lupus erythematosus with neuropsychiatric manifestations: report of a task force of the EULAR standing committee for clinical affairs. Ann Rheum Dis **69**: 2074-2082, 2010.
11. Appenzeller S, Pike GB, Clarke AE. Magnetic resonance imaging in the evaluation of central nervous system manifestations in systemic lupus erythematosus. Clin Rev Allergy Immunol **34**: 361-366, 2008.
12. Moritani T, Hiwatashi A, Shrier DA, Wang HZ, Numaguchi Y, Westesson PL. CNS vasculitis and vasculopathy. Efficasy and usefulness of diffusion-weighted echoplanar MR imaging. Clin Imaging **28**: 261-270, 2004.
13. Fujieda Y, Kataoka H, Odani T, et al. Clinical features of reversible posterior leukoencephalopathy syndrome in patients with systemic lupus erythematosus. Mod Rheumatol **21**: 276-281, 2011.
14. Barber CE, Leclerc R, Gladman DD, Urowitz MB, Fortin PR. Posterior reversible encephalopathy syndrome: an emerging disease manifestation in systemic lupus erythematosus. Semin Arthritis Rheum **41**: 353-363, 2011.
15. Price LC, Wort SJ, Perros F, et al. Inflammation in pulmonary arterial hypertension. Chest **141**: 210-221, 2012.
16. Sanchez O, Sitbon O, Jais X, Simonneau G, Humbert M. Immunosuppressive therapy in connective tissue disease-associated pulmonary arterial hypertension. Chest **130**: 182-189, 2006.
17. Jais X, Launay D, Yaici A, et al. Immunosuppressive therapy in lupus- and mixed connective tissue disease-associated pulmonary arterial hypertension. Arthritis Rheum **58**: 521-531, 2008.
18. Miyamichi-Yamamoto S, Fukumoto Y, Sugimura K, et al. Intensive immunosuppressive therapy improves pulmonary hemodynamics and long-term prognosis in patients with pulmonary arterial hypertension associated with connective tissue disease. Circ J **75**: 2668-2674, 2011.
19. Hardie RJ, Isenberg DA. Tetraplegia as a presenting feature of systemic lupus erythematosus complicated by pulmonary hypertension. Ann Rheum Dis **44**: 491-493, 1985.
20. Funauchi M, Shimadzu H, Tamaki C, et al. Survival study by organ disorders in 306 Japanese patients with systemic lupus erythematosus. Rheumatol Int **27**: 243-249, 2007.
21. Cefle A, Inanc M, Sayarlioglu M, et al. Pulmonary hypertension in systemic lupus erythematosus: relationship with antiphospholipid antibodies and severe disease outcome. Rheumatol Int **31**: 183-189, 2011.
22. Shirai T, Fujii H, Ono M, et al. A novel autoantibody against fibronectin leucine-rich transmembrane protein 2 expressed on the endothelial cell surface identified by retroviral vector system in systemic lupus erythematosus. Arthritis Res Ther **14**: R157, 2012.

OPEN ACCESS Freely available online

PLOS ONE

The Multicenter Study of a New Assay for Simultaneous Detection of Multiple Anti-Aminoacyl-tRNA Synthetases in Myositis and Interstitial Pneumonia

Ran Nakashima[1☯], Yoshitaka Imura[1☯], Yuji Hosono[1], Minae Seto[2], Akihiro Murakami[2], Kizuku Watanabe[3], Tomohiro Handa[3], Michiaki Mishima[3], Michito Hirakata[4], Tsutomu Takeuchi[5], Keishi Fujio[6], Kazuhiko Yamamoto[6], Hitoshi Kohsaka[7], Yoshinari Takasaki[8], Noriyuki Enomoto[9], Takafumi Suda[9], Kingo Chida[9], Shu Hisata[10], Toshihiro Nukiwa[10], Tsuneyo Mimori[1]*

