© Kamla-Raj 2019 Int J Hum Genet, 19(4): 179-186 (2019)
PRINT: ISSN 0972-3757 ONLINE: ISSN 2456-6330 DOI: 10.31901/24566330.2019/19.04.734

Expression Quantitative Trait Loci for *ALMS1* and Their Influence on the Symptoms of Alström Syndrome

Minju Choi and Chaeyoung Lee

Department of Bioinformatics and Life Science, Soongsil University, Seoul 06978, Republic of Korea

KEYWORDS *ALMS1*. Alström Syndrome. Expression Quantitative Trait Locus. Genome-Wide Association Study. Single Nucleotide Variant

ABSTRACT Genome-wide expression quantitative trait loci (eQTL) were identified for *ALMS1*, which is known as a causal gene for Alström syndrome. Genetic associations of nucleotide variants were examined with expression of the gene encoding *ALMS1* in lymphoblastoid cell lines derived from 373 Europeans. The analyses revealed six cis-regulatory eQTLs for *ALMS1* ($P < 5 \times 10^{-8}$), showing a strong signal ($P < 10^{-20}$) upstream of the gene. Transcriptome-wide association analysis with these eQTLs revealed that 16 genes shared eQTLs with *ALMS1* ($P < 1.02 \times 10^{-4}$). These genes included MOCS3, CDH3, C1orf100, CXCL14, CDK4, and PCBD1, which are involved in the symptoms of Alström syndrome. This study suggested six novel eQTLs for *ALMS1* and their associations with 16 additional genes. The candidate genes whose expressions were associated with eQTL for *ALMS1* might directly or indirectly cause symptoms of Alström syndrome. Further studies are required to understand their underlying regulatory mechanisms of the eQTLs to improve our understanding of the function of *ALMS1* and other causal genes for Alström syndrome.

INTRODUCTION

Alström syndrome is known as a rare monogenic recessive ciliopathy caused by mutations in *ALMS1* on chromosome 2 (2p13.1), which is accompanied by a variety of symptoms: retinal dystrophy, hearing loss, childhood obesity and diabetes, cardiomyopathy, renal dysfunction and multiple organ failure (Hearn et al. 2002; Álvarez-Satta et al. 2015; Marshall et al. 2015). *ALMS1* localizes to centrosomes and basal bodies of primary cilia and plays a critical role in intracellular transport and cell cycle control (Hearn et al. 2005; Zulato et al. 2011; Collin et al. 2012).

Genetic association studies have been conducted to identify mutant variants of the *ALMS1* that causes Alström syndrome, especially focusing on its exons. Most variants triggering the syndrome were observed within its exons 1, 5, 8, 10, 11, 12, 16, 18, and 20 (Marshall et al. 2007; Ichihara et al. 2013; Taşdemir et al. 2013; Casey et al. 2014). In particular, exons 8, 10, and 16 reported with more than 200 mutations have been considered as hotspots (Marshall et al. 2015). The variants were associated with specific symptoms of the syndrome. For example, patients with variants in exon 16 had retinal degeneration, urologic dysfunction, or diabetes mellitus (Marshall et al. 2007). Glutamic acid repeat polymorphism in exon 1 caused early-onset myocardial infarction (Ichihara et al. 2013).

Many symptoms of Alström syndrome have been observed with relevant endocrine and metabolic characteristics (Han et al. 2018). It is, however, difficult to understand the relationship between *ALMS1* and Alström syndrome. Specifically, it is questionable whether a variety of symptoms is caused by only one gene. Some genes other than *ALMS1* might also be involved in symptoms of the syndrome, possibly by influence of or interaction with *ALMS1*. Moreover, it is difficult to understand that each variant within the same gene may produce its specific symptoms. These are critical obstacles to understand underlying molecular mechanisms of Alström syndrome.

Objectives

Regulatory nucleotide sequence variants for expressions of *ALMS1* might provide some evidence to understand the symptoms of Alström syndrome. This study aimed to identify eQTLs for *ALMS1* and to examine associations between the eQTLs and expressions of other genes.

