






OPEN ACCESS

Prediction of causative genes in inherited retinal disorder from fundus photography and autofluorescence imaging using deep learning techniques

Yu Fujinami-Yokokawa ^{1,2,3,4} Hideki Ninomiya,² Xiao Liu,¹ Lizhu Yang,¹ Nikolas Pontikos ^{1,3,5} Kazutoshi Yoshitake,⁶ Takeshi Iwata,⁶ Yasunori Sato,^{4,7} Takeshi Hashimoto,^{4,8} Kazushige Tsunoda,⁹ Hiroaki Miyata,^{2,4} Kaoru Fujinami ^{1,3,5}
The Japan Eye Genetics Study (JEGC) Group

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bjophthalmol-2020-318544>).

For numbered affiliations see end of article.

Correspondence to

Dr Kaoru Fujinami, Laboratory of Visual Physiology/Ophthalmic Genetics, Tokyo Iryo Center, Meguro-ku, Tokyo, Japan; k.fujinami@ucl.ac.uk

Received 1 December 2020
Revised 12 March 2021
Accepted 28 March 2021
Published Online First
20 April 2021



► <http://dx.doi.org/10.1136/bjophthalmol-2021-319365>



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Fujinami-Yokokawa Y, Ninomiya H, Liu X, et al. *Br J Ophthalmol* 2021;**105**:1272–1279.

ABSTRACT

Background/Aims To investigate the utility of a data-driven deep learning approach in patients with inherited retinal disorder (IRD) and to predict the causative genes based on fundus photography and fundus autofluorescence (FAF) imaging.

Methods Clinical and genetic data from 1302 subjects from 729 genetically confirmed families with IRD registered with the Japan Eye Genetics Consortium were reviewed. Three categories of genetic diagnosis were selected, based on the high prevalence of their causative genes: Stargardt disease (*ABCA4*), retinitis pigmentosa (*EYS*) and occult macular dystrophy (*RP1L1*). Fundus photographs and FAF images were cropped in a standardised manner with a macro algorithm. Images for training/testing were selected using a randomised, fourfold cross-validation method. The application program interface was established to reach the learning accuracy of concordance (target: >80%) between the genetic diagnosis and the machine diagnosis (*ABCA4*, *EYS*, *RP1L1* and normal).

Results A total of 417 images from 156 Japanese subjects were examined, including 115 genetically confirmed patients caused by the three prevalent causative genes and 41 normal subjects. The mean overall test accuracy for fundus photographs and FAF images was 88.2% and 81.3%, respectively. The mean overall sensitivity/specificity values for fundus photographs and FAF images were 88.3%/97.4% and 81.8%/95.5%, respectively.

Conclusion A novel application of deep neural networks in the prediction of the causative IRD genes from fundus photographs and FAF, with a high prediction accuracy of over 80%, was highlighted. These achievements will extensively promote the quality of medical care by facilitating early diagnosis, especially by non-specialists, access to care, reducing the cost of referrals, and preventing unnecessary clinical and genetic testing.

INTRODUCTION

Inherited retinal disorders (IRDs) are an important cause of irreversible blindness worldwide, particularly among working-age adults and children.^{1–3} In England and Wales, an epidemiological survey

reported that IRD accounted for 20.2% of blindness in working-age adults in 2009–2010.² The provision of accurate diagnosis of IRD is still difficult or unavailable in most of the world due to the limited access to multidisciplinary teams of specialists who can perform specific clinical investigations, including retinal imaging as well as genetic testing, interpretation of genetic results (ie, genetic diagnosis), and counselling.⁴ There are currently few approved treatment approaches for IRD.^{5–7} Thus, it is widely recognised that the development of accurate gene-specific diagnosis and novel therapeutic interventions is crucial to meet the urgent unmet needs of people suffering from blindness due to IRD.^{5–7}

Recently, deep learning techniques have been successfully applied in various medical fields, and the utilisation of machine learning-assisted diagnosis has been extensively promoted.^{8–10} A deep convolutional neural network (CNN) was first reported in 2012.¹¹ Since then, several software frameworks have been developed for CNNs, which include TensorFlow, PyTorch, Caffe and Theano.⁹

In particular, deep learning based on clinical images has been rapidly developed to predict diagnosis of common ocular disorders such as diabetic retinopathy and age-related macular degeneration.^{8–9–12–13} Abramoff *et al*¹³ reported an artificial intelligence (AI) system with high diagnostic accuracy (sensitivity: 87.2%, specificity: 90.7%) for automated detection of diabetic retinopathy and diabetic macular oedema, based on fundus photographs and spectral-domain optical coherence tomographic (SD-OCT) images, demonstrating AI's ability to bring specialty-level diagnostics to primary care settings.

In contrast, AI-oriented bioinformatic engineering had never been applied to ophthalmic orphan diseases such as IRD until 2019, even though IRD is the most prevalent cause of blindness and requires the largest number of hospital visits,¹⁴ which leads to major health burden and high economic costs (£523.3 million in the UK).¹⁵ Our team first published the utility of the AI-guided diagnostic system in IRD based on SD-OCT images

in 2019, illustrating a high test accuracy (98.5%) for predicting prevalent causative IRD genes.¹⁶

The purpose of this study was to investigate the utility of deep learning and establish an AI-guided automatic diagnosis system mainly targeting non-specialists from fundus photographs and fundus autofluorescence (FAF) images in a large Japanese cohort with IRD.

MATERIALS AND METHODS

Informed consent was obtained to undergo clinical and genetic testing and to use medical data for the current study.