1 Department of Rheumatology and Clinical Immunology, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **2** Medical & Biological Laboratories Co., Ltd., Nagoya, Japan, **3** Department of Respiratory Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan, **4** Medical Education Center/Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan, **5** Division of Rheumatology, Department of Internal Medicine, Keio University School of Medicine, Tokyo, Japan, **6** Department of Allergy and Rheumatology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan, **7** Department of Medicine and Rheumatology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan, **8** Department of Internal Medicine and Rheumatology, Juntendo University School of Medicine, Tokyo, Japan, **9** Second Division, Department of Internal Medicine, Hamamatsu University School of Medicine, Hamamatsu, Japan, **10** Department of Respiratory Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan

Abstract

Objective: Autoantibodies to aminoacyl-tRNA synthetases (ARSs) are useful in the diagnosis of idiopathic inflammatory myopathy (IIM) with interstitial pneumonia (IP). We developed an enzyme-linked immunosorbent assay (ELISA) system using a mixture of recombinant ARS antigens and tested its utility in a multicenter study. **Methods:** We prepared six recombinant ARSs: GST-Jo-1, His-PL-12, His-EJ and GST-KS expressed in *Escherichia coli*, and His-PL-7 and His-OJ expressed in Hi-5 cells. After confirming their antigenic activity, with the exception of His-OJ, we developed our ELISA system in which the five recombinant ARSs (without His-OJ) were mixed. Efficiency was confirmed using the sera from 526 Japanese patients with connective tissue disease (CTD) (IIM n = 250, systemic lupus erythematosus n = 91, systemic sclerosis n = 70, rheumatoid arthritis n = 75, Sjögren's syndrome n = 27 and other diseases n = 13), 168 with idiopathic interstitial pneumonia (IIP) and 30 healthy controls collected from eight institutes. IIPs were classified into two groups; idiopathic pulmonary fibrosis (IPF) (n = 38) and non-IPF (n = 130). Results were compared with those of RNA immunoprecipitation. **Results:** Sensitivity and specificity of the ELISA were 97.1% and 99.8%, respectively when compared with the RNA immunoprecipitation assay. Anti-ARS antibodies were detected in 30.8% of IIM, 2.5% of non-myositis CTD, and 10.7% of IIP (5.3% of IPF and 12.3% of non-IPF). Anti-ARS-positive non-IPF patients were younger and more frequently treated with glucocorticoids and/or immunosuppressants than anti-ARS-negative patients. **Conclusion:** A newly established ELISA detected anti-ARS antibodies as efficiently as RNA immunoprecipitation. This system will enable easier and wider use in the detection of anti-ARS antibodies in patients with IIM and IIP.

Citation: Nakashima R, Imura Y, Hosono Y, Seto M, Murakami A, et al. (2014) The Multicenter Study of a New Assay for Simultaneous Detection of Multiple Anti-Aminoacyl-tRNA Synthetases in Myositis and Interstitial Pneumonia. PLoS ONE 9(1): e85062. doi:10.1371/journal.pone.0085062

Editor: Masataka Kuwana, Keio University School of Medicine, Japan

Received August 29, 2013; **Accepted** November 21, 2013; **Published** January 14, 2014

Funding: This work was supported by Grants-in-Aid for Scientific Research and for Challenging Exploratory Research from the Japan Society for the Promotion of Science, and grants for intractable diseases from the Ministry of Health, Labour and Welfare in Japan.The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: Murakami M. and Seto A. are the employees of Medical and Biological Laboratory Co., Ltd. (MBL). This study was performed under collaboration between the Aauthor 's' institutes and MBL. The all authors have declared that they have no other conflicts of interest.

* E-mail: mimorit@kuhp.kyoto-u.ac.jp

☯ These authors contributed equally to this work.

Introduction

A number of autoantibodies can be detected in sera from patients with idiopathic inflammatory myopathy (IIM), some of which are specific to IIM (known as myositis-specific autoantibodies: MSAs). Detection of these autoantibodies is closely associated with IIM clinical manifestations [1,2].

