pISSN 1598-2998, eISSN 2005-9256

Cancer Res Treat. 2020;52(4):1251-1261 https://doi.org/10.4143/crt.2020.140

Original Article Open Access

Genome-Wide Association Study for the Identification of Novel Genetic Variants Associated with the Risk of Neuroblastoma in Korean Children

Joon Seol Bae, PhD[1]
Ji Won Lee, MD, PhD[2]
Jung Eun Yoo, MD[3]
Je-Gun Joung, PhD[4]
Keon Hee Yoo, MD, PhD[2]
Hong Hoe Koo, MD, PhD[2]
Yun-Mi Song, MD, PhD[3]
Ki Woong Sung, MD, PhD[2]

[1]*Research Institute for Future Medicine, Departments of* [2]*Pediatrics and* [3]*Family Medicine,* [4]*Samsung Genome Institute, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea*

Correspondence: Ki Woong Sung, MD, PhD
Department of Pediatrics, Samsung Medical Center, Sungkyunkwan University School of Medicine, 81 Irwon-ro, Gangnam-gu, Seoul 06351, Korea
Tel: 82-2-3410-3529
Fax: 82-2-3410-0043
E-mail: kwsped@skku.edu

Received February 24 2020
Accepted June 29, 2020
Published Online June 30, 2020

*Joon Seol Bae and Ji Won Lee contributed equally to this work.

Purpose
Neuroblastoma (NB) is the most common extracranial solid tumor found in children. To identify significant genetic factors for the risk of NB, several genetic studies was conducted mainly for Caucasians and Europeans. However, considering racial differences, there is a possibility that genetic predispositions that contribute to the development of NB are different, and genome-wide association study has not yet been conducted on Korean NB patients.

Materials and Methods
To identify the genetic variations associated with the risk of pediatric NB in Korean children, we performed a genome-wide association analysis with 296 NB patients and 1,000 unaffected controls (total n=1,296) after data cleaning and filtering as well as imputation of non-genotyped single nucleotide polymorphisms (SNPs) using IMPUTE v2.3.2.

Results
After adjusting for multiple comparisons, we found 21 statistically significant SNPs associated with the risk of NB ($p^{corr} < 0.05$) within 12 genes (*RPTN*, *MRPS18B*, *LRRC45*, *KANSL1L*, *ARHGEF40*, *IL15RA*, *L1TD1*, *ANO7*, *LAMA5*, *OR7G2*, *SALL4*, and *NEUROG2*). Interestingly, out of these, 12 markers were nonsynonymous SNPs. The SNP rs76015112 was most significantly associated with the risk of NB (p=8.1E-23, p^{corr}=2.3E-17) and was located in the *RPTN* gene. In addition, significant nonsynonymous SNPs in *ADGRE1* were found in patients with *MYCN* amplification (rs7256147, p=2.6E-05). In high-risk group, rs7256147 was observed as a significant SNP (p=5.9E-06).

Conclusion
Our findings might facilitate improved understanding of the mechanism of pediatric NB pathogenesis. However, functional evaluation and replication of these results in other populations are still needed.

Key words
Neuroblastoma, Genetic variation, Genome-wide association study, *MYCN* amplification, High risk, Korean children

Introduction

Neuroblastoma (NB) is the most common extracranial solid tumor accounting for up to 6%-10% of all childhood cancers and it is one of the leading cause of cancer mortality in children [1]. NB arises from the precursor cells of the sympathetic nervous system or the adrenal medulla. Since NB is widely heterogeneous in its clinical phenotypes and treatment outcomes, current treatment protocols are based on risk stratification of NB. High-risk group is currently defined as *MYCN*-amplified tumors at any age or metastatic tumors in patients older than 18 months according to the International Neuroblastoma Risk Group (INRG) classification system [2]. The patients in the high-risk group show poor prognosis despite modern intensive multimodal treatment. N-myc proto-oncogene protein, also known as N-Myc, is encoded by the *MYCN* gene in humans. Since *MYCN* amplification was found to be a highly predictive marker of poor outcome in the NB patients by the INRG cohort, the *MYCN* status is used for risk stratification [3]. Based on this, approximately 40% of NB patients have been classified as high-risk [2,4]. In a previous study regarding familial NB, *PHOX2A, ALK, KIF1Bβ,* and *RAS*

mutations were reported as causal mutations [5]. However, while only 1%-2% of NB cases were familial, further genetic association studies are required to identify the predisposing genetic factors. In 2008, the first genome-wide association study (GWAS) was performed for Europeans in a case of sporadic NB [6]. They reported three variants of chromosome 6p22, which have been mapped to the genes *CASC-15* and *NBAT-1*. Several candidate genes such as *LMO1*, *BARD1*, *HACE1*, and *LIN28B* have been reported by subsequent GWAS [5]. Capasso et al. [7] investigated the genetic factor for the NB patients who developed high-risk tumors and they found that the locus in 6p22 was enriched. Additionally, they found several novel risk-related single nucleotide polymorphisms (SNPs) including intronic variant in *BARD1* gene [7]. Although previous GWAS and candidate genetic studies have provided considerable information about the genetics and understanding for NB, most of these studies were performed on Caucasian and African Americans. To expand the genetic studies for other various populations such as Asians, it is necessary to explain the genetic aspects of NB.

Due to the remarkable phenotypic heterogeneity of NB, the mild to moderate effects on the risk of NB development are still unclear. In this study, we have performed GWAS using NB to discover genetic variants in Korean children. To our knowledge, this is the first GWAS study in Korean NB patients.