Participants

Participants with a clinical diagnosis of IRD and available genetic data were studied until 2020 as part of the Japan Eye Genetics Consortium studies. A total of 1302 subjects from 729 families were enrolled for which genotype–phenotype association analyses were completed.^{17–20} Clinical diagnosis was performed based on comprehensive clinical investigations. A genetic diagnosis was obtained based on whole-exome sequencing with targeted analysis of 301 retinal diseases-associated genes.

The most common retinal diseases caused by the three major genes were selected. *EYS* (Mendelian Inheritance in Man [MIM]: 612424), *RP1L1* (MIM: 608581) and *ABCA4* (MIM: 601691) were the major genes. The proportions of *EYS*-associated retinal disease (*EYS* retinopathy; retinitis pigmentosa (RP) and others), *RP1L1* retinopathy (occult macular dystrophy) and *ABCA4* retinopathy (Stargardt disease and others) among all Japanese cases of IRD were 16%, 8% and 5%, respectively.^{18 20}

Fundus photography and FAF imaging

Fundus photography and FAF imaging were performed with the following equipment: TRC-50DX (Topcon, Tokyo, Japan), TRC-NW8 (Topcon), CR-2 PLUS AF Digital Non-Mydriatic Retinal Camera (Canon, Tokyo, Japan) and HRA II (excitation light 488 nm, barrier filter 500 nm; Heidelberg Engineering, Heidelberg, Germany).

Categories of genetic diagnosis

The fundus photographs and FAF images of four categories of patients were extracted. Encrypted clinical images, including fundus photographs and FAF images, were accessed by a certified ophthalmic genetic expert (KF). The categories, defined by genetic diagnosis, were as follows: category 1: *ABCA4* retinopathy; category 2: *RP1L1* retinopathy; category 3: *EYS* retinopathy; and category 4: normal. Typical images for each category were selected by a certified ophthalmic genetic expert.

Processes for training and testing

Fundus photographs and FAF images of both eyes were automatically or manually cropped into a standardised square shape (with aligning of the fovea at the centre) and adjusted to a spatial resolution of 72 pixels (pix)/inch (500×500 pix²). The images were cropped in a standardised way with the macro algorithm of the software (Adobe Photoshop CC V.20.0.4, San Jose, California). One image per eye at the latest examination was selected for the learning/testing processing by a certified ophthalmic genetic expert (KF). The algorithm for pipeline analyses was a web-based deep learning platform (MedicMind) which uses TensorFlow Inception V.3 (Alphabet, Mountain View, California). A CNN with an applied data set determined the learning parameters.

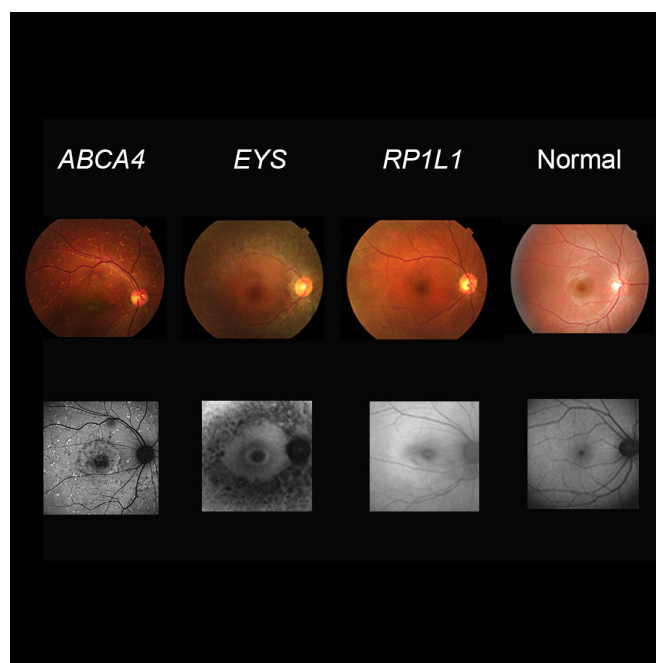


Figure 1 Representative cases from the four categories. Fundus photographs and fundus autofluorescence (FAF) images showing characteristic phenotypical features of the four categories are demonstrated: *ABCA4* retinopathy, *EYS* retinopathy, *RP1L1* retinopathy and normal. Fundus photographs: *ABCA4* retinopathy: macular atrophy and yellowish flecks at the level of RPE; *EYS* retinopathy: peripheral retinal atrophy, vessel attenuation and pigmentation at the level of RPE; *RP1L1* retinopathy: normal fundus appearance; normal: normal appearance. FAF images: *ABCA4* retinopathy: an area of low AF density at the macula and foci of abnormal AF; *EYS* retinopathy: diffuse abnormality of low AF in the peripheral retina and a ring of AF enhancement at the macula; *RP1L1* retinopathy: abnormal AF density at the fovea; normal: normal appearance. FAF, fundus autofluorescence; RPE, retinal pigment epithelium.

Four categories based on clinical and genetic diagnoses were applied: *ABCA4* retinopathy, *EYS* retinopathy, *RP1L1* retinopathy and normal (figure 1). After preparation of images for the four-gene categories, patients/subjects were randomly split into a training set and a test set at a 3:1 ratio (online supplemental figure 1). This random split for creating a training/test set was based on patients/subjects to avoid a confounding effect of the similarity between the eyes of the same patient/subject. Evaluations were conducted with a randomised, four-fold cross-validation method, and the accuracy of concordance (target: >80%) between the genetic diagnosis and the machine diagnosis was calculated through training to fix the application program interface (API) for further testing. Saliency maps of characteristic features based on fundus photographs and FAF images detected by API were investigated both in concordant and discordant cases between the machine diagnosis and the original genetic diagnosis.

