Whole exome sequencing in Korean patients with retinitis pigmentosa identified mutations in *EYS*

SeungHee Jung[1], Siyeon Kim[1], YoungChan Park[1,2], JongYoung Lee[2]*, InSong Koh[1]*

[1]Department of Biomedical Informatics, Hanyang University, Seoul, Korea

[2]One Omics, Seoul, Korea

Abstract

Retinitis pigmentosa(RP) is a rare hereditary retinal disease characterized by progressive degeneration of photoreceptors. The worldwide prevalence of RP is about 1 in 4000. The clinical features of RP are usually confined to the eyes, but some patients with RP have a syndrome affecting non-ocular organs. RP is genetically heterogeneous and can be inherited in autosomal dominant, autosomal recessive, or X-linked modes. More than 60 genes are identified as causes of RP (RetNet-Retinal Information Network). The eyes shut homolog (*EYS*) gene is located on chromosome 6q12 and is one of the largest genes expressed in the retina. The *EYS* protein is important for maintaining the normal morphology of photoreceptor cells. A high prevalence of *EYS*-related RP was reported in Asian countries (Korea, Japan, and China), but the relationship between *EYS* mutations and clinical phenotypes is still unclear. To identify RP related *EYS* variants in a Korean cohort, we performed a whole exome sequencing analysis only in non-syndromic RP patients. *EYS* mutations account for 12% of the total RP-related mutations in this study. The 4 *EYS* variants (c.2259+1G>A, c.4957dupA, c.6557G>A, c.8805C>A) are likely to be pathogenic and compound heterozygotes according to ACMG guidelines.

Materials and Methods

1. Data collection

157 Korean RP patient's genomics materials were extracted from blood. Sequencing libraries were prepared along with capture using xGen Exome Research Panel V2 (Integrated DNA Technologies). The quantitation of libraries was measured by 2100 Bioanalyzer (Agilent).The exome libraries were paired-end sequenced on the Illumina Novaseq 6000 sequencing system aiming throughput at least 100x coverage.

2. Bioinformatics analysis

bcl2fastq2 from the CASAVA software suite was used for conversion from BCL data to fastQ. The Sanger/phred33/illumine 1.8+ was used for the quality scale. To detect sequencing errors performed using FastQC. The remaining adapters were trimmed using cutadapt. Preprocessed sequencing reads were aligned against human reference genome GRCh37 using Burrow Wheeler Aligner(BWA) mem algorithms. To correct duplication that occurs from the PCR stage, mapped reads were processed using GATK's MarkDuplicate. Then, base quality score recalibration was conducted to avoid exaggerated quality scores for each base. GATK HaplotypeCaller was used to discover variants.

To prioritize confident variants, we filtered variants with QUAL < 30. Annotation was performed using variants annotation software SnpEff and SnpSift. Although we extracted all the variants in the RP cohort, we investigated more on variants in *EYS* gene region.

Results

1. Causative genes.

We investigated variants in genes associated with RP in 157 Korean RP patients. As in the previous study that discovered mutations in Korean retinal disease patients (Ma DJ,2021), *EYS* variants were the most frequently identified, accounting for 12%.(Fig.1). It was followed by *USH2A* with 8%, *CNGB1* with 5%, *PCARE* with 4%, *IMPG*, *ROM1*, *CDH23, MERTK, PRPF8*, and *RP1* with about 3% each. In the two large cohort studies of other Asian countries, Japan and China, it was reported that *EYS* was the most common mutation gene in Japanese and the third most common mutation gene in Chinese (Numa S,2020; Gao FJ,2019).

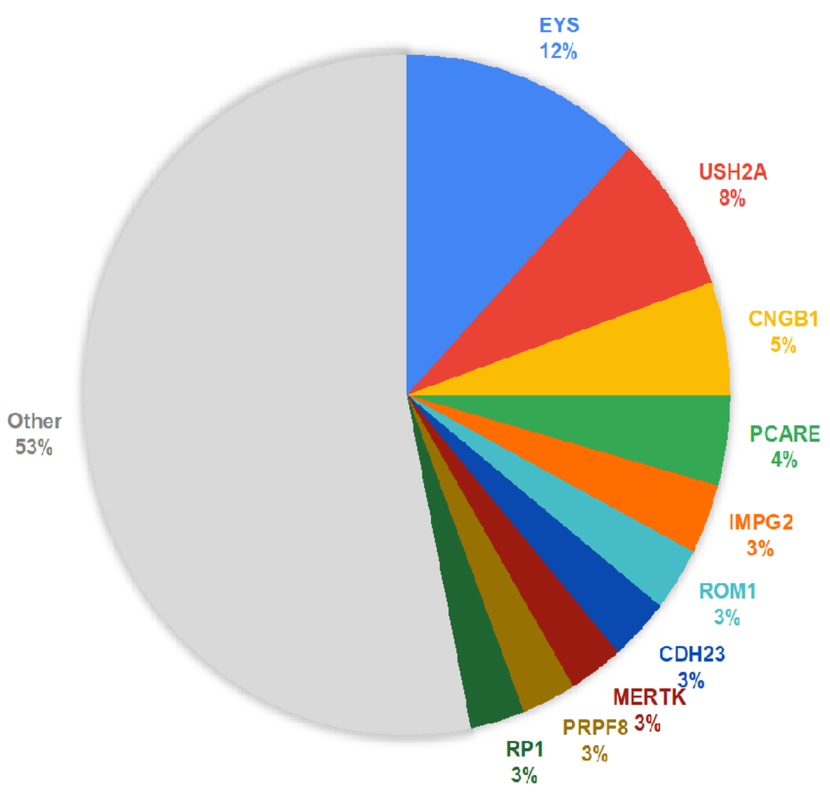

Figure 1. Mutation proportions of RP-related Genes of 157 Korean RP patients. Genes with less than 2% were expressed as "Other".

2. *EYS* variants investigation

In this study, we focused on the *EYS* gene, which is the largest gene in the retina as well as the main gene frequently identified in Asian RP patients. In *EYS* variants of Korean RP patients, we Identified 4 pathogenic variants (c.4957dupA, c.2259+1G>A, c.8805C> A, c.6557G>A) in 14 patients. Three patients with pathogenic variants were considered as compound heterozygotes for *EYS* variants (Table 1).

Patients	Nucleotide exchange	Amino acid exchange	Type
SM031	c.4957dupA		Frameshift
	c.2259+1G>A		Splice site
SM130	c.8805C>A	p.Tyr2935Ter	missense
	c.6557G>A	p.Gly2186Glu	missense
SM076	c.8805C>A	p.Gly2186Glu	missense
	c.2259+1G>A		

Table 1. *EYS* gene variants in three compound heterozyg--ous patients

Discussion

Identification of the causative mutations is the starting point of RP treatment. In this study, we used the WES strategy to detect RP-associated variants in *EYS* in 157 Korean RPbpatients. In the *EYS* gene, four pathogenic variants according to the ACMG criteria were found in 14 patients (about 8.9%). In addition, we found in 3 patients a mutation likely to be compound heterozygous.

Acknowledgement

This research was supported by NRF-2017M3C9A6047623.

HANYANG UNIVERSITY