Address for correspondence:
Chaeyoung Lee
Professor
Department of Bioinformatics and Life Science,
Soongsil University,
Seoul 06978, Republic of Korea
Telephone: +82-2-820-0455
E-mail: clee@ssu.ac.kr

MATERIAL AND METHODS

Subjects and Data

Transcriptional expression data produced by the GEUVADIS consortium were used to identify eQTLs for genes, including *ALMS1*. They were expressions of messenger RNAs in lymphoblastoid cell lines of 373 Europeans from CEPH (Utah residents with Northern and Western European ancestry from the CEPH collection), Finns, British, and Toscani in the 1000 Genomes Project (Lappalainen et al. 2013). The gene expression was quantified by summing all reads per kilobase of transcript per million (RPKM) for each gene. The corresponding genotypic data of the 373 individuals was retrieved from the 1000 Genomes Project Phase 1. They included 5,941,815 SNPs filtered by minor allele frequency <0.05.

Statistical Association Analysis

The researchers conducted a genome-wide association analysis to discover eQTLs for the *ALMS1* gene. A mixed model was employed for testing genetic association of each nucleotide variant with expression of *ALMS1* to avoid spurious association produced by population stratification (Shin and Lee 2015; Lee 2018). The analytical model was constructed under the additive model as follows:

$y=\mu 1+x\beta+g+e\varepsilon$

Where y is the vector of *ALMS1* expression, β is the fixed minor allele effect of the candidate nucleotide variant, and x is the vector for β, including elements of 0, 1, and 2 for homozygous genotype of the major allele, heterozygous genotype, and homozygous genotype of the minor allele. g is the vector of random polygenic effects with $g \sim N(0, G\sigma_g^2)$ where G is the genomic similarity matrix with elements of pairwise genomic similarity coefficients based on genotypes of nucleotide variants, and σ_g^2 is the polygenic variance component. The genomic similarity coefficient between individuals j and k was estimated as $g_{jk}=\frac{1}{n_v}\sum^{n_i}\frac{(\omega_{ij}-2p_i)(\omega_{ik}-2p_i)}{2f_i(1-p_i)}$ where n_v is the number of nucleotide variants that contribute to the genomic similarity, ω_{ij} and ω_{ik} represent the number of minor alleles for the nucleotide variant i, and p_i is the frequency of the minor allele. ε is the vector of random residuals with $\varepsilon \sim N(0,I\sigma_g^2)$ where I is the identity matrix, and σ_ε^2 is the residual variance component. The variance components for polygenic effects and residuals was estimated using average information restricted maximum likelihood, and then the fixed minor allele effect was tested using the mixed model equations with the estimates of variance components. The eQTL association analysis was conducted using Genome-wide Complex Trait Analysis (GCTA) version 1.26 (Yang et al. 2011). Multiple testing was applied with a significance threshold of $P = 5 \times 10^{-8}$. Linkage disequilibrium (LD) blocks was estimated using Haploview to determine independent association signals. Further association analysis was conducted transcriptome-wide to identify other genes which expressions were associated with the identified eQTLs. Genetic associations in the transcriptome-wide analysis were tested with a significance threshold of $P = 1.02 \times 10^{-4}$.

Functional Analysis

Functions of the SNPs associated with expression of *ALMS1* were investigated using freely available databases and software. RegulomeDB provided a score for each SNP based on how likely to affect binding and to link to expression of a gene target (http://www.regulomedb.org). HaploReg showed their regulatory functions as promoters and enhancers by regulatory chromatin states resulted from ChIP-Seq studies with H3K4me1 and H3K4me3 histone modifications (v4.1; http://compbio.mit.edu/HaploReg).

RESULTS

Genome-wide association analysis for expression of *ALMS1* revealed 489 SNPs in and around the gene on chromosome 2 ($P < 5 \times 10^{-8}$), and the association signals are presented in Figure 1. In particular, strong associations ($P < 10^{-20}$) were found with SNPs upstream of the gene. Six linkage disequilibrium blocks at the association signals were constructed in Figure 1D, indicating six eQTLs. Their representative SNPs are summarized in Table 1. The RegulomeDB showed that 27 nucleotide variants at the signals were likely to affect binding and linked to expression of a gene target in Table 2. In Table 3, the HaploReg showed