Materials and Methods

1. Study subjects

We screened patients who were diagnosed with NB between February 1998 and March 2017. After screening, 296 NB patients whose peripheral blood samples were already cryopreserved at Samsung Medical Center Biobank were enrolled in this study. The 77,472 exome chip genotypes of the healthy controls (n=1,000), without any history of tumor, were obtained from the National Biobank of Korea (No. 2018-019). Medical records were reviewed for obtaining detailed clinical and biological data such as the clinical features presented during diagnosis, tumor biology including *MYCN* amplification status and tumor histology by International Neuroblastoma Pathology Classification (INPC). During the study period, the patients were classified into high-risk group and non–high-risk group according to their age during diagnosis, tumor stage based on the International Neuroblastoma Staging System (INSS), and *MYCN* amplification status. In brief, stage 4 tumors in patients older than 1.5 years or with *MYCN*-amplified tumors were included in the high-risk group.

2. Genome-wide genotyping

Genomic DNA was extracted from the peripheral blood lymphocytes of the patients using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI), according to the manufacturer's protocol. Approximately 200 ng of genomic DNA was used to genotype each sample using the Illumina's Global Screening Array (GSA) BeadChip (Illumina, San Diego, CA). The samples were then processed according to the Illumina Infinium assay manual. Each sample was whole-genome amplified, fragmented, precipitated, and resuspended into an appropriate hybridization buffer. The denatured samples were then hybridized on a prepared GSA BeadChip for a minimum of 16 hours at 48°C. Following hybridization, the BeadChips were processed for the single-base extension reaction, staining, and imaging on an Illumina iScan system. The normalized bead intensity data obtained for each sample were uploaded onto the GenomeStudio software (Illumina) which converted the fluorescent intensities into SNP genotypes. The quality of sample was checked by sample call rate (> 95%). The quality of cluster for marker was measured by GenTrain scores, and then high-quality markers used this study (> 0.7).

3. Imputation and statistical analysis

The clinical variables were summarized using mean±standard deviation or median (range), as appropriate (Table 1). We performed imputation consisting of 1,296 samples. The prephasing of genotypes was conducted with SHAPEIT.v2.r837. We then imputed variants from the 1000 Genomes Project phase 3 references using IMPUTE v2.3.2. Markers with low imputation quality as call rate (< 98%), minor allele frequency (MAF < 1%), p-value of Hardy-Weinberg equilibrium (HWE; < 1E-5), duplicated markers, and ambiguous strand markers were excluded from the association analysis. In addition, low-quality samples (call rate < 95%) were applied for quality control. For genome-wide association analysis, genotype distributions were compared using logistic regression analyses with the HelixTree software (Golden Helix Inc., Bozeman, MT). To predict a protein damaging score for each nonsynonymous SNP, PolyPhen-2 program [8] was used (http://genetics.bwh.harvard.edu/pph2/index.shtml) according to the manual. Gene pathway analysis for significantly associated SNPs with the risk of NB was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) functional annotation tool (https://david.ncifcrf.gov/). In addition, biological network analysis was performed by GluGo, Cytoscape plug-in that visualizes non-redundant biological terms for large clusters of genes in a functionally grouped network (https://cytoscape.org/).

4. Ethical statement

This study was approved by the Institutional Review Board of Samsung Medical Center (IRB No. SMC 2015-06-068-006) and written informed consent was obtained from the parents or their guardians.

Table 1. Characteristics of the study subjects

Variable	Neuroblastoma	Healthy control
Total No. of subjects	296	1,000
Sex (male:female)	164:132	500:500
Age, median (min-max, yr)	2.1 (0.0-19.3)	61.4 (47-78)
***MYCN* amplification**	56 (18.9)	-
High-risk group	142 (48.0)	-
Clinical stage		
I	27 (9.1)	-
II	47 (15.9)	-
III	54 (18.2)	-
IV	160 (54.1)	-
IV-S	6 (2.0)	-
NA	2 (0.7)	-
Site of origin		
Retroperitoneum	221 (74.7)	-
Mediastinum	71 (24.0)	-
Other regions	4 (1.4)	-

Values are presented as number (%) unless otherwise indicated.

Results

1. Clinical characteristics

A total of 296 NB patients were recruited for the current study. In addition, genotypes of 1,000 normal healthy controls were obtained from the National Biobank of Korea. The average age of the NB patients was 2.1 years (range, 0.0 to 19.3 years). Among the 296 patients, 142 patients were stratified into high-risk group and *MYCN* amplification was seen in 56 patients. Table 1 shows the characteristics of the patients and healthy controls.

2. Association analysis and identification of novel susceptibility loci

A quantile-quantile plot for the association test between NB and healthy controls showed a significant deviation of measures at the tail (S1 Fig.) indicating potentially true associations between the SNPs and NB. A total of 281K markers were imputed from 535K genotypes of patients and 76K genotypes of healthy controls using strict quality control parameters (MAF > 1%, missing rate < 1%, or p for HWE < 1×10^{-5}). First, we tested the association between NB patients and healthy controls using logistic regression analysis. A total of 21 markers showed significant association with the risk of NB after adjusting for multiple comparisons (p^{corr} < 0.05) (Fig. 1A). The significant markers for the risk of NB are summarized in Table 2. The markers were located in *RPTN*, *MRPS18B*, *LRRC45*, *KANSL1L*, *ARHGEF40*, *IL15RA*, *L1TD1*, *ANO7*, *LAMA5*, *OR7G2*, *SALL4*, and *NEUROG2* genes. In addition, *PARP8*, *EPB41L3*, and *MAP4K1* genes were found to be the nearby genes for the markers such as rs7717033, rs1375128, rs10737958, rs2594708, rs2463796, rs35296988, rs3864235, rs32396, rs9964022, rs17847695, rs117910631, rs147260795, rs146801912, rs17847686, rs200216392, rs77270842, rs149013375, and rs74990833 (Table 2). Among all the markers, 12 markers were distributed in the coding region. Most of the nonsynonymous SNPs showed low MAF except of rs76015112 (MAF < 0.010). In rs76015112, the MAF of NB was lower than healthy controls (0.125 vs. 0.332) as shown in Table 2. However, the other markers in the coding region showed a high MAF in NB compared to the controls (Table 2). We analyzed the regional association of 400 kb around *RPTN* on chromosome 1q21.3 (Fig. 1B) and observed that the rs76015112 marker showed relatively robust association signal (p^{corr}=2.3E-17) (Table 2). The results of linkage disequilibrium (LD) analysis showed that the marker was unlikely to be in LD with the nearby genes (Fig. 1B).

