

Sequence Variations and Haplotypes of the *GJB2* Gene Revealed by Resequencing of 192 Chromosomes from the General Population in Korea

Hee-Jung Kim, MD · Chang-Hun Park, MD · Hee-Jin Kim, MD · Ki-O Lee, BS¹ · Hong-Hee Won, MS^{1,2}
 Moon-Hee Ko, MD³ · Hosuk Chu, MD³ · Yang-Sun Cho, MD³ · Won-Ho Chung, MD³
 Jong-Won Kim, MD · Sung Hwa Hong, MD³

Department of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul;

¹Samsung Biomedical Research Institute, Samsung Medical Center, Seoul; ²Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon; ³Department of Otorhinolaryngology and Head & Neck Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

Objectives. Hearing impairment (HI) is the most common sensory deficit in human. The Gap Junction Protein, Beta-2 (*GJB2*) gene encodes the protein connexin 26, and this gene accounts for up to half of the cases of autosomal recessive nonsyndromic HI. This study was conducted to obtain a set of sequence variations (SVs) of the *GJB2* gene among Koreans from the general population for making molecular genetic diagnoses and performing genetic counseling.

Methods. We resequenced the *GJB2* gene in 192 chromosomes from 96 adult individuals of Korean descent and who were without a history of hearing difficulty. The data of the SVs was obtained and the haplotypes were reconstructed from the data.

Results. Five SVs were observed, including a novel one (c.558G>A; p.T186T), with the allele frequencies ranging from 0.5% (1/192) to 41% (79/192). The linkage disequilibrium study and haplotype construction showed that some of the SVs are in tight linkage, resulting in a limited number of haplotypes.

Conclusion. We observed SVs of the *GJB2* gene with different allele frequencies, and a limited number of haplotypes were constructed. The data from this study can be used as reference data for *GJB2*-related hearing genetic studies, including studies on the founder effect and population genetics, and this data is particularly relevant to people of East Asian descent.

Key Words. *GJB2*, Resequencing, Sequence variations, Haplotype, Korea

• Received March 19, 2010
 Accepted after revision May 04, 2010

• Corresponding author : **Hee-Jin Kim, MD**
 Department of Laboratory Medicine & Genetics, Samsung Medical Center, Sungkyunkwan University School of Medicine, 50 Ilwon-dong, Gangnam-gu, Seoul 135-710, Korea
 Tel : +82-2-3410-2702, Fax : +82-2-3410-2719
 E-mail : heejinkim@skku.edu

• Corresponding author : **Sung Hwa Hong, MD**
 Department of Otorhinolaryngology - Head and Neck Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, 50 Ilwon-dong, Gangnam-gu, Seoul 135-710, Korea
 Tel: +82-2-3410-3574, Fax : +82-2-3410-3879
 E-mail : sunghwa3574.hong@samsung.com

INTRODUCTION

Hearing impairment (HI) is the most common form of sensory deficit worldwide, and it affects 0.05% to 0.1% of newborns (1). More than half of prelingual deafness is genetic and it most often shows an autosomal recessive inheritance pattern. Among the many causative genes and loci, the Gap Junction Protein, Beta-2 (*GJB2*) gene, which encodes the protein connexin 26, accounts for up to half of the cases of autosomal recessive nonsyndromic HI (DFNB1, MIM# 220290) (2). The *GJB2* gene is