Among MSAs, autoantibodies against aminoacyl-tRNA synthetases (ARSs) are the most frequently detected in adult IIM patients. To date, eight anti-ARS antibodies have been described. Anti-Jo-1 (histidyl-tRNA synthetase) [3,4] is the most common, occurring in approximately 20% of IIM patients [2,5]. Anti-PL-7 (threonyl) [6], anti-PL-12 (alanyl) [7,8], and anti-EJ (glycyl) [9] occur in ~3–4%, and anti-OJ (isoleucyl) [10] and anti-KS (asparaginyl) [11] occur in < 2% of IIM patients. Anti-tyrosyl- and anti-phenylalanyl-tRNA synthetases were also reported in one case each [12,13]. Patients with anti-ARSs show a spectrum of common clinical manifestations known as anti-synthetase syndrome (ASS), including myositis, interstitial pneumonia (IP), non-erosive arthritis, fever, Raynaud's phenomenon, and mechanic's

hands. Of note, the prevalence of IP in anti-ARS-positive patients is as high as 75–95% and IP sometimes precedes myositis [1,14,15]. Yoshifuji *et al.* reported that anti-ARS-positive patients with IP respond better to initial corticosteroid therapy but suffer from a significantly higher recurrence than anti-ARS-negative patients [1]. Therefore, anti-ARS antibodies are useful not only in diagnosing IIM but also in predicting late-onset myopathy in IP-proceeding patients and the clinical course of IP in myositis.

Currently, anti-ARS antibodies are detected using an enzyme-linked immunosorbent assay (ELISA), immunodiffusion or immunoprecipitation, but all of the antibodies are not routinely detected except for anti-Jo-1. To detect anti-ARS antibodies more readily, we established an ELISA system using a mixture of five recombinant ARS antigens: Jo-1, PL-7, PL-12, EJ, and KS. Our intention was to detect these autoantibodies simultaneously as "multiple anti-ARS antibodies". This ELISA system that we developed could be used to detect not only anti-ARS-positive myositis patients but also anti-ARS-positive idiopathic interstitial pneumonia (IIP) patients.

Materials and Methods

Patients

Serum samples were obtained from 694 Japanese adult patients with connective tissue disease (CTD) and IIP who had been followed at eight University Hospitals in Japan and 30 healthy volunteers. Patient diagnoses included IIM (n = 250), systemic lupus erythematosus (SLE) (n = 91), systemic sclerosis (SSc) (n = 70), rheumatoid arthritis (RA) (n = 75), SS (n = 27), other diseases (n = 13), and IIP (n = 168). The diagnoses of IIM, SSc, SLE, and RA were made on the basis of corresponding criteria proposed by Bohan and Peter [16] or the American College of Rheumatology [17,18,19]. IIP was defined as IP of unknown cause in which a patient did not fulfill classification criteria for any specific CTD or vasculitis, or whose lung disease was potentially caused by a drug or occupational-environmental exposure [20]. Patients with IIP were classified into two groups; an idiopathic pulmonary fibrosis (IPF) (n = 38; 12 by histological diagnosis) group and a non-IPF (n = 130; according to the typical radiographic patterns of chest high-resolution computed tomography) group.

All patients and healthy volunteers gave their written informed consent to participate in this study prior to sample collection that was performed in accordance with the Declaration of Helsinki. This study was approved by the Ethics Committee of Kyoto University Graduate School and Faculty of Medicine (Approval number: E544) and also by institutional review boards of all participating centers (Table S1).

Immunoprecipitation

The presence of anti-ARS antibodies was determined by RNA immunoprecipitation (RNA-IP) as previously described [21]. The immunoprecipitated RNA was resolved using urea-polyacrylamide gel electrophoresis and visualized using silver staining. Each anti-ARS antibody was identified according to its mobility and tRNA pattern compared with standard serum.

Construction of expression plasmids for ARS-encoding cDNAs

For the expression and purification of recombinant proteins, full-length cDNAs of PL-12, EJ, PL-7, Jo-1, KS, and OJ (GenBank accession Numbers: D32050, U09587, NM_152295, AY995220, and BC001687, respectively) were first amplified using RT-PCR with HeLa total mRNA as a template. CDNAs for PL-12 and EJ were inserted into pET30a(+) (Novagen, Madison, WI, USA) and expressed as C-terminal His-tagged proteins. CDNAs for Jo-1 and KS were subcloned into pGEX4T-1 and pGEX6P-1 (GE Healthcare UK Ltd, Buckinghamshire, England), respectively, and expressed as N-terminal GST fusion proteins. CDNAs for PL-7 and OJ were engineered with a cMyc-epitope tag and His-tag sequence at their 3′ ends, and inserted into the pFastBacDual vector for baculovirus expression (Invitrogen, Carlsbad, CA, USA). Correct construction of plasmids was confirmed using DNA sequencing.

