

Replication of the Association of the 6q22.31c Locus near *GJA1* with Pulse Rate in the Korean Population

Nam Hee Kim¹, Young Jin Kim¹, Ji Hee Oh¹, Yoon Shin Cho^{1,2*}

¹Division of Structural and Functional Genomics, Center for Genome Science, Korea National Institute of Health, Cheongwon 363-951, Korea, ²Department of Biomedical Science, Hallym University, Chuncheon 200-702, Korea

Pulse rate is known to be related to diverse phenotypes, such as cardiovascular diseases, lifespan, arrhythmia, hypertension, lipids, diabetes, and menopause. We have reported two genomewide significant genetic loci responsible for the variation in pulse rate as a part of the Korea Association Resource (KARE) project, the genomewide association study (GWAS) that was conducted with 352,228 single nucleotide polymorphisms typed in 8,842 subjects in the Korean population. *GJA1* was implied as a functionally causal gene for pulse rate from the KARE study, but lacked evidence of replication. To re-evaluate the association of a locus near *GJA1* with pulse rate, we looked up this signal in another GWAS conducted in a Health Examinee-shared cohort of 3,703 samples. Not only we were able to confirm two pulse rate loci (1q32.2a near *CD46* and 6q22.13c near *LOCL644502*) identified in the KARE GWAS, we also replicated a locus (6q22.31c) near *GJA1* by the lookup in the Health Examinee GWAS. Considering that the *GJA1*-encoded protein is a major component of cardiac gap junctions, a functional study might be necessary to validate its genuine molecular biological role in the synchronized contraction of the heart.

Keywords: association, genome-wide association study, pulse rate

Introduction

Pulse rate, also called heart rate, is the number of heartbeats per unit of time, commonly used as beats per minute (BPM). Pulse rate is varied in accordance with the body's need for oxygen changes, such as during exercise or sleep. Since various health-related phenotypes are closely correlated with pulse rate [1-3], measuring pulse rate is commonly used for diagnosing and accessing medical conditions.

Pulse rate is known to be positively correlated with triglycerides and negatively correlated with high-density lipoprotein cholesterol [3]. Numerous studies have reported an association between heart rate and total and/or cardiovascular mortality [1]. Metabolic syndrome is associated with poor exercise capacity and poor heart rate recovery in patients who have coronary disease [4]. Moreover, high pulse rate is negatively correlated with lifespan in mammals, including humans [2].

Despite the importance of pulse rate as a risk factor for various diseases, however, few genes are known to be associated with pulse rate, and the underlying mechanism for pulse rate remains unraveled. We previously reported pulse rate-associated loci from the Korea Association Resource (KARE) genomewide association study (GWAS), comprising 8,842 samples in the Korean population [5]. In this study, we extended the previous study and reexamined the relevance of pulse rate-associated single-nucleotide polymorphisms (SNPs) from an additional GWAS of 3,703 Korean samples.

Methods

Subjects

In the discovery stage, the association results for pulse rate were obtained from the KARE study. The information for subjects of the KARE GWAS is described elsewhere [5]. Subjects for the replication stage were recruited from the

Received May 3, 2012; Revised May 21, 2012; Accepted May 22, 2012

*Corresponding author: Tel: +82-33-248-2111, Fax: +82-33-256-3420, E-mail: yooncho33@hallym.ac.kr

Copyright © 2012 by The Korea Genome Organization

© It is identical to the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>).

Health Examinee cohorts, called HEXA-shared cohorts. The standardized examinations that were applied in this survey included 4,302 participants aged 40 to 70 years. The characteristics of the cohorts are described in Table 1.

The measurement of pulse rate

Pulse rate is the frequency of the heart beat (BPM). Pulse rate of KARE was measured 3 times while the participant was lying down, and the average value was used for the analysis (before the first measurement, participants rested for 5 min, and 3 measurements were taken at least 30 s apart). In HEXA-shared cohorts, the first measurement of the participants, who rested for 5 min, was used.

Genotyping and quality control

The methods for genotyping and quality control for the KARE GWAS have been described elsewhere [5]. For the HEXA GWAS, genomic DNA was extracted from 4,302 participants from an urban cohort and genotyped with Affymetrix Genome-Wide Human SNP array 6.0 (Affymetrix, Inc., Santa Clara, CA, USA). From 4,302 genotyped samples, we excluded samples with a low call rate ($\geq 95\%$, $n = 443$), high heterozygosity ($n = 25$), outliers ($n = 26$), and gender inconsistencies ($n = 8$), and those obtained from individuals who had developed any kind of cancer ($n = 61$) were excluded from subsequent analyses, along with related or identical individuals whose computed average pairwise identity-by-state value was higher than that estimated from cryptic first-degree relatives of samples ($n = 33$). SNP markers with a high missing genotype rate ($> 5\%$), minor allele frequency (MAF) (< 0.01), and significant deviation from Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$) were excluded, leaving a total of 627,659 markers to be examined in 3,703 individuals.