Integrative assessment of learning performance

Integrative assessment of learning performance was conducted based on each diagnostic category by sensitivity/specificity prediction results from fundus photographs and FAF images. The area under the curve (AUC) was calculated with the probabilities for overall prediction results of fundus photography and FAF imaging (JMP, SAS Institute, Cary, North Carolina).

RESULTS

Participants

A total of 417 images from 156 Japanese subjects were examined, including 115 molecularly proven patients caused by the three major causative genes and 41 normal subjects. The detailed filtration/selection flow of the examined images is presented in online supplemental figure 2.

Fundus photography

From 149 probands, 1107 images were reviewed in total. Adequate data quality of 200 images from 102 probands (68.5%) was confirmed. Forty-seven probands were excluded from the analyses due to the inadequate data set/quality detected by the certified ophthalmic genetic expert; the exclusion criteria included missing data, presence of recording artefacts and so on. Images of both eyes were included for 98 probands, and images of one eye were included in 4 probands (1 patient with *ABCA4* retinopathy, 2 with *EYS* retinopathy and 1 with *RP1L1* retinopathy).

The median age at examination of the 102 probands was 49.5 years (range, 11–89 years). There were 58 females (56.9%) and 44 males (43.1%). All 102 probands were originally from Japan (East Asia). Fifty-nine images from 30 age-matched control subjects without ocular disorders were selected for the analysis. One image from one normal subject was unavailable.

In total, a data set containing 259 images from 132 patients/subjects was applied for deep learning; this data set comprised

41 images of *ABCA4* retinopathy, 94 images of *EYS* retinopathy, 65 images of *RP1L1* retinopathy and 59 normal images.

The training and test results of deep learning performance based on fundus photographs for prediction of the three causative genes (*ABCA4*, *EYS*, *RP1L1*) are presented in table 1.

The results for training accuracy are summarised in online supplemental table 1. Sufficient training accuracy over 87.5% was obtained. The mean training accuracy of the four repeated experiments was 91.8%. The mean sensitivity per gene category was 88.2% for *ABCA4* retinopathy, 97.8% for *EYS* retinopathy, 95.6% for *RP1L1* retinopathy and 80.2% for normal. The mean specificity was 98.2% for *ABCA4* retinopathy, 97.5% for *EYS* retinopathy, 94.8% for *RP1L1* retinopathy and 98.5% for normal.

The detailed test accuracy is shown in table 1. The mean test accuracy of the four repeated experiments was 88.2%. The mean sensitivity was 88.2% for *ABCA4* retinopathy, 88.4% for *EYS* retinopathy, 94.4% for *RP1L1* retinopathy and 82.9% for normal. The mean specificity was 100% for *ABCA4* retinopathy, 98.1% for *EYS* retinopathy, 92.9% for *RP1L1* retinopathy and 96.7% for normal, according to the four replications of the experiment.

Overall, considerably high specificity (>95%) was revealed for *ABCA4* retinopathy, *EYS* retinopathy and normal. High specificity (>90%) was identified in *RP1L1* retinopathy. High sensitivity (>90%) was detected in *RP1L1* retinopathy.

Table 1 Detailed training and test results based on fundus photography images for predicting the causative genes in inherited retinal disorder

Original classification with genetic diagnosis	Learning results			Predicted results						
	Training images (n)	Testing images (n)	Total	<i>ABCA4</i> images (n)	<i>EYS</i> images (n)	<i>RP1L1</i> images (n)	Normal images (n)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Trial 1										
<i>ABCA4</i>	30	11	41	8	–	3	–	73	100	–
<i>EYS</i>	70	24	94	1	22	1	–	92	98	–
<i>RP1L1</i>	49	16	65	–	–	16	–	100	88	–
Normal	43	16	59	1	1	2	12	75	100	–
Total	192	67	259	10	23	22	12	–	–	87
Trial 2										
<i>ABCA4</i>	31	10	41	9	1	–	–	90	100	–
<i>EYS</i>	73	21	94	2	18	1	–	86	98	–
<i>RP1L1</i>	47	18	65	–	–	16	2	89	85	–
Normal	43	16	59	–	–	6	10	63	96	–
Total	194	65	259	11	19	23	12	–	–	82
Trial 3										
<i>ABCA4</i>	31	10	41	10	–	–	–	100	100	–
<i>EYS</i>	66	28	94	1	24	–	3	86	100	–
<i>RP1L1</i>	47	18	65	–	–	16	2	89	98	–
Normal	42	17	59	–	–	1	16	94	91	–
Total	186	73	259	12	9	12	14	–	–	90
Trial 4										
<i>ABCA4</i>	31	10	41	9	1	–	–	90	100	–
<i>EYS</i>	73	21	94	2	19	–	–	90	97	–
<i>RP1L1</i>	52	13	65	–	–	13	–	100	100	–
Normal	49	10	59	–	–	–	10	100	100	–
Total	205	54	259	11	20	13	10	–	–	94

In total, 132 subjects with molecularly confirmed inherited retinal disorders or no ocular diseases were ascertained: 21 with *ABCA4* retinopathy, 48 with *EYS* retinopathy, 33 with *RP1L1* retinopathy and 30 normal subjects. Subjects were randomly split following a 3:1 ratio into training and test sets.