3. Association analysis of NB subgroups

The NB patients were classified as either with *MYCN* amplification or without. Table 3 shows the significant SNPs after genome-wide association analysis between *MYCN*-amplified NB patients and the other NB patients. Interestingly, many SNPs within *ADGRE1* (synonyms: *EMR1*) gene showed significant associations (S2A Fig.). Among them, rs725614 and rs457857 had nonsynonymous SNPs as V589I (NM_001256253) and I424V (NM_001256253) (Table 3). In the regional association analysis of 200 kb around *ADGRE1* (synonyms: *EMR1*), we found that 32 SNPs were in tight LD (S2B Fig.). In addition, three significantly associated SNPs were found to be located in the intron region of *C2CD6* gene (rs143074421, rs77468686, and rs117943473).

In the following analysis, GWAS was performed with NB patients in high-risk group and non–high-risk group. The

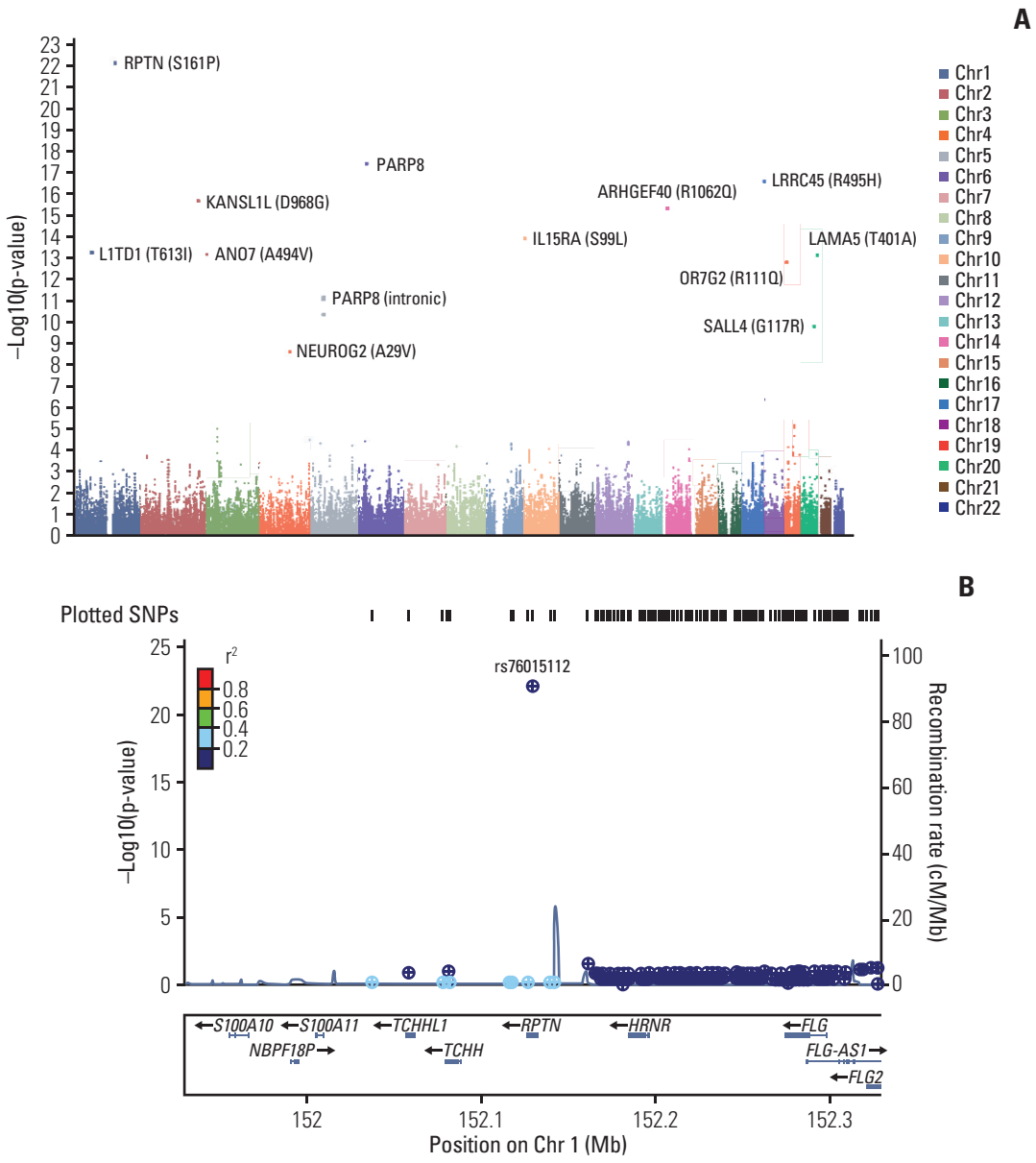

Fig. 1. (A) The p-values of genome-wide association study. The Manhattan plot shows the p-values for the risk of neuroblastoma using logistic regression analysis. x-axis represents the single nucleotide polymorphism (SNP) markers on each chromosome. The highest p-value (p=8.1E-23, p^{corr}=2.3E-17) was observed in rs76015112 on 1q21.3. (B) Regional association plots at the *RPTN*. Regional association plots including both genotyped and SNPs for the *RPTN* was generated by LocusZoom within 400 kb. The significance of association (−log10-transformed p-values) and the recombination rate are plotted. SNPs are colored to reflect pairwise linkage disequilibrium (r^2) with the most significantly associated genotyped SNP in the 1000 Genomes Project Phase 1 interim release Asian (ASN) population genotypes. The most significant genotyped SNPs are labeled and shown in purple.

top 30 significant SNPs are listed in Table 4. As a result, rs62296061 located intron region in *FGFGL1* gene was found to be highly associated with the high-risk group. Many SNPs within *ADGRE1* gene showed significant associations (S3A Fig.). In the rs725614 and rs457857 located in coding region, the MAF in high-risk group was lower than other groups (0.080 vs. 0.217 in rs725614, 0.109 vs. 0.233 in rs457857). In the regional association analysis of 200 kb around *ADGRE1*, 32 SNPs were found to be in tight LD (S3B Fig.). In Table 4, we have shown that *PON1* and *PON2* genes were located near seven markers such as rs11981667, rs17166829, rs73422040, rs11980347, rs17884252, rs17883750, and rs149643570.

Table 2. Significant SNP loci associated with neuroblastoma on the GWAS