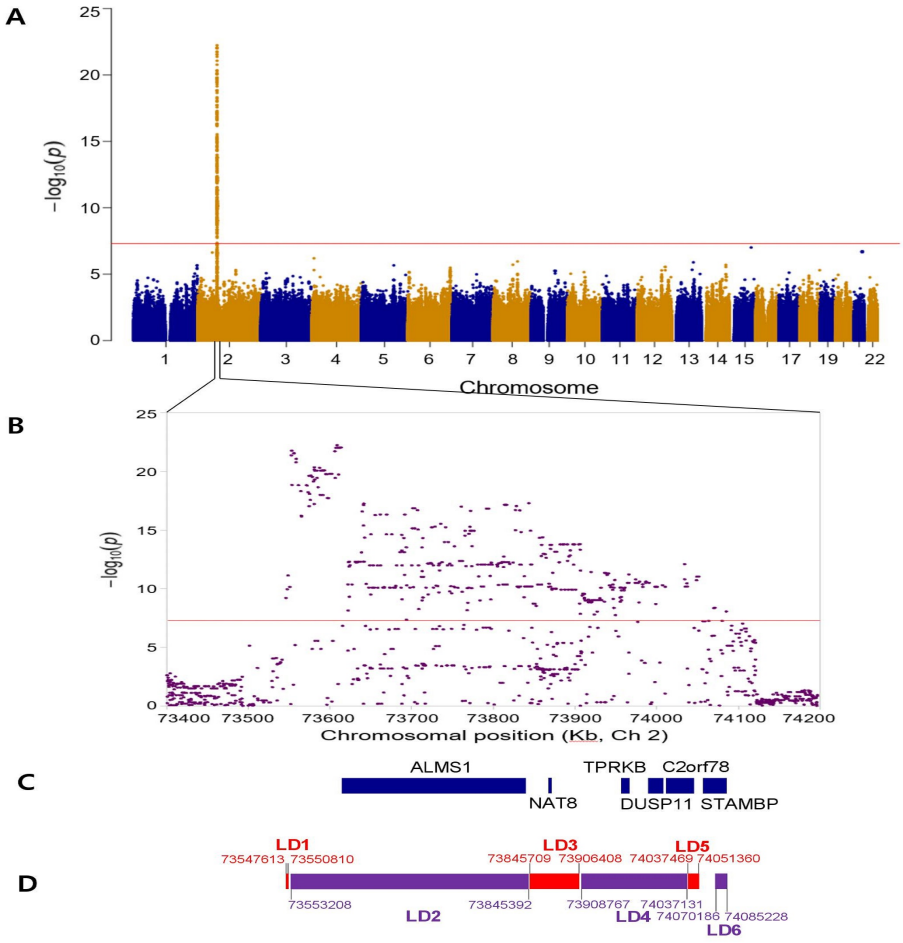

Fig. 1. Genome-wide expression quantitative trait loci (eQTLs) for *ALMS1*. A: Manhattan plot. The red horizontal line indicates the genome-wide significance threshold of $P = 5 \times 10^{-8}$. B: Regional plot for eQTLs. C: Genes located around eQTLs. D: Linkage disequilibrium blocks for eQTLs

Table 1: Expression quantitative trait loci (eQTLs) identified for *ALMS1* by genome-wide association analysis

eQTLs	*SNP*†	*Position (Ch2, bp)*	*Allele*§	*MAF*	*Beta*	*P*
1	rs6706562	73548787	G/C	0.24	0.248	7.47×10^{-12}
2	rs11693531	73608614	G/C	0.29	0.313	5.83×10^{-23}
3	rs13009035	73845709	T/C	0.36	0.239	1.14×10^{-15}
4	rs79814134	74034185	G/T	0.27	0.231	8.26×10^{-13}
5	rs732778	74049341	A/G	0.32	0.205	9.55×10^{-12}
6	rs1036113	74071151	T/C	0.25	0.199	4.31×10^{-9}

MAF: Minor allele frequency; SNP: Single nucleotide polymorphism
†representative SNPs
§minor allele/major allele