located on the long arm of chromosome 13, band 13q11-12, and it consists of two exons, which are a single protein-coding exon preceded by a noncoding exon. When searching for a gene disease-causing mutation in the *GJB2* gene, one may find not only mutations, but also other polymorphisms. Sequence variations (SVs) include mutations, polymorphisms and variations of undetermined significance. A mutation means a disease-causing change and a polymorphism is a non-disease-causing change or change found at a frequency of 1% or higher in the population. A variation of undetermined significance is a variant of unknown clinical significance. A single nucleotide polymorphism (SNP) is a DNA sequence variation that occurs when a single nucleotide in the genome differs at an appreciable frequency in the population. More than 100 mutations of different types have been reported (<http://www.hgmd.cf.ac.uk/ac/gene.php?gene=GJB2>) and the spectrum of mutations varies greatly among the different populations with population-specific founder mutations (2-7). While nonsense or frameshift mutations leading to premature truncation are apparently mutational in nature, the disease association of those SVs that leads to single amino acid substitutions (missense or nonsynonymous) has sometimes been confusing and controversial (8, 9), and mislabeling of variants of a nonpathogenic nature would potentially lead to providing incorrect genetic diagnoses and family counseling, as well as interfering with the genotype-phenotype correlations. The segregation patterns of the genotype and phenotype within families and in *in vitro* studies have been used to support or refute a disease association. In addition, the observation of the homozygous status of a nonsynonymous SV in individuals without apparent HI supports the polymorphic nature of the SV of interest with a minimal phenotypic contribution, although the contribution is not nil. Indeed, several nonsynonymous SVs of the *GJB2* gene have been deemed SNPs (<http://www.ncbi.nlm.nih.gov/SNP/>), while some of them were reported to be mutations (10). Not only are the mutation spectrums different between European descendants and Asian populations, but also the common SNPs and their frequencies are strikingly different between the two ethnic groups according to the genotype data from four populations and that is available at the International Haplotype Mapping Project (Hap Map project; <http://www.hapmap.org/>). Haplotype is the genetic constitution of individuals with respect to one member of a pair of allelic genes or sets of genes that are closely linked and tend to be inherited together (called linkage disequilibrium [LD]).

By resequencing the *GJB2* gene in 96 adult Korean control individuals (192 control chromosomes), we investigated the SVs occurring in the *GJB2* gene and we constructed the haplotypes thereof for future studies on the founder effect in the Korean population, and potentially in other East Asian populations as well.

MATERIALS AND METHODS

Resequencing of 96 control Korean individuals

Peripheral blood samples were collected from 96 Korean control individuals (mean age, 53.2 yr; range, 42 to 71 yr) who were recruited from the Health Promotion Center for a routine checkup at our institution. According to the medical records that revealed the findings from a comprehensive evaluation of the health status, none of these study subjects had significant HI. The number of study subjects (96) was based on a previous study that demonstrated that the sample size (number of chromosomes) of 210 would be needed to detect a rare SV occurring at a 1% frequency with 80% power at a 5% significance level (11). Genomic DNA was isolated from the leukocytes using the Wizard genomic DNA purification kit according to the manufacturer's instructions (Promega, Madison, WI, USA). The entire coding exon and the flanking intronic sequences of the *GJB2* gene were amplified by polymerase chain reaction using the three pairs of primers designed by the authors (available upon request). Cycle sequencing was performed with the BigDye Terminator Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA, USA) on the ABI 3100 Genetic Analyzer (Applied Biosystems). Sequence variations were analyzed with reference to the wild type sequence using the Sequencher program (Gene Codes Corp., Ann Arbor, MI, USA). The identified variations were described according to the guidelines by the Human Genome Variation Society (<http://www.hgvs.org/mutnomen/>).

Pairwise linkage disequilibrium and haplotype analyses

Pairwise linkage disequilibrium (LD) between the SVs with minor allele frequencies (MAF) greater than 2% was analyzed by two parameters, $|D'|$ and r^2 . D is a commonly used coefficient for LD. D' is introduced to normalize D by dividing it with the theoretical maximum for the observed allele frequencies ($D'=D/D_{max}$). r^2 is another common way of representing LD (12). We used the graphical overview of linkage disequilibrium (GOLD) statistical package to obtain $|D'|$ and r^2 (13). The haplotypes from the SVs were reconstructed by the PHASE program, ver. 2.1.1. (14).

RESULTS

The percent coverage of the resequencing of the *GJB2* gene was 100% in the 192 control chromosomes. We found a total of five SVs; four nonsynonymous SVs and one novel synonymous SV (Table 1). The two most commonly occurring nonsynonymous SVs were V27I and E114G. V27I was observed in the heterozygous state in 49 individuals and in the homozygous state in 15 with a MAF at 0.411. E114G was observed in the heterozygous state in 27 individuals and in the homozygous

Table 1. Sequence variations of the GJB2 gene and their observed frequencies in 192 chromosomes from the general population in Korea as compared with the data from previous studies