The commercially available deep learning tool, MedicMind, was applied to this four-class classification program.

The accuracy for each trial and the sensitivity and specificity for each category of each trial were calculated during the learning process, and this procedure was repeated four times with randomly assigned training/test sets to control for selection bias.

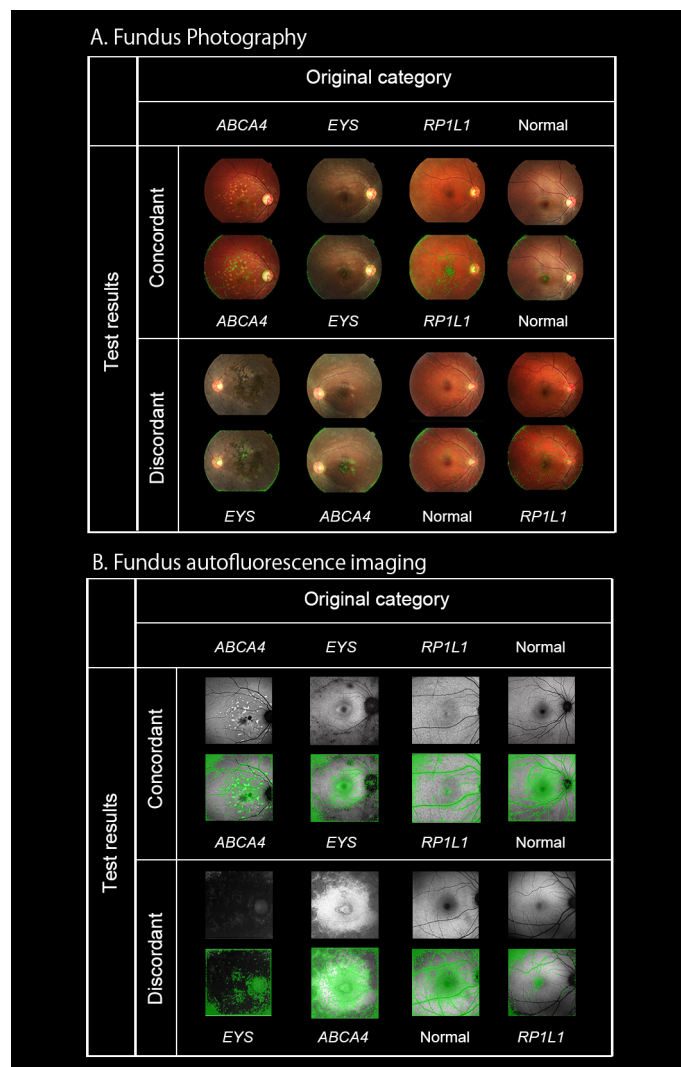


Figure 2 Saliency maps of the characteristic features of eight representative fundus photographs and eight fundus autofluorescence images detected by the application program interface. Saliency maps of characteristic features based on fundus photographs detected by the application program interface developed with deep learning are presented, indicating concordant and discordant results between the machine diagnosis and the original genetic diagnosis. (A) Fundus photography. Characteristic features of macular atrophy and flecks in *ABCA4* retinopathy and peripheral atrophic areas and attenuated vessels in *EYS* retinopathy were demonstrated in concordant cases. Macular vessels in *RP1L1* retinopathy and no particular findings in the normal category were identified in the concordant cases. Peripheral atrophic changes in *ABCA4* retinopathy and depigmented changes at the macula in *EYS* retinopathy were featured in the discordant cases. Macular vessels in a normal subject and no particular findings in *RP1L1* retinopathy were noted in the discordant cases. (B) Fundus autofluorescence (FAF) imaging. Characteristic features of an area of low AF density at the macula and foci of abnormal AF in *ABCA4* retinopathy and diffuse abnormality of low AF in the peripheral retina and a ring of AF enhancement at the macula in *EYS* retinopathy were demonstrated in the concordant cases. Vessels are intensively featured in *RP1L1* retinopathy, and vessels and macular pigment are prominent in normal category. A widespread area of low AF in *ABCA4* retinopathy and an area of low AF at the macula and diffuse peripheral atrophies of low AF in *EYS* retinopathy were featured in the discordant cases. Vessels and macular pigment were featured in *RP1L1* retinopathy, and vessels were prominent in a normal subject in discordant cases.

Saliency maps of the characteristic features of eight representative fundus photographs detected by the API are presented in [figure 2](#). Characteristic features of each category were identified in the concordant cases: macular atrophy and flecks in *ABCA4* retinopathy, peripheral atrophic area and attenuated vessels in *EYS* retinopathy, macular vessels in *RP1L1* retinopathy, and no particular findings in the normal category. The discordant cases demonstrated atypical features: peripheral atrophic changes in *ABCA4* retinopathy, depigmented changes at the macula in *EYS* retinopathy, no particular findings in *RP1L1* retinopathy and macular vessels in a normal subject.

FAF imaging

From 149 probands, 1002 images were reviewed. Adequate data quality of 115 images from 59 probands (39.6%) was confirmed. Ninety probands were excluded from the analysis due to the inadequate data set/quality detected by the certified ophthalmic genetic expert; the exclusion criteria included missing data, presence of recording artefacts and so on. Images of both eyes were included for 56 probands, and images of one eye were included for 4 probands (1 patient with *ABCA4* retinopathy, 1 with *EYS* retinopathy and 1 with *RP1L1* retinopathy).