Table 3: Functions of nucleotide variants located within the expression quantitative trait loci (eQTLs) identified for *ALMS1* using HaploReg

eQTL	*SNP*	*Promoter histone mark*	*Enhancer histone mark*	*DNAse*	*Proteins bound*	*Motifs changed*
2	rs2091022	BLD				BCL
	rs11126395		BLD			Pou5f1
	rs11126396		BLD	BLD		5 altered motifs
	rs11898362		BLD			7 altered motifs
	rs7604588		LIV	BLD	4 bound proteins	Hic1,LUN-1
	rs6546824		PANC			AP-4,NF-E2,Nrf-
2	rs6716776		PANC	BRN,BRN	CTCF,SMC3, RAD21	
	rs1851546		ESRD			AP-1,EBF,Roaz
	rs1522926		ESRD			4 altered motifs
	rs6728128		ESC, BLD		8 bound proteins	DMRT3, PLZF,Z fp187
	rs7567017		BLD		GATA1	14 altered motifs
	rs1852644		IPSC			5 altered motifs
	rs6747145		FAT			4 altered motifs
	rs6748040		FAT			GATA,HDAC2
	rs7576824	ESRD				5 altered motifs
	rs12465867	14 tissues	13 tissues		FOXA1,FOXA2	Irx
	rs35885918	21 tissues	ESC, BLD, GI		FOXA1,FOXA2	
	rs11126399	24 tissues		49 tissues	20 bound proteins	7 altered motifs
	rs6706235	24 tissues		37 tissues	11 bound proteins	
	rs6546829	24 tissues		21 tissues	HAE2F1	
	rs13428235	ESRD				RFX5
	rs7580825		BLD			4 altered motifs
	rs2421546		MUS			
	rs2421548		MUS, BONE			4 altered motifs
	rs6755217		MUS, BONE			8 altered motifs
	rs12995433		MUS, BONE			GATA
	rs6755241		MUS, BONE			4 altered motifs
	rs4852927		MUS, BONE			11 altered motifs
	rs7564890		4 tissues	BRN,LNG		HDAC2,Ik-2,NF-AT
	rs6730785		4 tissues			Myb
	rs7598396		GI			SIX5
	rs6753344		GI	GI		Pax-4
	rs6740223		GI	GI,GI		Elf5
	rs6707722		GI			Foxp1, Nkx2, Nkx3
	rs10190219			GI		4 altered motifs
	rs7584575			ADRL		
	rs7598901		MUS			Smad
	rs6546836		BLD			Myb,SMC3
	rs6546837		BLD			8 altered motifs
	rs6546838		BLD			
	rs6724782		BLD			
	rs780395		ESRD			BCL,LBP-1
	rs780394		ESRD			4 altered motifs
	rs13392872	ESRD, ESC, IPSC	ESC, IPSC		Egr-1, Zfp161	
	rs73947808		5 tissues	5 tissues	CJUN,JUND	7 altered motifs
	rs56154726		BRN			Hsf
	rs9309473		4 tissues			GATA,TATA
	rs17009149			BRN		Hoxa9,NF-kappa B,Pdx1
	rs7566315		IPSC, ESC			Cdc5,Evi-1
	rs7586463	SPLN				CEBPB,Hdx

Table 3: Contd...