Expression and purification of recombinant ARSs

Expression and purification of His-tagged recombinant proteins: PL-12 and EJ were expressed in *Escherichia coli* BL-21(DE3) codon plus RIL bacteria (Stratagene, La Jolla, CA, USA). Competent cells were transformed with the vectors and the cells were incubated on Luria-Bertani (LB) agar plates containing 50 μg/mL kanamycin for 15 h at 37°C. A single colony was cultured in LB liquid medium containing kanamycin at 37°C. Addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside to the medium was used to induce expression of recombinant PL-12 and EJ proteins. After a 2-h incubation, cells were harvested using centrifugation and resuspended in ice-cold phosphate buffered saline (PBS) at pH 7.5. The cells were sonicated and soluble cell lysates containing the His-tagged recombinant proteins were separated using centrifugation.

PL-7 and OJ were expressed in baculovirus-infected Hi-5 cells. Each of the expression vectors was transfected into SF-9 cells using Cellfectin (Invitrogen), and the baculovirus stock was prepared from the transfectant culture supernatant. Hi-5 cells infected with baculovirus were incubated for 72 h at 26°C and were harvested using centrifugation, and soluble cell lysates containing recombinant proteins were prepared as described above.

Soluble His-tagged recombinant ARSs were purified using immobilized metal ion affinity chromatography. Cell extracts were applied to TALON® Metal Affinity Resin columns (Clontech, Palo Alto, CA, USA), and the columns were washed with PBS containing 10 mM imidazole. Purified PL-12, EJ, PL-7, and OJ were eluted with PBS containing 50 mM imidazole.

Expression and purification of recombinant GST-ARS fusion proteins: Jo-1 and KS were also expressed in *E. coli* BL-21(DE3) codon plus RIL bacteria in the presence of ampicillin. Transformation, cultivation, induction, and extraction of soluble cell proteins were performed as described for PL-12 and EJ proteins. Soluble GST-Jo-1 and GST-KS fusion proteins were purified on Glutathione Sepharose 4B columns (GE Healthcare UK Ltd.) and eluted with Tris-HCl (pH 8.0) containing 15 mM GSH.

Immunoblotting of recombinant antigens

Purified recombinant ARS antigens were subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to a polyvinylidene difluoride (PVDF) membrane as described by Towbin *et al.* [22] with minor modifications. After blocking with 5% skimmed milk, the membrane was incubated for 60 min with serum diluted 1:100 and then incubated for 60 min with a 1:1000 dilution of goat anti-human IgG conjugated to peroxidase (Code No. 208, MBL, Nagoya, Japan). Immunoreactive bands were detected using the Western Blot Detection System WEST-one (iNtRON Biotechnology, Gyeonggi-do, Korea).

ELISA

For detection of each ARS autoantibody, purified recombinant ARSs were individually coated on 96-well microtiter plates (Maxisorp; Nunc, Rochester, NY, USA). PL-12, EJ, PL-7, and

Jo-1 were diluted in PBS to a final concentration of 2.5 μg/mL, and KS to 5.0 μg/mL. Each diluent was added at 100 μL/well and incubated overnight at 4°C. The plates were washed twice with PBS, and blocked with PBS containing 1% bovine serum albumin (BSA) and 5% sucrose overnight at 4°C. Sera from patients and normal healthy donors were diluted 1:100 in PBS containing 0.15% Tween 20 (PBS-T), 1% casein enzymatic hydrolysate, and 0.2 mg/mL *E. coli* extract, and 100 μL was applied to each well. After incubation for 60 min at room temperature (RT), the wells were washed four times with PBS-T. Goat anti-human IgG conjugated to peroxidase (Code No. 208, MBL) was diluted 1:7000 in 20 mM HEPES, 135 mM NaCl, 1% BSA, and 0.1% hydroxyphenylacetic acid (peroxidase stabilizer), and 100 μL was added to each well. After incubation for 60 min at RT, the wells were washed four times with PBS-T, and 3,3',5,5'-tetramethylbenzidine substrate was then added. After a 30-min incubation at RT, the reaction was stopped by adding 100 μL of 0.25 N sulfuric acid and absorbance was read at 450 nm (A_{450}).