Marker	Chromosome	Position	Transcript(s)	Gene (nearby)	In-exon	Mutation(s)	Alleles	MAF	MAF (case)	MAF (control)	OR (95% CI)	p-value	p^{corr}	Protein damaging prediction (score)
rs76015112	1	152,129,094	NM_001122965	*RPTN*	Exon	Missense_S161P	A>G	0.284	0.125	0.332	0.3 (0.2-0.4)	8.1E-23	2.3E-17	Benign (0.01)
rs148828689	6	30,593,528	NM_014046	*MRPS18B*	Exon	Missense_A244V	C>T	0.030	0.093	0.012	7.8 (4.7-12.9)	4.0E-18	5.7E-13	Benign (0.00)
rs117249618	17	79,987,501	NM_144999	*LRRC45*	Exon	Missense_R495H	G>A	0.029	0.088	0.011	7.8 (4.7-13.1)	2.6E-17	2.5E-12	Probably damaging (1.00)
rs117674897	2	210,887,734	NM_152519	*KANSL1L*	Exon	Missense_D968G	T>C	0.034	0.096	0.015	6.1 (3.9-9.7)	2.2E-16	1.5E-11	Probably damaging (0.88)
rs114591848	14	21,550,212	NM_018071	*ARHGEF40*	Exon	Missense_R1062Q	G>A	0.029	0.086	0.012	7.0 (4.2-11.6)	4.7E-16	2.6E-11	Probably damaging (0.99)
rs77226427	10	6,002,518	NR_046362, NM_001243539, NM_001256765, NM_002189, NM_172200	*IL15RA*	Exon	Silent, Missense_S96L, Missense_S218L, Missense_S132L, Missense_S99L	G>A	0.055	0.130	0.033	3.7 (2.6-5.1)	1.2E-14	5.7E-10	Benign (0.01)
rs2886644	1	62,676,284	NM_001164835, NM_019079	*L1TD1*	Exon	Missense_T613I, Missense_T613I	C>T	0.054	0.128	0.032	3.5 (2.5-4.8)	5.9E-14	2.4E-09	Benign (0.14)
rs57677160	2	242,149,010	NM_001001891	*ANO7*	Exon	Missense_A494V	C>T	0.030	0.081	0.015	6.0 (3.7-9.7)	6.8E-14	2.4E-09	Benign (0.00)
rs4925229	20	60,921,643	NM_005560	*LAMA5*	Exon	Missense_T401A	C>T	0.045	0.113	0.026	3.9 (2.7-5.6)	8.0E-14	2.5E-09	Benign (0.40)
rs62621389	19	9,213,651	NM_001005193	*OR7G2*	Exon	Missense_R111Q	C>T	0.033	0.086	0.017	5.3 (3.4-8.4)	1.6E-13	4.4E-09	Benign (0.00)
rs7717033	5	49,982,726	-	*(PARP8)*	-	-	A>T	0.316	0.429	0.282	2.0 (1.6-2.5)	7.4E-12	1.4E-07	-
rs1375128	5	50,014,674	-	*(PARP8)*	-	-	G>A	0.316	0.429	0.282	2.0 (1.6-2.5)	7.4E-12	1.4E-07	-
rs10737958	5	50,023,374	-	*(PARP8)*	-	-	A>C	0.316	0.429	0.282	2.0 (1.6-2.5)	7.4E-12	1.4E-07	-
rs2594708	5	50,024,668	-	*(PARP8)*	-	-	G>A	0.316	0.429	0.282	2.0 (1.6-2.5)	7.4E-12	1.4E-07	-
rs2463796	5	50,025,690	-	*(PARP8)*	-	-	C>G	0.316	0.429	0.282	2.0 (1.6-2.5)	7.4E-12	1.4E-07	-
rs35296988	5	50,046,753	-	*(PARP8)*	-	-	GA>G	0.315	0.429	0.282	2.0 (1.6-2.5)	8.7E-12	1.5E-07	-
rs3864235	5	49,945,474	-	*(PARP8)*	-	-	T>C	0.316	0.429	0.282	2.0 (1.6-2.5)	8.8E-12	1.5E-07	-
rs32396	5	50,106,439	-	*(PARP8)*	-	-	G>A	0.314	0.423	0.282	2.0 (1.6-2.4)	4.5E-11	7.0E-07	-
rs77538589	20	50,408,673	NM_020436	*SALL4*	Exon	Missense_G117R	C>T	0.046	0.103	0.030	3.3 (2.3-4.7)	1.4E-10	2.1E-06	Benign (0.00)
exm419398	4	113,436,546	NM_024019	*NEUROG2*	Exon	Missense_A29V	G>A	0.097	0.162	0.078	2.5 (1.8-3.3)	2.3E-09	3.2E-05	Benign (0.01)