Table 2 Detailed training and test results based on fundus autofluorescence images for predicting the causative genes in inherited retinal disorder

Original classification with genetic diagnosis	Learning results			Predicted results						
	Training images (n)	Testing images (n)	Total	<i>ABCA4</i> images (n)	<i>EYS</i> images (n)	<i>RP1L1</i> images (n)	Normal images (n)	Sensitivity (%)	Specificity (%)	Accuracy (%)
Trial 1										
<i>ABCA4</i>	27	10	37	10	–	–	–	100	100	–
<i>EYS</i>	25	10	35	–	7	–	3	70	100	–
<i>RP1L1</i>	33	10	43	–	–	7	3	70	97	–
Normal	31	12	43	–	–	1	11	92	80	–
Total	116	42	158	10	7	8	17	–	–	83
Trial 2										
<i>ABCA4</i>	27	10	37	9	1	–	–	90	100	–
<i>EYS</i>	27	8	35	–	8	–	–	100	97	–
<i>RP1L1</i>	31	12	43	–	–	9	3	75	97	–
Normal	32	11	43	–	–	1	10	91	90	–
Total	117	41	158	9	9	10	13	–	–	88
Trial 3										
<i>ABCA4</i>	28	9	37	9	–	–	–	100	91	–
<i>EYS</i>	25	10	35	3	7	–	–	70	100	–
<i>RP1L1</i>	32	11	43	–	–	6	5	55	100	–
Normal	32	11	43	–	–	–	11	100	83	–
Total	117	41	158	12	7	6	16	–	–	81
Trial 4										
<i>ABCA4</i>	29	8	37	8	–	–	–	100	89	–
<i>EYS</i>	28	7	35	3	3	1	–	43	100	–
<i>RP1L1</i>	33	10	43	–	–	6	4	60	92	–
Normal	34	9	43	–	–	1	8	89	84	–
Total	124	34	158	11	3	8	12	–	–	74

In total, 82 subjects with molecularly confirmed inherited retinal disorders or no ocular diseases were ascertained: 19 with *ABCA4* retinopathy, 18 with *EYS* retinopathy, 22 with *RP1L1* retinopathy and 23 normal subjects. Subjects were randomly split following a 3:1 ratio into training and test sets.

The commercially available deep learning tool, MedicMind, was applied to this four-class classification program.

The accuracy for each trial and the sensitivity and specificity for each category of each trial were calculated during the learning process, and this procedure was repeated four times with randomly assigned training/test sets to control for selection bias.

The median age at examination of the 59 probands was 47.0 years (range, 11–85 years). There were 31 females (52.5%) and 28 males (47.5%). All 59 probands were originally from Japan (East Asia). Forty-three images from 23 age-matched control subjects without no ocular disorders were selected for the analysis. Three images from three normal subjects were unavailable.

In total, a data set containing 158 images of 82 patients/subjects was applied for deep learning, comprising 37 images of *ABCA4* retinopathy, 35 images of *EYS* retinopathy, 43 images of *RP1L1* retinopathy and 43 normal images.

The training and test results for FAF-based prediction of the three causative genes (*ABCA4*, *EYS*, *RP1L1*) are presented in [table 2](#).

The results for training accuracy are summarised in online supplemental table 2. Sufficient training accuracy over 81.3% was obtained. The mean training accuracy of the four repeated experiments was 90.7%. The mean sensitivity per gene category was 100% for *ABCA4* retinopathy, 80.7% for *EYS* retinopathy, 85.0% for *RP1L1* retinopathy and 85.7% for normal. The mean specificity was 95.1% for *ABCA4* retinopathy, 100% for *EYS* retinopathy, 98.9% for *RP1L1* retinopathy and 93.2% for normal.

The detailed test accuracy is shown in [table 2](#). The mean test accuracy of the four repeated experiments was 81.3%. The mean sensitivity was 97.5% for *ABCA4* retinopathy, 70.7% for *EYS* retinopathy, 64.9% for *RP1L1* retinopathy and 92.9% for normal. The mean specificity was 94.8% for *ABCA4* retinopathy, 99.2%

for *EYS* retinopathy, 96.3% for *RP1L1* retinopathy and 84.3% for normal, according to the four replications of the experiment.

Overall, a considerably high level of sensitivity (>95%) was revealed in *ABCA4* retinopathy, and high sensitivity (>90%) was identified in *RP1L1* retinopathy. Considerably high specificity (>95%) was detected in *ABCA4* retinopathy, *EYS* retinopathy and *RP1L1* retinopathy. Low sensitivity (<80%) was found in *EYS* retinopathy. Considerably low sensitivity (<70%) was observed in *RP1L1* retinopathy.

Saliency maps of the characteristic features of eight representative FAF images detected by API are presented in [figure 2](#). Characteristic features of each category were identified in the concordant cases: an area of low AF density at the macula and foci of abnormal AF in *ABCA4* retinopathy, diffuse abnormality of low AF in the peripheral retina and a ring of AF enhancement at the macula in *EYS* retinopathy, retinal vessels in *RP1L1* retinopathy, and vessels and macular pigment in the normal category. The discordant cases demonstrated atypical features: a widespread area of low AF in *ABCA4* retinopathy, an area of low AF at the macula and diffuse peripheral atrophies of low AF in *EYS* retinopathy, vessels and macular pigment in *RP1L1* retinopathy, and vessels in normal category.