Sequence variations	Allele frequency (het/hom*)	Polymorphism vs. mutation	HapMap data				Allele frequency (het/hom) of the control chromosomes from previous studies				
			CHB	JPT	CEU	YRI	Park et al. (16)	Kudo et al. (4)	Ohtsuka et al. (19)	Shi et al. (18)	Snoeckx et al. (20)
							Koreans (n=100)	Japanese (n=203)	Japanese (n=147)	Chinese (n=50)	Indonesian (n=100)
p.V271* (c.79G>A)	0.411 (49/15)	Polymorphism (rs2274084)	0.375	0.398	0	0	0.40 (56/12)	0.36	NA	0.39 (27/6)	0.035
p.V371 (c.109G>A)	0.010 (2/0)	Possibly Mutation	-	-	-	-	0	0.03	0.007	0.03 (3/0)	0.015
p.E114G* (c.341A>G)	0.193 (27/5)	Polymorphism (rs2274083)	0.289	0.102	0	0	0.20 (36/2)	0.14	NA	0.31 (19/6)	0.025
p.T123N (c.368C>A)	0.021 (4/0)	Possibly Rare polymorphism	-	-	-	-	0	0	0.007	0.01 (1/0)	0
p.T186T (c.558G>A)	0.005 (1/0)	Rare polymorphism	-	-	-	-	0	0	NA	0	0
p.I203T (c.608T>C)	0.026 (5/0)	Possibly Rare polymorphism	-	-	-	-	0.015 (3/0)	0.06	NA	0	0

*Haplotype-tagging single nucleotide polymorphisms in the HapMap dataset. *Number of individuals who were heterozygous or homozygous for the variant of interest.

CHB: Han Chinese in Beijing, China; JPT: Japanese in Tokyo, Japan; CEU: Utah residents with ancestry from northern and western Europe CEPH (Centre d'Etude du Polymorphisme Humain) families; YRI: Yoruba in Ibadan, Nigeria; NA: not available.

Table 2. Pairwise linkage disequilibrium between the four loci with minor allele frequencies greater than 0.01 (1%)

Locus		Linkage disequilibrium between two loci A & B	
A	B	ID'1*	r ² *
V271 (79G>A)	E114G (109G>A)	0.943	0.304
V271 (79G>A)	T123N (368C>A)	1	0.03
V271 (79G>A)	I203T (608T>C)	1	0.019
E114G (109G>A)	T123N (368C>A)	0.19	0.003
E114G (109G>A)	I203T (608T>C)	0.998	0.006
T123N (368C>A)	I203T (608T>C)	1	0.001

*ID'1 and r² are coefficients of linkage disequilibrium (12).

sate in 5 with a MAF at 0.193. Two individuals were shown to be heterozygous for V371 (MAF, 0.010), and four individuals were heterozygous for T123N (MAF, 0.021). Five individuals were heterozygous for I203T (MAF, 0.026). We found a novel synonymous SV, T186T, from 558G>A in a single individual. The genotype frequencies of each of the above SVs were in Hardy-Weinberg equilibrium.

The analysis of pairwise LD between the SVs with a MAF greater than 0.01 (1%) showed that the two most frequent SNPs, V271 and E114G, are tightly linked ($|D'|=0.943$, $r^2=0.304$) (Table 2). In addition, T123N and I203T are in complete linkage with V271. This LD profile is reflected by the

Table 3. Haplotypes reconstructed from the genotype data of the four loci with minor allele frequency greater than 10% and their frequencies

Haplotypes	Combination of nucleotides*	Count	Frequency
Ht 1	G-A-C-T	107	0.557
Ht 2	G-A-C-c	5	0.026
Ht 3	G-g-C-T	1	0.005
Ht 4	a-A-C-T	39	0.203
Ht 5	a-A-a-T	4	0.021
Ht 6	a-g-C-T	36	0.188
	Sum	192	1

The boldfaced haplotypes are the three common haplotypes with population frequencies greater than 0.10 (10%).

*The order of physical location is V271-E114G-T123N-I203T, and the major alleles are in capitals.

results from the haplotype analyses (Table 3). The inference of haplotypes by the PHASE algorithm reconstructed six haplotypes (Ht 1-Ht 6) from the genotype data of the four loci with MAFs greater than 1% (V271-E114G-T123N-I203T in order by the physical location). The common haplotypes occurring at frequencies greater than 10% (0.1) in the 96 control Koreans were Ht 1 (G-A-C-T; frequency, 0.557), Ht 4 (a-A-C-T; frequency, 0.203) and Ht 6 (a-g-C-T; frequency, 0.188) (bold; major alleles are in capitals). This shows that the E114G variant was occurring in tight linkage with V271, which was also supported by the

high LD values between the two loci. The other rare haplotypes also show that T123N was occurring exclusively in alleles with V27I (a)-E114E (A)-I203I (T) (Ht 5), and I203T was occurring in alleles with V27V (G)-E114E (A)-T123T (C) (Ht 2).