eQTL	*SNP*	*Promoter histone mark*	*Enhancer histone mark*	*DNAse*	*Proteins bound*	*Motifs changed*
3	rs4852939		BLD			
	rs12472502				FOXA1,FOXA2	5 altered motifs
	rs6546862		OVRY			RXRA
	rs4852950		LIV, GI	GI		
	rs4852951	LIV, ADRL	10 tissues	LIV		GR,NRSF
	rs4852952	BLD, LIV, ADRL		11 tissues	LIV,VAS	SP1,TR4 RFX5, SRF
	rs4852953	BLD, LIV, ADRL		11 tissues	LIV,VAS	SP1,TR4 NF-I, RFX5,SRF
	rs4852954	BLD, LIV, ADRL		10 tissues	7 tissues	14 bound proteins 6 altered motifs
	rs12991161		8 tissues			5 altered motifs
	rs13017182		ESRD			Bbx,Hbp1,Pax-8
	rs13027816		LIV			4 altered motifs
	rs7570683		BLD			CTCF,RXRA
	rs4341964		BLD			4 altered motifs
	rs2421582		BLD	BRN		5 altered motifs
	rs12620018		BLD			
4	rs17350125		BLD, GI			10 altered motifs
	rs4852976		BLD, GI			MIF-1
	rs11126415		BLD			Pou2f2, Pou6f1, TATA
	rs6748233			LNG		Foxa,Pax-4,STAT
	rs4852978		BLD, ADRL			Pax-4,Pbx-1, RFX5
	rs11891140	10 tissues	6 tissues	8 tissues	HEY1,CFOS,NFYA	
	rs7591112	5 tissues	8 tissues	23 tissues	6 bound proteins	9 altered motifs
	rs12713789		PANC			LRH1,Pax-5
	rs12713790		PANC			Pax-5
	rs12713791		PANC			
	rs6749841		PANC, LIV			5 altered motifs
	rs6750494		PANC, LIV			Zfp281
	rs6750515		PANC, LIV	PANC		11 altered motifs
	rs6750877		GI, LIV			HEY1,TBX5
	rs12615807		GI			11 altered motifs
	rs2421675				CTCF,POL2	6 altered motifs
	rs2421674			MUS	CTCF	5 altered motifs
	rs12713793		ESRD			AP-2,AhR:: Arnt,AhR
	rs11894953	24 tissues		53 tissues	34 bound proteins	
	rs17350188	24 tissues	ESRD, BLD	31 tissues	4 bound proteins	Hsf
	rs11126416	VAS	BLD, ADRL	6 tissues	GATA2	ERalpha-a,Myc
	rs11126417		BLD, ADRL, VAS	VAS		
	rs12713798		12 tissues	LNG,HRT,VAS		7 altered motifs
	rs12624267	HRT	14 tissues	4 tissues		Myc
	rs2421559		BLD, BRN			GR,HNF1
	rs2272051	24 tissues		42 tissues	17 bound proteins	
	rs3813228	24 tissues	GI	53 tissues	32 bound proteins	4 altered motifs
	rs13006598	ESC, LNG, LIV	4 tissues	LIV	USF1	10 altered motifs
	rs13383097		BLD			4 altered motifs
	rs7597302		BLD			6 altered motifs
5	rs2462127	BLD		BLD		
	rs732778	BLD		BLD		Cart1,Pdx1
	rs10203740			32 tissues	4 bound proteins	6 altered motifs
6	rs7594485		GI, LIV			BAF155

VAS: Vascular; HRT: Heart; ESC: ES-13 cells; IPSC: Induced pluripotent stem cells; SPLN: Spleen; LNG: Lung; LIV: Liver; BLD: Blood; PANC: Pancreas; End-stage renal disease; FAT: Fat; GI: Gastrointestinal tract; MUS: Muscle; BONE: Bone; BRN: Brain; OVRY: Ovary; ADRL: Adrenal gland

Table 2: RegulomeDB functional score for nucleotide variants located within the expression quantitative trait loci (eQTLs) identified for *AL-MS1*

eQTL	*SNP*	*Score*[a]
2	rs12478346	1f
	rs6706409	1f
	rs7576824	1f
	rs10496191	1f
	rs1083922	1f
	rs7598660	1f
	rs1052161	1f
	rs4346412	1f
3	rs4852939	1f
	rs12472502	1d
	rs4852954	1b
	rs13017182	1f
	rs2421586	1f
4	rs7606947	1f
	rs12052539	1f
	rs11891140	1f
	rs7591112	1f
	rs2421674	1f
	rs17009372	1f
	rs12713793	1d
	rs11894953	1f
	rs17350188	1f
	rs35791379	1f
	rs12713798	1f
	rs12624267	1f
	rs2272051	1f
	rs3813228	1f

[a]The smaller the number and the nearer the letter a, the stronger the function
1b: eQTL + TF binding + any motif + DNase Footprint + DNase peak
1d: eQTL + TF binding + any motif + DNase peak
1f: eQTL + TF binding /DNase peak