Integrative assessment of learning performance

Integrative assessment of learning performance was conducted based on each diagnostic category by using the sensitivity and

Table 3 Integrative assessment of overall prediction results of fundus photography and fundus autofluorescence (FAF) imaging

		<i>ABCA4</i>	<i>EYS</i>	<i>RP1L1</i>	Normal
Fundus photography	Sensitivity (%)	88	88	94	83
	Specificity (%)	100	98	93	97
Fundus autofluorescence imaging	Sensitivity (%)	98	71	65	93
	Specificity (%)	95	99	96	84

High specificity (>90%) was identified for both fundus photography and FAF imaging in *ABCA4* retinopathy, *EYS* retinopathy and *RP1L1* retinopathy. High sensitivity (>90%) of FAF imaging in *ABCA4* retinopathy, of fundus photography in *RP1L1* retinopathy and of FAF imaging in the normal category was detected. Low sensitivity (<80%) was identified in *EYS* retinopathy, and considerably low sensitivity (<60%) was detected in *RP1L1* retinopathy.

specificity of the prediction results from the two clinical imaging modalities (table 3).

High specificity (>90%) was identified for both fundus photography and FAF imaging in *ABCA4* retinopathy, *EYS* retinopathy and *RP1L1* retinopathy. High sensitivity (>90%) for FAF imaging in *ABCA4* retinopathy, for fundus photography in *RP1L1* retinopathy and for FAF imaging in normal category was detected. Low sensitivity (<80%) was identified in *EYS* retinopathy, and considerably low sensitivity (<70%) was detected in *RP1L1* retinopathy.

The AUC for the overall prediction results was 0.708 for fundus photography and 0.703 for FAF imaging (online supplemental figure 3).

DISCUSSION

The performance of a deep learning method to establish AI-guided automatic diagnosis systems based on fundus photography and FAF imaging was evaluated in a large Japanese cohort of patients with IRD. The accuracy of the automatic prediction of classification for each major causative gene was obtained based on 417 clinical images, with a mean overall sensitivity and specificity of 85.0% and 95.3%.

Potential efficacy of AI-guided prediction

The potential efficacy of this automatic screening/diagnostic system for major IRD genes based on fundus photography and FAF imaging was illustrated. Fundus photographs and FAF images demonstrated high specificity (>90%) in the identification of *ABCA4*, *EYS* and *RP1L1* retinopathies. The high sensitivity of fundus photographs for *RP1L1* retinopathy and FAF images for *ABCA4* retinopathy was also identified. These favourable results support the utility of the AI-guided diagnostic API in IRD.

This is the first report using deep learning technology for fundus photographs and FAF images in molecularly proven retinal orphan disorders. Since IRD exhibits strong gene-characteristic retinal features not strongly influenced by environmental factors, these cases present an ideal application of AI to assist clinical and genetic diagnoses. Recently, Miere *et al*²¹ reported the high gross accuracy (0.95) of AI-guided prediction of four phenotypic categories based on FAF images: RP, Best disease, Stargardt disease and normal. Together with this previous result, clinical imaging should be powerful to automatically predict the phenotypic category or the major causative genes in IRD. The three causative genes selected based on the disease prevalence in the current study typically show different clinical presentations described as Stargardt disease, RP and ocular macular dystrophy. This selection bias should give the advantage in predicting the causative gene. However, the phenotypic spectra of these three genes

overlap. It is challenging to establish perfect genotype–phenotype correlations in some cases, which supports the rationale of categorisation-based deep learning to predict the causative genes in these heterogeneous disease groups.

In addition, deducing the presence of IRD is difficult due to the very limited experience of general ophthalmologists; hence, AI-guided assessment could assist in the accurate identification of patients. Furthermore, given the high specificity in distinguishing between the normal category (96.7%) and the other abnormal categories, fundus photography can be considered a first screening method.

The API established in the current study is able to provide a real-time AI-guided diagnosis accessible from anywhere, which is promising to improve the quality of medical care by facilitating early diagnosis, reducing the cost for referrals, and preventing unnecessary clinical and genetic testing.

Interpretation of AI-guided diagnosis

ABCA4 retinopathy

The high specificity of both fundus photography and FAF imaging and the high sensitivity of fundus photography were identified. These prediction results are in accord with two previously reported clinical features of *ABCA4* retinopathy: (1) the preceding abnormalities can only be detected by FAF imaging^{22–23}; and (2) the fundus appearance in the end stage mimics RP.^{24–26} Subtle changes in FAF are crucial for early detection of the abnormal metabolic activity caused by the failure of *ABCA4* protein function.^{23–27} Morphological changes in the photoreceptor often occur in the central retina in the early stage.²⁸ Cases of *ABCA4* retinopathy in the end stage show entire drastic depigmented or pigmented changes, resembling the changes during the end stage of *EYS* retinopathy.^{24–26 29–31}

EYS retinopathy

The high specificity of both fundus photography and FAF imaging and the low sensitivity of fundus photography were identified. The prediction results for *EYS* retinopathy based on fundus photographs in the current study were compatible with previously reported features,^{19–32} while the prediction results based on FAF were not. The characteristic feature of peripheral retinal atrophy with vessel attenuation and pigmentation is frequently detected by fundus photographs.^{19–32} Some of the characteristic features were beyond the scope of the images, which were cropped in the process of standardisation; this limitation may have led to low sensitivity, with some diagnostic features missing from the fundus photographs. In the current study, the cropping process was mandatory. Expanding the use of a wide-field fundus/FAF recording could improve the sensitivity and specificity of the method for detecting abnormalities located in the peripheral retina.