SNP, single nucleotide polymorphism; GWAS, genome-wide association study; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.

Table 3. Logistic analysis of neuroblastoma with *MYCN* amplification

Marker	Chromosome	Position	Transcript(s)	Gene (nearby)	In-exon	Mutation(s)	Alleles	MAF	MAF (with *MYCN*)	MAF (without *MYCN*)	OR (95% CI)	p-value
rs143074421	2	202,378,651	-	*C2CD6*	-	-	A>G	0.014	0.063	0.002	32.7 (3.9-272.0)	2.3E-05
rs77468686	2	202,318,941	-	*C2CD6*	-	-	C>A	0.014	0.063	0.002	32.4 (3.9-269.6)	2.4E-05
rs117943473	2	202,256,710	-	*C2CD6*	-	-	A>G	0.014	0.063	0.002	32.3 (3.9-268.4)	2.5E-05
rs7256147	19	6,921,868	NM_001256253, NM_001974, NM_001256252, NM_001256255, NM_001256254	*ADGRE1*[a]	Exon	Missense_V589I, Missense_V589I, Missense_V537I, Missense_V412I, Missense_V448I	G>A	0.152	0.036	0.180	0.2 (0.1-0.5)	2.6E-05
rs1917448	11	103,498,150	-	LOC105369463	-	-	G>A	0.458	0.634	0.413	2.5 (1.6-3.8)	3.0E-05
rs62123078	19	6,913,435	-	*ADGRE1*	-	-	A>C	0.179	0.054	0.206	0.2 (0.1-0.5)	3.3E-05
rs62123079	19	6,913,436	-	*ADGRE1*	-	-	A>C	0.179	0.054	0.206	0.2 (0.1-0.5)	3.3E-05
rs62123080	19	6,913,437	-	*ADGRE1*	-	-	G>C	0.179	0.054	0.206	0.2 (0.1-0.5)	3.3E-05
rs62123081	19	6,913,438	-	*ADGRE1*	-	-	T>A	0.179	0.054	0.206	0.2 (0.1-0.5)	3.3E-05
rs466649	19	6,913,310	-	*ADGRE1*	-	-	T>C	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs465642	19	6,913,350	-	*ADGRE1*	-	-	T>C	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs461352	19	6,913,398	-	*ADGRE1*	-	-	C>T	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs457857	19	6,913,811	NM_001256253, NM_001974, NM_001256252, NM_001256255, NM_001256254	*ADGRE1*	Exon	Missense_I424V, Missense_I424V, Missense_I372V, Missense_I247V, Missense_I283V	G>A	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs462913	19	6,913,878	-	*ADGRE1*	-	-	T>G	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs455476	19	6,914,099	-	*ADGRE1*	-	-	T>C	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs460955	19	6,914,933	-	*ADGRE1*	-	-	G>A	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs677767	19	6,915,230	-	*ADGRE1*	-	-	A>G	0.177	0.054	0.203	0.2 (0.1-0.5)	3.5E-05
rs34406206	19	6,922,014	-	ADGRE1	-	-	C>T	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs72986353	19	6,922,093	-	ADGRE1	-	-	T>C	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs12975999	19	6,922,418	-	ADGRE1	-	-	A>C	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs11671182	19	6,922,504	-	*ADGRE1*	-	-	T>C	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs11671186	19	6,922,574	-	*ADGRE1*	-	-	T>C	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs7249799	19	6,922,815	-	*ADGRE1*	-	-	T>G	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs3890539	19	6,923,073	-	*ADGRE1*	-	-	A>G	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs35615093	19	6,923,498	-	*ADGRE1*	-	-	T>C	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs67011688	19	6,923,667	-	*ADGRE1*	-	-	A>G	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05
rs57675929	19	6,924,125	-	*ADGRE1*	-	-	G>C	0.147	0.036	0.175	0.2 (0.1-0.5)	3.8E-05

(Continued to the next page)

Table 3. Continued

Marker	Chromosome	Position	Transcript(s)	Gene (nearby)	In-exon	Mutation(s)	Alleles	MAF	MAF (with *MYCN*)	MAF (without *MYCN*)	OR (95% CI)	p-value
rs10421295	19	6,912,894	-	*ADGRE1*	-	-	A>G	0.176	0.054	0.202	0.2 (0.1-0.5)	4.1E-05
rs11669085	19	6,915,052	-	*ADGRE1*	-	-	T>C	0.175	0.054	0.201	0.2 (0.1-0.5)	4.7E-05
rs35090409	19	6,915,887	-	*ADGRE1*	-	-	G>A	0.172	0.054	0.198	0.2 (0.1-0.6)	6.4E-05

MAF, minor allele frequency; OR, odds ratio; CI, confidence interval. [a]Synonyms: EMR1.