Statistical analysis

Association analyses were performed using PLINK (<http://pngu.mgh.harvard.edu/~purcell/plink/>) [6], SAS version 9.2 (SAS institute Inc., Cary, NC, USA), and R statistics package (version 2.7.1) (<http://r-project.org/>). The SNP association for pulse rate was tested by linear regression analysis with an additive model, adjusting for age, sex, and recruitment area. The KARE GWAS and replication GWAS were combined by an inverse-variance meta-analysis method, assuming fixed effects, with Cochran's Q test used to assess between-study heterogeneity [7]. Meta-analysis calculations were performed using METAL (<http://www.sph.umich.edu/csg/abecasis/Metal>) in the R program (version 2.7.1).

Table 1. Characteristics of cohorts

Sample characteristics	KARE (mean \pm SD)	HEXA-shared (mean \pm SD)
Age (y)	52.22 \pm 8.92	53.21 \pm 8.32
Male/female	4,183/4,659	1,651/2,052
Pulse rate (BPM)	63.73 \pm 7.52	69.14 \pm 8.87

KARE, Korea Association Resource; HEXA, Health Examinee; BPM, beats per minute.

Results

The discovery GWAS for pulse rate included 352,225 SNPs profiled with Affymetrix Genome-Wide human SNP array 5.0 in 8,842 individuals from the Ansung and Ansan cohorts as described previously [5]. Characteristics of the study, including age, sex, and trait summaries, are presented in Table 1. The association analysis identified several genomic locations as potentially associated with pulse rate. Unlike the previous GWAS [5], we did not apply a threshold with a $p < 0.01$ in both Ansung and Ansan cohort in addition to a combined $p < 1 \times 10^{-5}$ for the SNP selection for the replication study so that we could include more promising SNPs for their association with pulse rate. We observed 61 significant SNPs by applying only a threshold with a combined $p < 1 \times 10^{-5}$. After excluding SNPs by cluster plot inspection and removing singletons, 3 independent SNPs (that is, with pairwise linkage disequilibrium statistics $r^2 < 0.2$ and $MAF \geq 0.05$ within a 500-kb window of the genomic region) were finally chosen to validate in the independent population - i.e., the HEXA cohort, comprising 3,703 individuals. In the replication study, all 3 SNPs were confirmed for their association with pulse rate at a $p < 0.05$ (Table 2).

The first pulse rate signal located to 1q32.2a (rs12731740, $MAF = 0.103$, $p_{overall} = 2.20 \times 10^{-8}$). Three genes, such as *CD46*, *Clorf132*, and *CD34*, appear near this locus [5]. The physiological relevance of the 3 genes in relation to pulse rate is not clear [8, 9]. Thus, fine mapping surrounding this locus by targeted resequencing is highly recommended to discover the causal variants for pulse rate. The second pulse rate SNP is rs12110693 ($MAF = 0.494$, $p_{overall} = 1.10 \times 10^{-8}$) on 6q22.31c. The only gene located near this SNP is *LOC644502*, which has no known function. The third locus associated with pulse rate is at 6q22.31c (rs2892767, $MAF = 0.261$, $p_{overall} = 2.50 \times 10^{-7}$). This locus is near *GJA1* (Gap junction protein, alpha 1, 43 kDa), which is a member of the connexin gene family. The protein encoded by *GJA1* is a component of gap junctions, which are intercellular channels that provide a route for the diffusion of low-

Table 2. Variants that associate with variation in pulse rate

SNP ID	CHR	Position	Cytoband	Candidate gene	Allele (minor/major)	MAF	KARE (n = 8,842)		HEXA-share (n = 3,703)		Combined META (n = 12,545)	
							Effect size (beta ± SEM)	p-value	Effect size (beta ± SEM)	p-value	Effect size (beta ± SEM)	p-value
rs12731740	1	206091443	1q32.2a	CD46, LOC148696, CD34	T/C	0.103	0.993 ± 0.195	3.80E-07	0.815 ± 0.342	0.0173	0.95 ± 0.1696	2.20E-08
rs2892767	6	121791626	6q22.31c	GJA1	A/C	0.261	0.598 ± 0.136	1.08E-05	0.622 ± 0.232	0.0074	0.6 ± 0.1172	2.50E-07
rs12110693	6	122199969	6q22.31c	LOC644502	T/C	0.494	0.572 ± 0.118	1.32E-06	0.632 ± 0.207	0.0023	0.59 ± 0.1026	1.10E-08

Effect sizes are shown as beta estimates from linear regression models, adjusting for age and sex. SNP, single nucleotide polymorphism; CHR, chromosome; MAF, minor allele frequency; KARE, Korea Association Resource; HEXA, Health Examinee.

molecular-weight materials from cell to cell.

Discussion

In this study, we performed a genomewide scan for pulse rate and identified 3 loci accounting for the variation in pulse rate. Two of 3 loci were previously reported and replicated once again in an independent HEXA-shared cohort. The other one (rs2892767) near *GJA1* was previously implicated for its association with pulse rate but was not chosen for the replication study, since the SNP did not meet the selection criteria in the discovery stage [5]. In the present study, SNP rs2892767 near *GJA1* was taken forward to the replication study using the independent dataset from the HEXA-shared cohort and was nicely confirmed for its association with pulse rate.