For simultaneous detection of five ARS autoantibodies, purified recombinant ARSs were diluted and mixed together in PBS and coated on plates. The final concentrations of PL-12, EJ, PL-7, Jo-1, and KS were 1.25 μg/mL, 0.63 μg/mL, 1.25 μg/mL, 0.63 μg/mL, and 2.5 μg/mL, respectively. The total protein concentration of the mixture was 6.25 μg/mL. ELISA plate preparation and assays were performed as described above. Conversion from A_{450} to a unit value (U/mL) was calculated using the following formula:

$$Unit\ Value\,(U/mL) = \frac{A_{450} < Sample > - A_{450} < Blank >}{A_{450} < Positive > - A_{450} < Blank >} \times 100$$

A_{450}<Positive> is the absorbance for an anti-Jo-1-positive patient serum that corresponds to a 100 U/mL value. A_{450}<Blank> is the background absorbance of buffer that does not contain serum. A_{450}<Sample> is the absorbance of a tested serum. The cutoff point was defined at 25 U/mL based on the analysis of the receiver operating characteristic curve in this multicenter study.

Statistical analysis

Statistical analyses were performed using StatView version 5.0 software. Clinical information of anti-ARS-negative and positive non-IPF patients was compared using the two-sample t-test or the Fisher's exact test.

Results

Autoantigen preparation

We first prepared six recombinant His-tagged ARS antigens, which were all expressed in *E. coli*. Immunoblot analysis showed that four of them, Jo-1, PL-12, EJ, and KS, were identified by their corresponding autoantibodies as well as by using an ELISA, whereas PL-7 and OJ reacted weakly with their corresponding autoantibodies (data not shown). Because we hypothesized that poor antigenic activity of recombinant PL-7 and OJ was due to a lack of posttranslational modification or proper structural folding, we prepared both fusion proteins expressed in eukaryotic Hi-5 cells using the baculovirus system. We confirmed antigenic activity of the new recombinant PL-7 using an ELISA (Fig.ô 1a) but the activity was lost when examined using immunoblotting (Fig.ô 1c). Recombinant PL-7, denatured using urea or SDS, had weaker antigenic activity than non-denatured PL-7, showing that the 3-dimensional protein structure played an important role in the reaction between the threonyl-tRNA synthetase and the anti-PL-7 antibody (Fig.ô 1a). Because of this antigenic characteristic of PL-7, we decided to prepare other recombinant ARSs, without denaturing reagents, as soluble polypeptides in PBS. Because His-Jo-1 and His-KS were insoluble, they were expressed as GST-recombinant proteins. ELISA revealed that the five newly prepared ARS antigens, His-PL-12, His-EJ, GST-Jo-1, GST-KS, and His-PL-7, displayed suitable antigenic reactivity. Immunoblotting also showed that four of the five ARS antigens, except for His-PL-7, had sufficient antigenic activity (Fig.ô 1b and c).

The recombinant OJ expressed in Hi-5 cells had weak antigenic activity, as confirmed using both immunoblotting and an ELISA (data not shown), suggesting that it is difficult to prepare a recombinant OJ as a single polypeptide that retains antigenic activity.

Establishing an ELISA system for simultaneous detection of five ARS antibodies

To detect multiple ARS antibodies simultaneously, we developed an ELISA system using a mixture of the five recombinant ARSs except for OJ. We tested a variety of antigen mixtures to estimate the most appropriate ratio and concentration to use, and we found that anti-ARS-positive sera showed reactivity with all five different ARSs with the highest sensitivity and specificity occurring at antigen concentrations of 0.63, 1.25, 1.25, 0.63, and 2.5 μg/mL (6.25 μg/mL in total) for histidyl-, threonyl-, alanyl-, glycyl-, and asparaginyl-tRNA synthetases, respectively. To assess potential cross-reactivity, we compared the absorbance values (A_{450}) obtained using an ELISA on every single recombinant ARS with those obtained with the new ELISA using the ARS mixture. When tested using a single-peptide-ELISA, each of the five anti-ARS antibodies showed reactivity with only its corresponding autoantigen. Samples positive for anti-PL-7, PL-12, or KS antibodies showed higher A_{450} values with the new mixed-peptide-ELISA than with the single-peptide ELISA, whereas the samples positive for anti-Jo-1 or EJ antibodies showed no significant difference in A_{450} values obtained with the two ELISAs. Such differences in A_{450} values may be due to different peptide-coating efficiencies because the total peptide concentration was higher in the mixed-peptide-ELISA than in the single-peptide ELISA (data not shown).