Table 4: Transcriptome-wide associations with the expression quantitative trait loci (eQTLs) identified for *ALMS1*[a]

eQTL	*Gene*	*Chromosome*	*Beta*	*P*
1	LRRC8D	1	0.563	1.00×10^{-04}
	ALMS1	2	0.248	7.47×10^{-12}
	TPRKB	2	-0.65	9.57×10^{-06}
	FEM1C	5	-0.249	8.52×10^{-05}
	CXCL14	5	0.137	9.11×10^{-05}
	TYW1	7	1.257	9.43×10^{-05}
	DSCC1	8	-0.353	3.93×10^{-05}
	CDK4	12	-4.498	1.02×10^{-04}
	WWWP2	16	-1.714	9.83×10^{-05}
	SNPH	20	0.038	2.39×10^{-05}
2	ALMS1	2	0.313	5.83×10^{-23}
	TPRKB	2	-0.539	5.40×10^{-05}
	TYW1	7	1.287	9.67×10^{-06}
	C8orf55	8	-0.383	6.70×10^{-05}
	CDH3	16	-0.049	7.23×10^{-05}
	MOCS3	20	-0.123	4.53×10^{-05}
3	ALMS1	2	0.239	1.14×10^{-15}
	TPRKB	2	-0.583	1.84×10^{-06}
	CDH3	16	-0.046	4.21×10^{-05}
	MOCS3	20	-0.109	8.13×10^{-05}
4	C1orf100	1	-0.02	3.10×10^{-05}
	ALMS1	2	0.231	8.26×10^{-13}
	TPRKB	2	-0.542	3.95×10^{-05}
	ADPRHL1	13	0.076	4.92×10^{-05}
5	ALMS1	2	0.205	9.55×10^{-12}
6	ALMS1	2	0.199	4.31×10^{-09}
	SERPINI1	3	0.192	1.25×10^{-05}
	PCBD1	10	-1.778	8.14×10^{-05}

[a]Associations with significance threshold of $P = 1.02 \times 10^{-4}$ are presented

that 99 nucleotide variants might be a promoter, an enhancer, or a protein binding site in a variety of tissues.

Transcriptome-wide association analysis with the identified eQTLs showed that the eQTLs also have associations ($P < 1.02 \times 10^{-4}$) with expressions of 16 additional genes in Table 4. They were all located in chromosomes other than chromosome 2 except for TPRKB gene. The rs6706562 was associated with expressions of nine additional genes, whereas no extra gene was found for the signal with rs732778. The expressions of TPRKB, MOCS3, CDH3, and TYW1 were associated with multiple eQTLs.

DISCUSSION

The current study identified six eQTLs for regulating expression of *ALMS1* known as the causal gene for Alström syndrome. They were all located near or within the gene on chromosome 2. They are suspected as cis-regulatory variants for expression of *ALMS1*.

The potential regulatory signals might influence the syndrome or its symptoms. rs10206899 in eQTL3 ($P = 5.0\times10^{-5}$) was associated with serum creatinine, a marker of kidney function, and it was also associated with chronic kidney disease, a symptom of the syndrome (Chambers et al. 2010). The metabolite X-11787 influencing chronic kidney disease and diabetes (another symptom) was associated with rs6546857 in eQTL2 ($P = 9.58\times10^{-24}$) and rs13538 in eQTL3 ($P = 1.71\times10^{-23}$) (Yu et al. 2013).

The current study showed that the six eQTLs associated with *ALMS1* expression were further associated with the expression of other genes with the same or quite similar functions as the

following symptoms of Alström syndrome. Both eQTL2 and eQTL3 were associated with the expression of MOCS3 in the current study, and its missense variant (rs7269297, $P = 3.08 \times 10^{-6}$) was identified for association with susceptibility to chronic obstructive pulmonary disease (Jackson et al. 2016). CDH3 identified for association of its expression with eQTL2 and eQTL3 was upregulated in idiopathic pulmonary fibrosis (Selman et al. 2006). The expression of C1orf100 associated with eQTL4 was associated with decreased lung function ($P < 10^{-4}$) (Forno et al. 2012). CXCL14 associated with eQTL1 enhanced insulin-dependent glucose uptake in adipocytes as a growth hormone-induced gene, and its expression caused high-fat diet-induced obesity (Takahashi et al. 2007). The CDK4 associated with eQTL1 inactivated growth-suppressive function of the retinoblastoma gene product (pRb) by phosphorylation (Kato et al. 1993). Mutation of PCBD1 associated with eQTL6 may cause early-onset diabetes in patients with mild neonatal hyperphenylalaninemia (Simaite et al. 2014). Further studies are required to understand relationship between these candidate genes and Alström syndrome. Differential expression of the genes between controls and patients with Alström syndrome can be preferentially examined.