Two SNPs, rs2892767 near *GJA1* and rs12110693 near *LOC644502*, are 408 kb apart from each other and are located in a separate LD, indicating that both are independent SNPs. The independence between both SNPs was further tested by conditional analyses. We performed multiple linear regression analyses, conditioning the SNPs to each other to find a significantly stronger proxy for pulse rate. However, no dramatic decrease in statistical significance was observed, and no interaction was found between rs2892767 and rs12110693 (Table 3). These results further validate that the 2 SNPs are independent of each other.

Gap junctions are thought to have a crucial role in the synchronized contraction of the heart and in embryonic development. It is also known that gap junctions, clusters of cell-to-cell channels composed of connexins, mediate the orderly spread of electrical excitation throughout the heart [10]. The *GJA1*-encoding protein has been reported to be involved in the development and function of the heart by allowing direct cell-to-cell exchange of molecules, which mediate multiple signaling events [11]. Mutations in *GJA1* cause Mendelian inherited hypoplastic left heart syndrome [12]. Cho *et al.* [5] previously reported the same locus near *GJA1* to be responsible for pulse rate without evidence of replication. In the present study, we successfully replicated

Table 3. The effect of SNP adjustment on pulse rate

Adjustment	rs2892767 ^a		rs12110693 ^b	
	Effect size (beta ± SEM)	p-value	Effect size (beta ± SEM)	p-value
Age, sex	0.598 ± 0.136	1.00E-05	0.572 ± 0.118	1.00E-06
Age, sex, SNP	0.571 ± 0.136	2.97E-05	0.592 ± 0.118	5.85E-07

SNP, single nucleotide polymorphism.

^aAdjust with rs12110693; ^bAdjust with rs2892767.

the locus in an independent replication study.

As indicated from the functional role of *GJA1*, it is a strong candidate gene for pulse rate. Thus, further efforts, such as fine mapping or imputation based on large-scale sequencing data near rs2892767 on 6q22.31, would reveal the causal variant of this locus. In addition, functional studies for *GJA1* would be worthwhile to facilitate a full understanding of its biological role in pulse rate.

Our study was relatively underpowered to discover additional novel loci and demonstrated that large-scale genomewide meta-analysis comprising hundreds of thousands samples is required for understanding the missing heritability of pulse rate.

Acknowledgments

This research was supported by a grant from the Korea Centers for Disease Control and Prevention (4845-301, 4851-302) and an intramural grant from the Korea National Institute of Health (2011-N73005-00). Y.S.C. acknowledges support from Hallym University Research Fund 2012 (HRF-201203-008) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (2012R1A2A1A03006155).

References

1. Perret-Guillaume C, Joly L, Benetos A. Heart rate as a risk factor for cardiovascular disease. *Prog Cardiovasc Dis* 2009;52: 6-10.
2. Zhang GQ, Zhang W. Heart rate, lifespan, and mortality risk. *Ageing Res Rev* 2009;8:52-60.
3. Mounier-Véhier C, Lequeuche B, Carré A. Heart rate and lipids. *Ann Cardiol Angeiol (Paris)* 1998;47:425-428.
4. Spies C, Otte C, Kanaya A, Pipkin SS, Schiller NB, Whooley MA. Association of metabolic syndrome with exercise capacity and heart rate recovery in patients with coronary heart disease in the heart and soul study. *Am J Cardiol* 2005;95: 1175-1179.
5. Cho YS, Go MJ, Kim YJ, Heo JY, Oh JH, Ban HJ, et al. A large-scale genome-wide association study of Asian populations uncovers genetic factors influencing eight quantitative traits. *Nat Genet* 2009;41:527-534.
6. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-575.
7. Ioannidis JP, Patsopoulos NA, Evangelou E. Heterogeneity in meta-analyses of genome-wide association investigations. *PLoS One* 2007;2:e841.
8. Wang G, Liszewski MK, Chan AC, Atkinson JP. Membrane cofactor protein (MCP; CD46): isoform-specific tyrosine phosphorylation. *J Immunol* 2000;164:1839-1846.
9. Dooley DC, Oppenlander BK, Xiao M. Analysis of primitive CD34- and CD34+ hematopoietic cells from adults: gain and loss of CD34 antigen by undifferentiated cells are closely linked to proliferative status in culture. *Stem Cells* 2004;22: 556-569.
10. Wang B, Wen Q, Xie X, Liu S, Liu M, Tao Y, et al. Mutation analysis of Connexin43 gene in Chinese patients with congenital heart defects. *Int J Cardiol* 2010;145:487-489.
11. Britz-Cunningham SH, Shah MM, Zuppan CW, Fletcher WH. Mutations of the Connexin43 gap-junction gene in patients with heart malformations and defects of laterality. *N Engl J Med* 1995;332:1323-1329.
12. Eijgelsheim M, Newton-Cheh C, Sotoodehnia N, de Bakker PI, Müller M, Morrison AC, et al. Genome-wide association analysis identifies multiple loci related to resting heart rate. *Hum Mol Genet* 2010;19:3885-3894.