4. Assessment of gene functional annotation and biological network analyses

The result of Gene Ontology (GO) analysis was listed in S4 Table. A total of 37 significant GO terms were identified. Among them, 10 biological pathways such as lipid transporter activity, kidney morphogenesis, regulation of macrophage derived foam cell differentiation, positive regulation of macrophage derived foam cell differentiation, sensory perception of taste, negative regulation of cell division, foam cell differentiation, macrophage derived foam cell differentiation, dicarboxylic acid catabolic process, and synaptic membrane adhesion was maintained the signals after multiple correction ($p^{corr} < 0.05$). S5 Fig. shows the result of biological network related to other biological functions. In the regulation of macrophage derived foam cell differentiation, it was closely linked to positive regulation of macrophage derived foam cell differentiation. To examine the biological function, we performed GO analysis using the DAVID. We have shown the result of GO analysis for the identified significant SNPs using the risk of NB using DAVID in S4 Table. Ten GO terms were observed have significantly corrected p-value (< 0.05). These were lipid transporter activity, kidney morphogenesis, regulation of macrophage derived foam cell differentiation, positive regulation of macrophage derived foam cell differentiation, sensory perception of taste, negative regulation of cell division, foam cell differentiation, macrophage derived foam cell differentiation, dicarboxylic acid catabolic process, and synaptic membrane adhesion.

Discussion

In the current study, we investigated novel genetic susceptibility markers for the risk of NB and its subgroups such as *MYCN* amplification group and high-risk group. We used the imputed markers from the genotypes of Illumina's GSA BeadChip through strict quality control. The Illumina GSA BeadChip, which was recently launched, contains highly optimized multi-ethnic clinical markers from well-defined databases of known diseases such as ClinVar, the Pharmacogenomics Knowledgebase (PharmGKB), and the National Human Genome Research Institute (NHGRI)-EBI database.

According to the INRG, four categories (very low-risk, low-risk, intermediate-risk, and high-risk) could be classified by seven clinical and biological factors [4]. Over the past 10 years, several genetic studies have been performed to identify somatic and germline variants affecting the onset and survival rate of NB. In the first GWAS, *CASC-15* and *NBAT-1* genes found on chromosome 6p22 were identified as the susceptible genes. Interestingly, we identified significant markers in cases with high-risk or *MYCN*-amplified NB. Subsequently, it was found that rs6939340 was most significantly associated with the risk of sporadic NB [6]. In the next GWAS, several novel risk SNPs

Table 4. Logistic analysis of neuroblastoma with high-risk grade

Marker	Chromosome	Position	Transcript(s)	Gene(s)	In-exon	Mutation(s)	Alleles	MAF	MAF (high risk)	MAF (no high risk)	OR (95% CI)	p-value
rs62296061	4	1,014,172	-	*FGFRL1*	-	-	G>A	0.039	0.078	0.007	13.6 (3.1-59.3)	2.8E-06
rs11981667	7	94,963,270	-	(PON1, PON3)	-	-	C>G	0.039	0.004	0.067	-	5.2E-06
rs17166829	7	94,966,716	-	(PON1, PON3)	-	-	G>T	0.039	0.004	0.067	-	5.2E-06
rs73422040	7	94,972,055	-	(PON1, PON3)	-	-	A>G	0.039	0.004	0.067	-	5.2E-06
rs11980347	7	94,977,637	-	(PON1, PON3)	-	-	G>A	0.039	0.004	0.067	-	5.2E-06
rs17884252	7	94,985,267	-	(PON1, PON3)	-	-	C>A	0.039	0.004	0.067	-	5.2E-06
rs17883750	7	94,995,345	-	(PON1, PON3)	-	-	A>G	0.039	0.004	0.067	-	5.2E-06
rs149643570	7	95,012,509	-	(PON1, PON3)	-	-	A>G	0.039	0.004	0.067	-	5.2E-06
rs7256147	19	6,921,868	NM_001256253, NM_001974, NM_001256252, NM_001256255, NM_001256254	*ADGRE1*	Exon	Missense_V589I, Missense_V589I, Missense_V412I, Missense_V448I	G>A	0.152	0.080	0.217	0.3 (0.2-0.6)	5.9E-06
rs34406206	19	6,922,014	-	*ADGRE1*	-	-	C>T	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs72986353	19	6,922,093	-	*ADGRE1*	-	-	T>C	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs12975999	19	6,922,418	-	*ADGRE1*	-	-	A>C	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs11671182	19	6,922,504	-	*ADGRE1*	-	-	T>C	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs11671186	19	6,922,574	-	*ADGRE1*	-	-	T>C	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs7249799	19	6,922,815	-	*ADGRE1*	-	-	T>G	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs3890539	19	6,923,073	-	*ADGRE1*	-	-	A>G	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs35615093	19	6,923,498	-	*ADGRE1*	-	-	T>C	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs67011688	19	6,923,667	-	*ADGRE1*	-	-	A>G	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs57675929	19	6,924,125	-	*ADGRE1*	-	-	G>C	0.147	0.080	0.209	0.3 (0.2-0.6)	1.4E-05
rs466649	19	6,913,310	-	*ADGRE1*	-	-	T>C	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs465642	19	6,913,350	-	*ADGRE1*	-	-	T>C	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs461352	19	6,913,398	-	*ADGRE1*	-	-	C>T	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs457857	19	6,913,811	NM_001256253, NM_001974, NM_001256252, NM_001256255, NM_001256254	*ADGRE1*	Exon	Missense_I424V, Missense_I424V, Missense_I372V, Missense_I247V, Missense_I283V	G>A	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs462913	19	6,913,878	-	*ADGRE1*	-	-	T>G	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs455476	19	6,914,099	-	*ADGRE1*	-	-	T>C	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs460955	19	6,914,933	-	*ADGRE1*	-	-	G>A	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05
rs677767	19	6,915,230	-	*ADGRE1*	-	-	A>G	0.177	0.109	0.233	0.4 (0.3-0.6)	7.4E-05