To identify eQTLs for *ALMS1*, the current study employed an analytical model with random effects to explain polygenic effects on gene expression. As a result, the method can reduce false positive eQTLs. Bayesian inference based on the mixed model was not applied to the current analysis because of expensive computing cost. Intensive computing is required for the Bayesian method to numerically marginalize joint posterior distribution using a Markov chain Monte Carlo. Nevertheless, the Bayesian method may have a benefit in reflecting uncertainty in unknown parameters (Lee 2019).

Researchers in investigating many symptoms with the candidate genes who shared eQTLs with *ALMS1* should pay a careful attention. First, although it is hard to cause all the symptoms of Alström syndrome by only one gene, various symptoms might be explained partially by different proteins produced from *ALMS1* with tissue specific functions. Efforts on profiling of eQTLs for *ALMS1* by tissues would be useful to verify the possibility. Secondly, the current study identified eQTLs based on differential expression of *ALMS1* while many studies for Alström syndrome have dealt with the patients with lacked or nonfunctional (for example, truncated) proteins of *ALMS1*. Thus, the researchers cannot exclude a possibility that such a difference produces false positives by such an extrapolation. Thirdly, there was a limitation of the current study where eQTLs were identified for expressions of *ALMS1* in only lymphoblastoid cells. Studies with symptom-specific cells of Alström syndrome would help understand the detailed mechanisms of the syndrome.

CONCLUSION

Alström syndrome is famous for its causal gene, *ALMS1*. However, the single gene can hardly explain its various symptoms. The researchers cannot exclude a possibility that the candidate genes whose expressions were associated with eQTL for *ALMS1* might produce the symptoms of Alström syndrome, such as pulmonary dysfunction, retinal dystrophy, obesity, and diabetes. Since association with expression of *ALMS1* was more significant than that with any other genes, *ALMS1* might be involved in regulating expressions of these genes.

RECOMMENDATIONS

Further studies are needed to reveal underlying regulatory mechanisms of the eQTLs shared by *ALMS1* and the candidate genes, and this would be a key to understand the molecular mechanisms of the syndrome.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2018R1A2B6004867).

REFERENCES

Álvarez-Satta M, Castro-Sánchez S, Valverde D 2015. Alström syndrome: Current perspectives. *Appl Clin Genet,* 8: 171-179.

Casey J, McGettigan P, Brosnahan D, Curtis E, Treacy E, Ennis S et al. 2014. Atypical Alstrom syndrome with novel ALMS1 mutations precluded by current