Clinical significance of anti-ARS ELISA in CTD

To confirm the efficiency of this newly established ELISA, we screened a total of 694 serum samples from patients with various CTDs and IIP, and 30 healthy controls. The results were compared between the ELISA and the RNA-IP assay (Fig.ô 2). A total of 102 samples were positive for anti-ARS antibodies using the ELISA and all of them, except for one, were identified to have any anti-ARS, other than anti-OJ, using the RNA-IP assay (Tableô 1). The sensitivity and specificity of the new ELISA in the detection of anti-ARS antibodies (including anti-OJ) compared with the RNA-IP technique were 97.1% and 99.8%, respectively (Tableô 1). Anti-ARS antibodies were detected in 30.8% (77/250) of IIM and 2.5% (7/276) of other CTDs (Tableô 2). None of the healthy controls were positive (Fig.ô 2). In IIM, 30.8% (33/107) of polymyositis (PM), 35.5% (33/93) of dermatomyositis (DM), 13.0% (3/23) of amyopathic DM, and 33.3% (1/3) of overlap myositis were positive for anti-ARS antibodies (Tableô 3). Among the 95 anti-ARS-positive IIM patients, 85 (89.4%) had IP, 54 (56.8%) arthralgia/arthritis, 24 (25.3%) had mechanic's hand, 37 (38.9%) had high fever, and 31 (32.6%) had Raynaud's phenomenon, which were consistent with previous reports [15]. The prevalence of these ASS symptoms was significantly higher in the anti-ARS-positive patients than in the negative patients (data