(Continued to the next page)

Table 4. Continued

Marker	Chromosome	Position	Transcript(s)	Gene(s)	In-exon	Mutation(s)	Alleles	MAF	MAF (high risk)	MAF (no high risk)	OR (95% CI)	p-value
rs10216960	8	82,939,088	-	(SNX16, LOC105375929)	-	-	T>C	0.466	0.380	0.543	0.5 (0.4-0.7)	7.4E-05
rs116876380	20	61,444,523	NM_007346	*OGFR*	EXON	Missense_S519L	C>T	0.030	0.004	0.053	0.1 (0.0-0.5)	8.3E-05
rs10421295	19	6,912,894	-	-	-	-	A>G	0.176	0.109	0.232	0.4 (0.3-0.7)	9.6E-05

MAF, minor allele frequency; OR, odds ratio; CI, confidence interval.

including *BARD1* gene on chromosome 2q35 were identified [7]. Other GWAS conducted using in familial and sporadic cases of NB have reported novel additional risk SNPs in candidate genes such as *DUSP12*, *HSD17B12*, *DDX4*, *IL31RA*, *LMO1*, *HACE1*, *LIN28B*, *SPAG16*, *NEFL*, *TP53*, *CPZ*, *MLFL*, *CDKN1B*, *KIF15*, and *MMP20* [9-19]. Among them, SNPs in *BARD1* and *LMO1* candidate genes were replicated in other cohorts due to its association by large cohort study [19,20]. In the current study, we also identified significant associations between the genes and the risk of NB (p=0.001, rs3768716, rs2070094 in *BARD1* gene; p=0.0002, rs110419 in *LMO1* gene).

However, as the previous GWAS studies were performed mostly with Caucasian and Africans, there is a serious lack of GWAS studies for Asian populations. In the current study, we identified novel genetic markers for comparing the risk between children with NB and healthy controls with no tumor record in Korean population. Using high-quality markers and strict criteria, imputed markers were used. As a result, we found 21 statistically significant markers associated with the risk of NB ($p^{corr} < 0.05$) such as *RPTN*, *MRPS18B*, *LRRC45*, *KANSL1L*, *ARHGEF40*, *IL15RA*, *L1TD1*, *ANO7*, *LAMA5*, *OR7-G2*, *SALL4*, and *NEUROG2*. Interestingly, out of these 21 significant SNPs ($p^{corr} < 0.05$), 12 SNPs were nonsynonymous (average MAF=0.064). Except rs76015112, rs77226427, and rs2886644, most nonsynonymous SNPs showed rare MAF above 5%. The most significantly associated marker was found to be rs76015112 which was located on the *RPTN* gene (p=8.1E-23, p^{corr}=2.3E-17). The *RPTN* gene encodes for repetin, an extracellular epidermal matrix protein consisting of 784 amino acids. This protein is rich in glutamine with EF-hands of the S100 type and contributes to the formation of the cornified envelope [21]. However, this gene has not yet been reported to be associated with NB. Interestingly, three nonsynonymous SNPs (rs117249618, rs117674897, and rs114591848) in *LRRC45*, *KANSL1L*, and *ARHGEF40* genes showed high protein function damaging scores (1.00, 0.88, and 0.99). Moreover, the MAF of the SNPs was higher in NB than in controls (Table 2). This implies that the SNPs may be risk factors for the onset of NB. Candidate markers within *BAL1*, *LMO1*, *MLF1*, and *HACE1* previously reported to be related with the risk of NB in Caucasian and European were not replicated in this study. The reason is presumed to be due to ethnic difference and platform difference.

NB patients were further divided into two subgroups, *MYCN*-amplified group and high-risk group, and then logistic regression analysis was performed for each of the subgroups. In these two subgroups, we found that many significant SNPs were located in the *ADGRE1* gene. Interestingly, two nonsynonymous SNPs (rs7256147 and rs457857) were identified in the *ADGRE1* gene (Tables 3 and 4). The frequencies of the two nonsynonymous SNPs in *MYCN*-amplified or high-risk tumors were lower than those in *MYCN* non-amplified or non-high-risk tumors, respectively. This

implies that the SNPs might have a protective role against the risk of NB. *ADGRE1* gene has been renamed as *EMR1* gene, which encodes for epidermal growth factor–like module containing mucin-like hormone receptor 1. *ADGRE1* gene encodes for proteins belonging to a group of hormone receptor with seven transmembrane segments [22]. Therefore, the mutated product of the *ADGRE1* gene might act as a neurotransmitter and influence the biological function of the signal transduction. When calculated using the GAS Power Calculator (http://csg.sph.umich.edu/abecasis/cats/gas_power_calculator), the expected power was 0.449. This result suggests that this study used insufficient samples, and that further studies using a large number of samples through multi-center collaboration in Asia are needed to verify the significant markers found by GWAS analysis.

To our knowledge, this is the first GWAS study in Korean NB patients. We discovered novel susceptible SNPs for the risk of NB. Of them, 12 nonsynonymous SNPs were identified. When the protein damaging prediction was performed by PolyPhen-2 algorithm, three SNPs showed high protein activity damaging scores (> 0.8). Additionally, we performed GWAS for two subgroups, *MYCN*-amplified group and the high-risk group, and identified the significantly associated SNPs in the *ADGRE1* gene. The identified variations may helpful to investigate the potential function for the onset of NB by functional assay. Our study may provide a new direction in the formulation of medication for the risk of NB in Korean population.

