

A novel geneX mutation R69H identified in a Chinese family with autosomal dominant infantile cataract

Hongfu Zhang

Abstract

Congenital cataract is a significant cause of hereditary visual impairment in childhood. To identify a gene responsible for autosomal dominant congenital cataract (ADCC) in a large four-generation, non-consanguineous Chinese family, we performed whole-exome sequencing on samples from three patients and one normal person in family. Whole-exome sequencing analysis identified a mutation in *gene X*, R69H, that co-segregated with the disease in the family but did not exist in >1126 controls. The variant's frequency is very low in public databases including GnomAD_exome-All, ExAC-ALL but not found in East Asian population. Bioinformatics analysis indicate R69H is Deleterious and very conservation. LOD scores max is nearly significance with 2.81, calculated by parametric linkage analysis with the autosomal dominant model. These data suggest that rare *geneX* variant are associated with ADCC. *geneX*, a RNA-binding protein that acts as a pre-mRNA splicing factor is not previously implicated in ADCC. Lens fiber cell defective differentiation and cataract formation at 72 hpf, with Morpholino micro-injection into zebrafish single cell embryo to knock down gene Xa, gene X homologous gene. This study identify a new gene for infantile cataract, which may provide insights into the pathogenic mechanisms of cataracts.

Result

Fig. 1. Pedigree structure of a Chinese family affected with ADCC

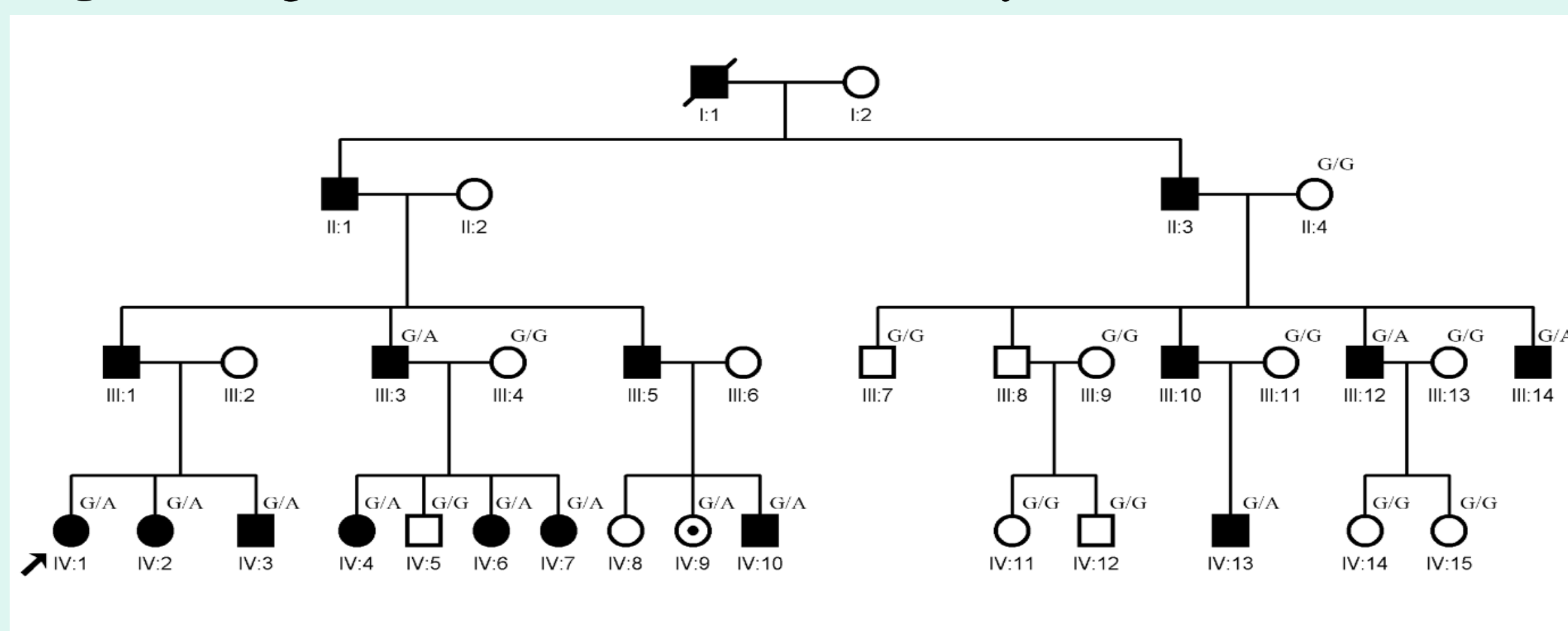


Fig. 2. Identification of a *gene X* mutation co-segregating with ADCC

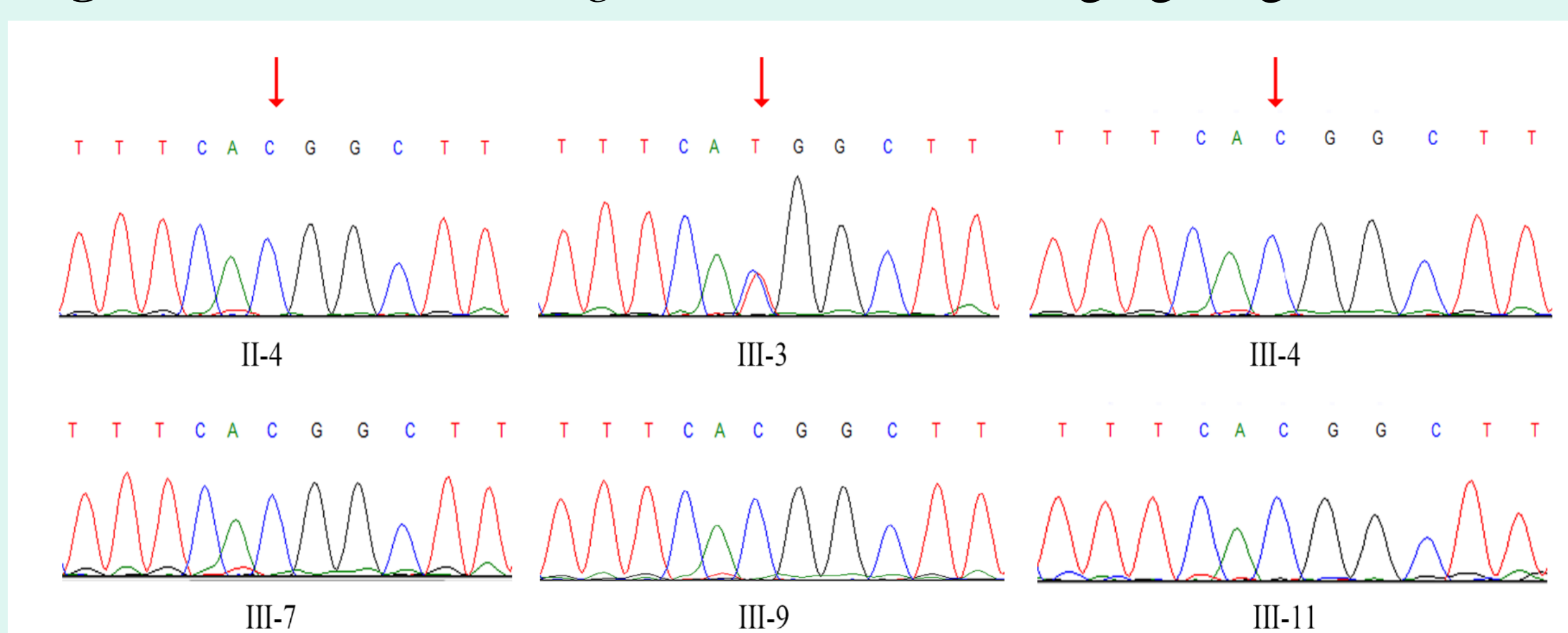