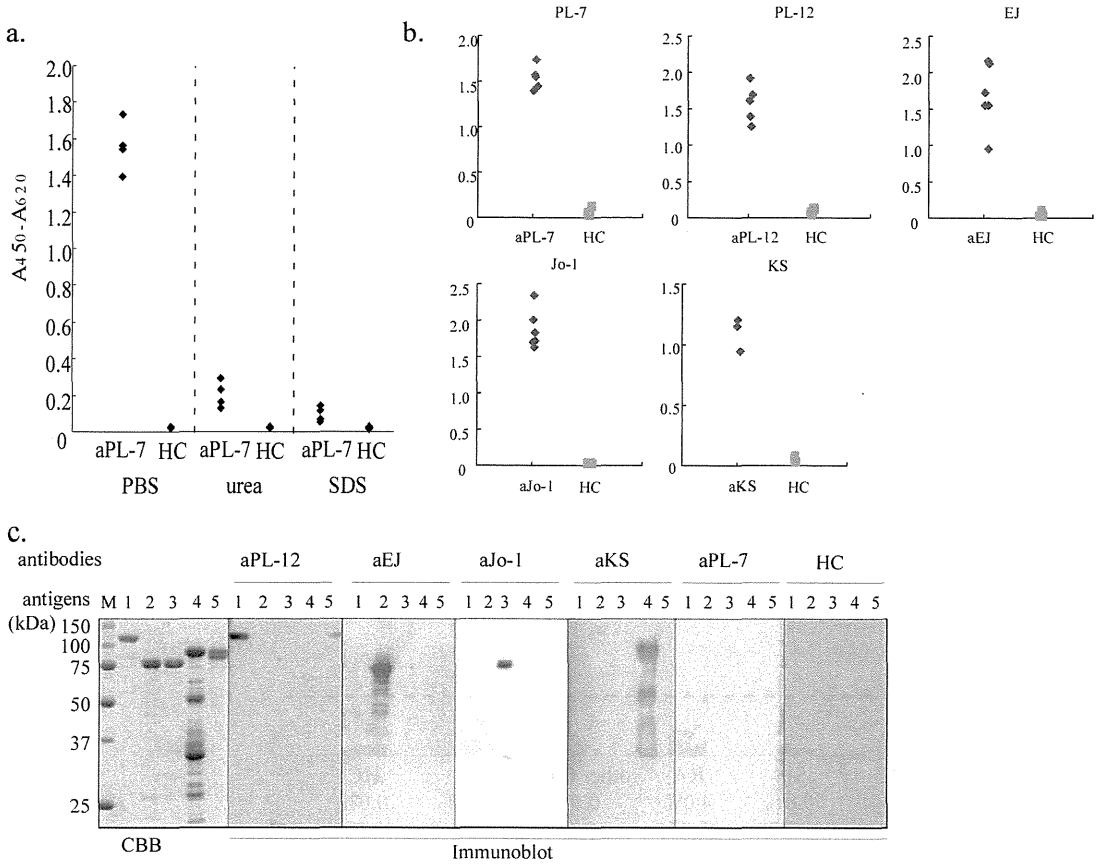


Figure 1. Antigenic activity of recombinant autoantigens a. Antigenic activity of PL-7 in various conditions. Left, purified recombinant PL-7 was eluted and diluted in PBS and coated on ELISA plates. Middle and Right, purified recombinant PL-7 was eluted in PBS and diluted in 8M urea and 2 × SDS sample buffer, respectively, and then coated onto ELISA plates. b. Five recombinant ARS antigens (His-PL-12, His-EJ, GST-Jo-1, GST-KS, and His-PL-7) were prepared as soluble polypeptides in PBS and their antigenic activity was tested in an ELISA using sera from five patients each containing corresponding autoantibodies (only GST-KS was tested using sera from three patients). Six healthy controls were used in each ELISA. c. Purified recombinant ARS antigens were electrophoresed on SDS-PAGE and transferred to a PVDF membrane followed by immunoblot analysis. CBB; Coomassie Brilliant Blue staining of gels, M; molecular weight marker, HC; healthy control, Lane 1; His-PL-12, Lane 2; His-EJ, Lane 3; GST-Jo-1, Lane 4; GST-KS and Lane 5; His-PL-7.
doi:10.1371/journal.pone.0085062.g001

not shown). There were seven anti-ARS-positive patients with other CTDs; two SSc patients were positive for anti-PL-12, two SLE patients were positive for anti-KS or anti-PL-12, and three RA patients were positive for anti-KS, anti-OJ or anti-PL-12.

Table 1 Comparison of the results between the new ELISA system and RNA-IP.

		RNA-IP		
		+	-	
anti-ARS ELISA	+	101*	1*	102
	-	0* (3)†	622* (619)†	622
	total	101* (104)†	623* (620)†	724

*The results detecting the five anti-ARS antibodies (anti-Jo-1, PL-12, EJ, KS, and PL-7) are described (sensitivity: 100%, specificity: 99.8%).
†Numbers in parenthesis are the results detecting all anti-ARS antibodies (including anti-OJ) (sensitivity: 97.1%, specificity: 99.8%).
doi:10.1371/journal.pone.0085062.t001

Clinical significance of anti-ARS ELISA in IIP

Anti-ARS antibodies were positive in 10.7% (18/168) of IIP patients. Only two patients (5.6%) with IPF were positive for anti-ARS; conversely, 16 patients (12.1%) with non-IPF were positive for anti-ARS antibodies (Tableô 2). To investigate whether the anti-ARS-positive IIP were clinically distinct from anti-ARS-negative IIP patients, we compared clinical backgrounds and treatments between anti-ARS-positive and negative non-IPF patients (Tableô 4). The anti-ARS-positive patients were significantly younger and a higher proportion was female ($p<0.01$), and they were treated more frequently with glucocorticoids (GC) or the combination of GC and immunosuppressants ($p<0.05$ and $p<0.01$, respectively).

Discussion

Among MSAs/myositis-associated autoantibodies (MAAs), anti-ARSs are the most frequently detected (28–37% [1,23,24]) in adult IIM patients, and anti-ARS-positive patients develop common characteristic symptoms known as ASS. Not only IIM but also apparent IIP patients can be positive for anti-ARS antibodies because IP often precedes myositis [1,14,20,25]. Both myopathy