DISCUSSION

More than 50% of prelingual deafness has a genetic background that is most often autosomal recessive and nonsyndromic. Despite the extreme genetic heterogeneity, the *GJB2* gene has been reported to account for roughly half of autosomal recessive nonsyndromic HI in European descendants and there is a high rate of mutation carriers in the general population due to a founder mutation, 35delG (6). *GJB2*-related hereditary HI also occurs in other populations, with 235delC being a prevalent mutation in East Asians, including Koreans (7, 15-17).

The variant with the highest frequency in our data was V27I, with 15 individuals having the homozygous genotype. V27I is a common variant that exclusively occurs in Asian populations (Table 1) (4, 18). Previous reports have also described V27I in the homozygous state in control individuals, and it has the reference SNP number rs2274084 (16). Another common variant, E114G, has also been assigned the reference SNP number rs2274083. Although its disease association has been controversial, the occurrence of homozygous E114G in a sufficient number of control individuals in this study, as well as in other studies, strongly supports its polymorphic nature (4, 16, 18). According to the HapMap project and the dbSNP database at the NCBI website (<http://www.ncbi.nlm.nih.gov/projects/SNP/>), V27I and E114G are haplotype-tagging SNPs. The HapMap data shows that these two SNPs are observed only in Han Chinese and Japanese, and not in European descendants or Africans. We also noted that the MAF of E114G in Koreans from our data (0.193) falls in between those of Chinese (0.289) and Japanese (0.102), while the MAF of V27I in Koreans (0.411) is comparable to those in Chinese (0.375) and Japanese (0.398). However, a recent multi-center dataset still labeled V27I and E114G as mutations (10). Three rare nonsynonymous variants (V37I, T123N and I203T) were observed in this study only in the heterozygous state. These variants have also been rarely reported in other populations (4, 18-20), and they have been designated as missense mutations according to the mutation database. We found a novel synonymous SV in codon 186 in a single individual (heterozygous for T186T). Whether T186T is an extremely rare, silent SNP or it is a deleterious mutation resulting from aberrant splicing is unknown. There were no alleles with deleterious mutations, including 235delC, which is the most common mutation in East Asian populations, including Koreans, and in our control individuals (4, 16).

Since the phenotypic spectrum of *GJB2* mutations encompasses early-onset, prelingual deafness to late-onset deafness,

the control individuals in our study, all of whom were adults (age>41 yr), are thought to better represent a control population for the *GJB2* gene than the control population in the previous study, which involved newborns with normal audiologic screening as control subjects (16). However, it was a limitation of this study that we could not provide objective audiologic data for the study subjects; we labeled them as controls based on their medical records that showed no evidence of HI being documented according to their comprehensive health status evaluations. Last, the data from this study showing tight linkage between SVs and the limited number of haplotypes occurring in the Korean population can be used as reference data for studies on *GJB2*-related hearing genetics, including the founder effect and population genetics, and the data is particularly relevant to people of East Asian descent.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ACKNOWLEDGMENTS

This study was supported by the Samsung Medical Center Clinical Research Development Program grant, #CRS-106-56-2.