RP1L1 retinopathy

The high specificity of both fundus photography and FAF imaging and the considerably low sensitivity of FAF imaging were identified. The prediction results of *RP1L1* retinopathy based on fundus photographs were more accurate than expected because the absence of fundus abnormality is one of the diagnostic criteria of autosomal dominant *RP1L1* retinopathy.³³ In general, it is assumed to be challenging to distinguish patients with *RP1L1* retinopathy from normal subjects based on fundus photographs. The prediction results for *RP1L1* retinopathy based on FAF are consistent with previously reported features^{34–35}; weak or subtle abnormalities are found in half of the patients. In

a previous study, 9 of 23 patients (39.1%) with occult macular dystrophy showed FAF abnormalities, while 14 patients (60.9%) exhibited no FAF abnormalities.²⁰ One unique feature of *RP1L1* retinopathy is the presence of intact retinal pigment epithelium, which is a potential reason why patients with *RP1L1* retinopathy do not present marked FAF abnormalities even with photoreceptor dysfunction at the macula.^{34 36}

Intriguingly, saliency maps demonstrated characteristic findings of macular vessels on fundus photographs and macular and arcade vessels on FAF images in the current study. These striking findings imply potential characteristic features of *RP1L1* retinopathy on retinal images that have not been reported, although confounding factors such as age and hypertension could not be completely excluded. Further comprehensive analyses with annotation data may elucidate the novel features of *RP1L1* retinopathy (figure 2).

Limitations

The current study involved several limitations that should be considered. The cohort size was relatively small both in the affected groups and the control group. Patients with IRD can be subcategorised into more detailed subsets based on the characteristic features by identifying the presence/absence of macular/mid-peripheral/peripheral atrophies, flecks, and functional deterioration of the macula and/or generalised rod/cone system. As an alternative, subdivision into typical or atypical could be useful for clinicians to interpret the quality of machine learning. AI-guided prediction of causative genes based on these phenotypic subcategories/subdivisions would improve the quality of diagnosis of ophthalmic genetic experts and provide opportunities to discover novel genotype–phenotype associations/correlations through the study processes. However, due to the limited cohort size, the phenotypic subcategory/subdivision-based machine learning was unavailable in the current study. In addition, the phenotypic subcategorisation/subdivision is challenging to perform before application of the API for non-specialists, who are the main target of the established diagnostic system in the current study. Thus, expanding the cohort size by including East Asian patients could allow us to further design/conduct deep learning of detailed phenotypic data for ophthalmic genetic experts to delineate the disease mechanism.

The severity of disease was not classified in the current study. Therefore, a selection bias may have existed. More extensive cohort studies with standardised data including demographics (family history, inheritance, symptoms and so on) and ocular conditions (visual acuity, visual field and so on) from subjects of various ethnicities could potentially expand the utility of our approach.

There are over 300 genes that can cause IRD; therefore, the four-class classifier developed in the current study cannot be usefully applied in practice to predict the status of a specific gene from fundus photographs and FAF images in patients with other retinal disease-associated gene diagnoses or a novel gene diagnosis. A more clinically applicable approach may be to train several one-versus-rest classifiers to distinguish one gene from all other genes.

There is also a current interpretability limitation to AI classifiers that use deep learning. The accuracy of training was variable among the training groups; thus, improving training accuracy in any training group set could potentially improve the quality of API in future studies. Moreover, future comparison studies of the quality between the machine diagnosis and the diagnosis by various human experts (eg, ophthalmic genetic experts, retinal

specialists, general ophthalmologists, ophthalmology residents, ophthalmic technicians, clinical geneticists and others) could prove the performance of AI-guided machine diagnosis in the real world.

CONCLUSION

This study illustrated a novel application of deep neural networks in the prediction of the major causative genes (30%) in IRD retinopathies from fundus photographs and FAF, with high accuracy for fundus photographs and relatively high accuracy for FAF. These achievements will extensively promote the improvement of medical care quality, reduce cost, enrich the education of non-specialists and support the application of personalised medicine.

Author affiliations

- ¹Laboratory of Visual Physiology, Division of Vision Research, National Institute of Sensory Organs, National Hospital Organization Tokyo Medical Center, Tokyo, Japan
²Department of Health Policy and Management, School of Medicine, Keio University, Tokyo, Japan
³UCL Institute of Ophthalmology, UCL, London, UK
⁴Graduate School of Health Management, Keio University, Tokyo, Japan
⁵Division of Inherited Eye Disease, Medical Retina, Moorfields Eye Hospital, London, UK
⁶Division of Molecular and Cellular Biology, National Institute of Sensory Organs, National Hospital Organization Tokyo Medical Center, Tokyo, Japan
⁷Department of Preventive Medicine and Public Health, Keio University School of Medicine, Tokyo, Japan
⁸Sports Medicine Research Center, Keio University, Tokyo, Japan
⁹Division of Vision Research, National Institute of Sensory Organs, National Hospital Organization Tokyo Medical Center, Tokyo, Japan

Twitter Nikolas Pontikos @npontikos

Acknowledgements English proof-reading has been conducted by Springer Nature Author Services. We deeply thank Professor Hong Sheng Chiong, MB BCH BAO, PGDipOphthBS, CEO of oDocs Eye Care, Otago, New Zealand, who provided support in the application of the software (MedicMind).