diagnostic criteria. *Eur J Hum Genet,* 57: 55-59.
Chambers JC, Zhang W, Lord GM, van der Harst P, Lawlor DA, Sehmi JS et al. 2012. Genetic loci influencing kidney function and chronic kidney disease. *Nat Genet,* 42: 373-375.
Collin GB, Marshall JD, King BL, Milan G, Maffei P, Jagger DJ et al. 2012. The Alström syndrome protein, ALMS1, interacts with α-actinin and components of the endosome recycling pathway. *PloS One,* 7: e37925.
Forno E, Lasky-Su J, Himes B, Howrylak J, Ramsey C, Brehm J et al. 2012. Genome-wide association study of the age of onset of childhood asthma. *J Allergy Clin Immunol,* 130: 83-90.
Han JC, Reyes-Capo DP, Liu CY, Reynolds JC, Turkbey E, Turkbey IB et al. 2018. Comprehensive endocrine-metabolic evaluation of patients with Alström syndrome compared with BMI-matched controls. *ýJ Clin Endocrinol Metab,* 103: 2707-2719.
Hearn T, Renforth GL, Spalluto C, Hanley NA, Piper K, Brickwood S et al. 2002. Mutation of *ALMS1*, a large gene with a tandem repeat encoding 47 amino acids, causes Alström syndrome. *Nat Genet,* 31: 79-83.
Hearn T, Spalluto C, Phillips VJ, Renforth GL, Copin N, Hanley NA et al. 2005. Subcellular localization of ALMS1 supports involvement of centrosome and basal body dysfunction in the pathogenesis of obesity, insulin resistance, and type 2 diabetes. *Diabetes,* 54: 1581-1587.
Ichihara S, Yamamoto K, Asano H, Nakatochi M, Sukegawa M, Ichihara G et al. 2013. Identification of a glutamic acid repeat polymorphism of ALMS1 as a novel genetic risk marker for early-onset myocardial infarction by genome-wide linkage analysis. *Circ Cardiovasc Genet,* 6: 569-578.
Jackson VE, Ntalla I, Sayers I, Morris R, Whincup P, Casas JP et al. 2016. Exome-wide analysis of rare coding variation identifies novel associations with COPD and airflow limitation in MOCS3, IFIT3 and SERPINA12. *Thorax,* 71: 501-509.
Kato JY, Matsushime H, Hiebert SW, Ewen ME, Sherr CJ 1993. Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4. *Genes Dev,* 7: 331-342.
Lappalainen T, Sammeth M, Friedländer MR, 't Hoen PA, Monlong J, Rivas MA et al. 2013. Transcriptome and genome sequencing uncovers functional variation in humans. *Nature,* 501: 506-511.
Lee C 2018. Genome-wide expression quantitative trait loci analysis using mixed models. *Front Genet,* 9: 341.
Lee C 2019. Bayesian inference for mixed model-based genome-wide analysis of expression quantitative trait loci by Gibbs sampling. *Front Genet,* 10: 199.
Marshall JD, Hinman EG, Collin GB, Beck S, Cerqueira R, Maffei P et al. 2007. Spectrum of ALMS1 variants and evaluation of genotype phenotype correlations in Alström syndrome. *Hum Mutat,* 28: 1114-1123.
Marshall JD, Hinman EG, Collin GB, Beck S, Cerqueira R, Maffei P et al. 2015. Alström syndrome: Mutation spectrum of *ALMS1*. *Hum Mutat,* 36: 660-668.
Selman M, Pardo A, Barrera L, Estrada A, Watson SR, Wilson K et al. 2006. Gene expression profiles distinguish idiopathic pulmonary fibrosis from hypersensitivity pneumonitis. *Am J Respir Crit Care Med,* 173: 188-198.
Shin J, Lee C 2015. A mixed model reduces spurious genetic associations produced by population stratification in genome-wide association studies. *Genomics,* 105: 191-196.
Simaite D, Kofent J, Gong M, Rüschendorf F, Jia S, Arn P et al. 2014. Recessive mutations in PCBD1 cause a new type of early-onset diabetes. *Diabetes,* 63: 3557-3564.
Takahashi M, Takahashi Y, Takahashi K, Zolotaryov FN, Hong KS, Iida K et al. 2007. CXCL14 enhances insulin-dependent glucose uptake in adipocytes and is related to high-fat diet-induced obesity. *Biochem Biophys Res Commun,* 364: 1037-1042.
Taşdemir S, Güzel-Ozantürk A, Marshall JD, Collin GB, Ozgül RK, Narin N et al. 2013. A typical presentation and a novel mutation in ALMS1: Implications for clinical and molecular diagnostic strategies for Alström syndrome. *Clin Genet,* 83: 96-98.
Yang J, Lee SH, Goddard ME, Visscher PM 2011. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet,* 88: 76-82.
Yu B, Zheng Y, Alexander D, Manolio TA, Alonso A, Nettleton JA et al. 2013. Genome wide association study of a heart failure related metabolomic profile among African Americans in the atherosclerosis risk in communities (ARIC) study. *Genet Epidemiol,* 37: 840-845.
Zulato E, Favaretto F, Veronese C, Campanaro S, Marshall JD, Romano S et al. 2011. ALMS1-deficient fibroblasts over-express extra-cellular matrix components, display cell cycle delay and are resistant to apoptosis. *PLoS One,* 6: e19081.

Paper received for publication in July, 2019
Paper accepted for publication in October, 2019