REFERENCE

1. Marazita ML, Ploughman LM, Rawlings B, Remington E, Arnos KS, Nance WE. Genetic epidemiological studies of early-onset deafness in the U.S. school-age population. *Am J Med Genet.* 1993 Jun 15;46(5):486-91.
2. Estivill X, Fortina P, Surrey S, Rabionet R, Melchionda S, D'Agruma L, et al. Connexin-26 mutations in sporadic and inherited sensorineural deafness. *Lancet.* 1998 Feb 7;351(9100):394-8.
3. Morell RJ, Kim HJ, Hood LJ, Goforth L, Friderici K, Fisher R, et al. Mutations in the connexin 26 gene (*GJB2*) among Ashkenazi Jews with nonsyndromic recessive deafness. *N Engl J Med.* 1998 Nov 19;339(21):1500-5.
4. Kudo T, Ikeda K, Kure S, Matsubara Y, Oshima T, Watanabe K, et al. Novel mutations in the connexin 26 gene (*GJB2*) responsible for childhood deafness in the Japanese population. *Am J Med Genet.* 2000 Jan 17;90(2):141-5.
5. Hwa HL, Ko TM, Hsu CJ, Huang CH, Chiang YL, Oong JL, et al. Mutation spectrum of the connexin 26 (*GJB2*) gene in Taiwanese patients with prelingual deafness. *Genet Med.* 2003 May-Jun;5(3):161-5.
6. Van Laer L, Coucke P, Mueller RF, Caethoven G, Flothmann K, Prasad SD, et al. A common founder for the 35delG *GJB2* gene mutation in connexin 26 hearing impairment. *J Med Genet.* 2001 Aug;38(8):515-8.
7. Yan D, Park HJ, Ouyang XM, Pandya A, Doi K, Erdenetungalag R, et al. Evidence of a founder effect for the 235delC mutation of *GJB2* (connexin 26) in east Asians. *Hum Genet.* 2003 Dec;114(1):44-50.
8. Griffith AJ, Chowdhry AA, Kurima K, Hood LJ, Keats B, Berlin CI,

- et al. Autosomal recessive nonsyndromic neurosensory deafness at DFNB1 not associated with the compound-heterozygous *GJB2* (connexin 26) genotype M34T/167delT. *Am J Hum Genet.* 2000 Sep; 67(3):745-9.
9. Feldmann D, Denoyelle F, Loundon N, Weil D, Garabedian EN, Couderc R, et al. Clinical evidence of the nonpathogenic nature of the M34T variant in the connexin 26 gene. *Eur J Hum Genet.* 2004 Apr; 12(4):279-84.
 10. Snoeckx RL, Huygen PL, Feldmann D, Marlin S, Denoyelle F, Waligora J, et al. *GJB2* mutations and degree of hearing loss: a multicenter study. *Am J Hum Genet.* 2005 Dec;77(6):945-57.
 11. Collins JS, Schwartz CE. Detecting polymorphisms and mutations in candidate genes. *Am J Hum Genet.* 2002 Nov;71(5):1251-2.
 12. Hudson RR. Linkage disequilibrium and recombination. In: Balding DJ, Bishop M, Cannings C, editors. *Handbook of statistical genetics.* Chichester: John Wiley & Sons; 2001. p. 309-24.
 13. Abecasis GR, Cookson WO. GOLD: graphical overview of linkage disequilibrium. *Bioinformatics.* 2000 Feb;16(2):182-3.
 14. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet.* 2001 Apr;68(4):978-89.
 15. Abe S, Usami S, Shinkawa H, Kelley PM, Kimberling WJ. Prevalent connexin 26 gene (*GJB2*) mutations in Japanese. *J Med Genet.* 2000 Jan;37(1):41-3.
 16. Park HJ, Hahn SH, Chun YM, Park K, Kim HN. Connexin26 mutations associated with nonsyndromic hearing loss. *Laryngoscope.* 2000 Sep;110(9):1535-8.
 17. Dai P, Yu F, Han B, Yuan Y, Li Q, Wang G, et al. The prevalence of the 235delC *GJB2* mutation in a Chinese deaf population. *Genet Med.* 2007 May;9(5):283-9.
 18. Shi GZ, Gong LX, Xu XH, Nie WY, Lin Q, Qi YS. *GJB2* gene mutations in newborns with non-syndromic hearing impairment in Northern China. *Hear Res.* 2004 Nov;197(1-2):19-23.
 19. Ohtsuka A, Yuge I, Kimura S, Namba A, Abe S, Van Laer L, et al. *GJB2* deafness gene shows a specific spectrum of mutations in Japan, including a frequent founder mutation. *Hum Genet.* 2003 Apr;112(4): 329-33.
 20. Snoeckx RL, Djelantik B, Van Laer L, Van de Heyning P, Van Camp G. *GJB2* (connexin 26) mutations are not a major cause of hearing loss in the Indonesian population. *Am J Med Genet A.* 2005 Jun 1;135(2): 126-9.