Electronic Supplementary Material

Supplementary materials are available at Cancer Research and Treatment website (https://www.e-crt.org).

Conflicts of Interest

Conflicts of interest relevant to this article was not reported.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (NRF-2016R1A2B1012908) and by grants from the National R&D Program for Cancer Control, Ministry of Health & Welfare, Republic of Korea (No. 1520210 and No. 1720270). This study was conducted with bioresources from National Biobank of Korea, the Centers for Disease Control and Prevention, Republic of Korea (2018-019).

References

1. Latimer E, Anderson G, Sebire NJ. Ultrastructural features of neuroblastic tumors in relation to morphological, and molecular findings; a retrospective review study. BMC Clin Pathol. 2014;14:13.
2. Cohn SL, Pearson AD, London WB, Monclair T, Ambros PF, Brodeur GM, et al. The International Neuroblastoma Risk Group (INRG) classification system: an INRG Task Force report. J Clin Oncol. 2009;27:289-97.
3. Lee JW, Son MH, Cho HW, Ma YE, Yoo KH, Sung KW, et al. Clinical significance of MYCN amplification in patients with high-risk neuroblastoma. Pediatr Blood Cancer. 2018;65:e27257.
4. Pinto NR, Applebaum MA, Volchenboum SL, Matthay KK, London WB, Ambros PF, et al. Advances in risk classification and treatment strategies for neuroblastoma. J Clin Oncol. 2015;33:3008-17.
5. Barr EK, Applebaum MA. Genetic predisposition to neuroblastoma. Children (Basel). 2018;5:119.
6. Maris JM, Mosse YP, Bradfield JP, Hou C, Monni S, Scott RH, et al. Chromosome 6p22 locus associated with clinically aggressive neuroblastoma. N Engl J Med. 2008;358:2585-93.
7. Capasso M, Devoto M, Hou C, Asgharzadeh S, Glessner JT, Attiyeh EF, et al. Common variations in BARD1 influence susceptibility to high-risk neuroblastoma. Nat Genet. 2009;41:718-23.
8. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248-9.
9. Chang X, Zhao Y, Hou C, Glessner J, McDaniel L, Diamond MA, et al. Common variants in MMP20 at 11q22.2 predispose to 11q deletion and neuroblastoma risk. Nat Commun. 2017;8:569.
10. Hungate EA, Applebaum MA, Skol AD, Vaksman Z, Diamond M, McDaniel L, et al. Evaluation of genetic predisposition for MYCN-amplified neuroblastoma. J Natl Cancer Inst. 2017;109:djx093.
11. Capasso M, McDaniel LD, Cimmino F, Cirino A, Formicola D, Russell MR, et al. The functional variant rs34330 of CDKN1B is associated with risk of neuroblastoma. J Cell Mol Med. 2017;21:3224-30.
12. McDaniel LD, Conkrite KL, Chang X, Capasso M, Vaksman Z, Oldridge DA, et al. Common variants upstream of MLF1 at 3q25 and within CPZ at 4p16 associated with neuroblastoma. PLoS Genet. 2017;13:e1006787.
13. Diskin SJ, Capasso M, Diamond M, Oldridge DA, Conkrite K, Bosse KR, et al. Rare variants in TP53 and susceptibility to neuroblastoma. J Natl Cancer Inst. 2014;106:dju047.
14. Capasso M, Diskin S, Cimmino F, Acierno G, Totaro F, Petrosino G, et al. Common genetic variants in NEFL influence gene expression and neuroblastoma risk. Cancer Res. 2014;74:6913-24.
15. Latorre V, Diskin SJ, Diamond MA, Zhang H, Hakonarson H, Maris JM, et al. Replication of neuroblastoma SNP association at the BARD1 locus in African-Americans. Cancer Epidemiol Biomarkers Prev. 2012;21:658-63.
16. Diskin SJ, Capasso M, Schnepp RW, Cole KA, Attiyeh EF, Hou C, et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. Nat Gen-

et. 2012;44:1126-30.
17. Nguyen le B, Diskin SJ, Capasso M, Wang K, Diamond MA, Glessner J, et al. Phenotype restricted genome-wide association study using a gene-centric approach identifies three low-risk neuroblastoma susceptibility loci. PLoS Genet. 2011; 7:e1002026.
18. Matthews JM, Lester K, Joseph S, Curtis DJ. LIM-domain-only proteins in cancer. Nat Rev Cancer. 2013;13:111-22.
19. Wang K, Diskin SJ, Zhang H, Attiyeh EF, Winter C, Hou C, et al. Integrative genomics identifies LMO1 as a neuroblastoma oncogene. Nature. 2011;469:216-20.
20. Capasso M, Diskin SJ, Totaro F, Longo L, De Mariano M, Russo R, et al. Replication of GWAS-identified neuroblastoma risk loci strengthens the role of BARD1 and affirms the cumulative effect of genetic variations on disease susceptibility. Carcinogenesis. 2013;34:605-11.
21. Huber M, Siegenthaler G, Mirancea N, Marenholz I, Nizetic D, Breitkreutz D, et al. Isolation and characterization of human repetin, a member of the fused gene family of the epidermal differentiation complex. J Invest Dermatol. 2005;124:998-1007.
22. Baud V, Chissoe SL, Viegas-Pequignot E, Diriong S, N'Guyen VC, Roe BA, et al. EMR1, an unusual member in the family of hormone receptors with seven transmembrane segments. Genomics. 1995;26:334-44.