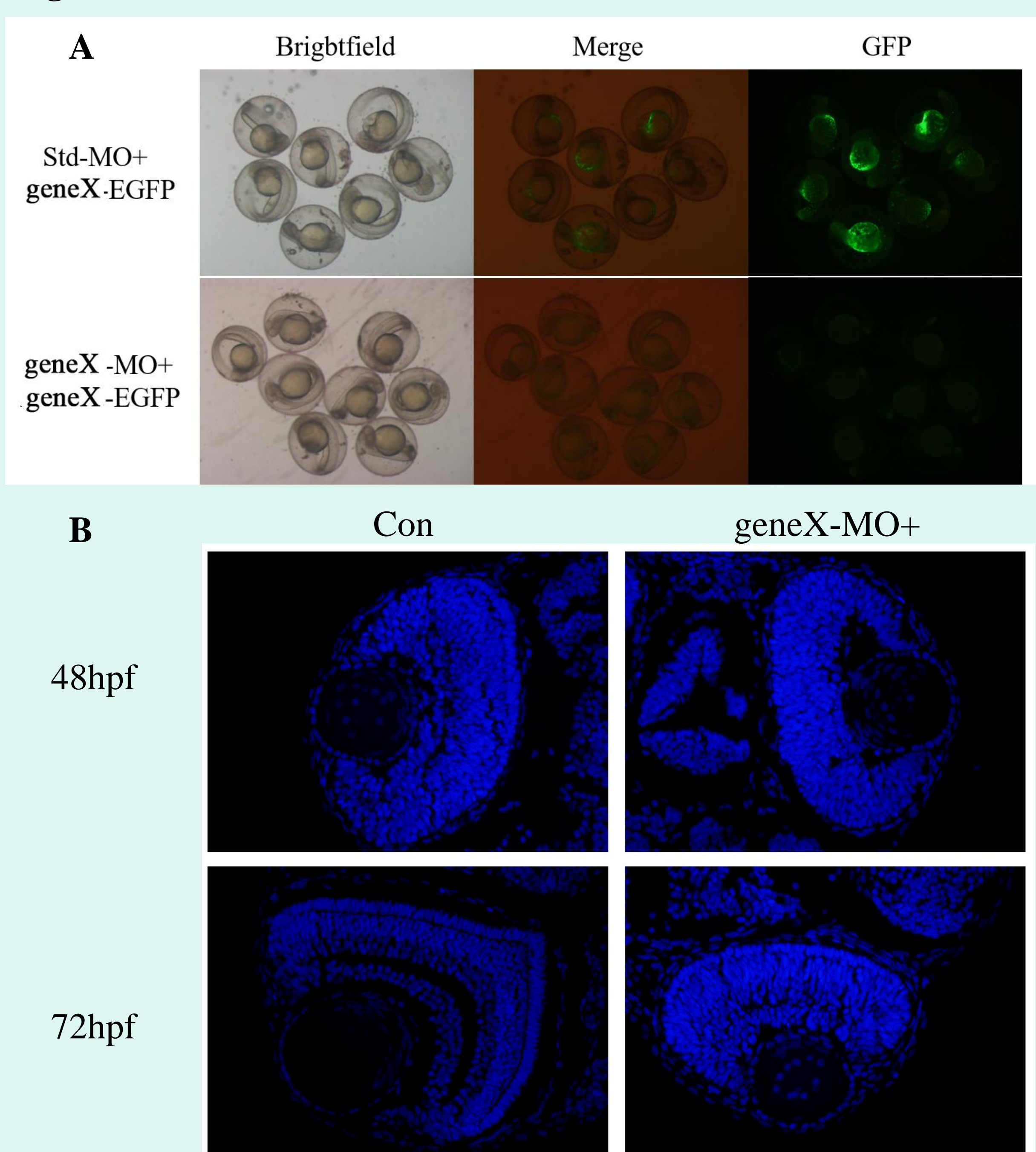


Fig. 3. Lens fiber cell defective differentiation and cataract formation



Tab. 1 LOD scores were calculated using parametric linkage analysis

Recombination Rate	c.G206A:p.R69H
0.00	2.81
0.05	2.58
0.10	2.34
0.15	2.08
0.20	1.81
0.25	1.51
0.30	1.20
0.35	0.86
0.40	0.52
0.45	0.21
0.50	0.00

Tab. 2. Bioinformatics analysis of gene X variant identified in this study

Category	Name	
Population databases	1000G_ALL	-
	ESP6500si_ALL	-
	ExAC_ALL	0.00007
	ExAC_EAS	-
	GnomAD_exome	0.00006
Missense prediction	dbSNP	-
	ClinVar	-
	SIFT	T
	Polyphen2_HVAR	D
	LRT	D
	MutationTaster	D
	PROVEAN	T
	Fathmm_MKL_coding	D
	MetaSVM	T
	M-CAP	D
conservation prediction	DANN_rankscore	0.938
	GERP++_RS_rankscore	0.949
	phyloP100way Vertebrate_rankscore	0.672
	phastCons100way Vertebrate	1
	SiPhy_29way_logOdds_rankscore	0.985