Collaborators The Japan Eye Genetics Study (JEGC) Group: Chair's Office: National Institute of Sensory Organs, Takeshi Iwata, Kazushige Tsunoda, Kaoru Fujinami, Shinji Ueno, Kazuki Kuniyoshi, Takaaki Hayashi, Mineo Kondo, Atsushi Mizota, Nobuhisa Naoi, Kei Shinoda, Shuhei Kameya, Hiroyuki Kondo, Taro Kominami, Hiroko Terasaki, Hiroyuki Sakuramoto, Satoshi Katagiri, Kei Mizobuchi, Natsuko Nakamura, Go Mawatari, Toshihide Kurihara, Kazuo Tsubota, Yozo Miyake, Kazutoshi Yoshitake, Toshihide Nishimura, Yoshihide Hayashizaki, Nobuhiro Shimozaawa, Masayuki Horiguchi, Shuichi Yamamoto, Manami Kuze, Shigeki Machida, Yoshiaki Shimada, Makoto Nakamura, Takashi Fujikado, Yoshihiro Hotta, Masayo Takahashi, Kiyofumi Mochizuki, Akira Murakami, Hiroyuki Kondo, Susumu Ishida, Mitsuru Nakazawa, Tetsuhisa Hatase, Tatsuo Matsunaga, Akiko Maeda, Kosuke Noda, Atsushi Tanikawa, Syuji Yamamoto, Hiroyuki Yamamoto, Makoto Araie, Makoto Aihara, Toru Nakazawa, Tetsuju Sekiryu, Kenji Kashiwagi, Kenjiro Kosaki, Carninci Piero, Takeo Fukuchi, Atsushi Hayashi, Katsuhiko Hosono, Keisuke Mori, Kouji Tanaka, Koichi Furuya, Keichirou Suzuki, Ryo Kohata, Yasuo Yanagi, Yuriko Minegishi, Daisuke Iejima, Akiko Suga, Brian P Rossmiller, Yang Pan, Tomoko Oshima, Mao Nakayama, Megumi Yamamoto, Naoko Minematsu, Daisuke Mori, Yusuke Kijima, Kentaro Kurata, Norihiro Yamada, Masayoshi Itoh, Hideya Kawaji, Yasuhiro Murakawa, Ryo Ando, Wataru Saito, Yusuke Murakami, Hiroaki Miyata, Lizhu Yang, Yu Fujinami-Yokokawa, Xiao Liu, Gavin Arno, Nikolas Pontikos, Mihori Kita, Hiroshi Hirose, Katsuyuki Sakai, Yasumasa Otori, Kazuki Yamazawa, Satomi Inoue, Takayuki Kinoshita.

Contributors KF and NP have full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Concept and design: YF-Y, NP, HM, KF. Acquisition, analysis or interpretation of data: all authors. Drafting of the manuscript: YF-Y, KF. Critical revision of the manuscript for important intellectual content: all authors. Statistical analysis: YF-Y, NP. Obtained funding: KF. Administrative, technical or material support: YF-Y, NP, HM, KF. Supervision: HM, KF.

Funding The Laboratory of Visual Physiology, Division of Vision Research, National Institute of Sensory Organs, National Hospital Organization Tokyo Medical Center, Tokyo, Japan is supported by grants from Astellas Pharma (NCT03281005), outside the submitted work. KF is supported by grant from Grant-in-Aid for Young Scientists (A) of the Ministry of Education, Culture, Sports, Science and Technology, Japan (16H06269), grant from Grant-in-Aid for Scientists to support international collaborative studies of the Ministry of Education, Culture, Sports, Science and Technology, Japan (16KK01930002), grant from the National Hospital Organization

Network Research Fund (H30-NHO-Sensory Organs-03), grant from the Japan Agency for Medical Research and Development (18992608), grant from the Ministry of Health, Labour and Welfare (18ek0109355h0001), grant from the Foundation Fighting Blindness Alan Lattes Career Development Program (CF-CL-0416-0696-UCL), grant from the Health Labour Sciences Research Grant, the Ministry of Health Labour and Welfare (201711107A), and grants from the Great Britain Sasakawa Foundation Butterfield Awards.

Disclaimer The funding sources had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Competing interests Individual investigators who participate in the sponsored project(s) are not directly compensated by the sponsor but may receive salary or other support from the institution to support their effort on the project(s). KF is a paid consultant of Astellas Pharma, Kubota Pharmaceutical Holdings and Acucela, Janssen Pharma, Novartis Pharma, and NightstaRx. KF reports personal fees from Astellas Pharma, personal fees from Kubota Pharmaceutical Holdings, personal fees from Acucela, personal fees from NightstaRx, personal fees from Santen, personal fees from Foundation Fighting Blindness, personal fees from Foundation Fighting Blindness Clinical Research Institute, personal fees from Japanese Ophthalmological Society, and personal fees from Japan Retinitis Pigmentosa Society.

Patient consent for publication Not required.

Ethics approval The protocol of this study adhered to the tenets of the Declaration of Helsinki and was approved by the local ethics committee, the National Hospital Organization Tokyo Medical Center (H30-1218002; R20-051).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Raw data were generated from the Japan Eye Genetics Consortium database and obtained with permission from the Japan Eye Genetics Consortium at <https://niso.kankakuki.go.jp/opkarte/>. Derived data supporting the findings of this study are available from the corresponding author on request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Yu Fujinami-Yokokawa <http://orcid.org/0000-0001-7555-5797>

Nikolas Pontikos <http://orcid.org/0000-0003-1782-4711>

Kaoru Fujinami <http://orcid.org/0000-0003-4248-0033>

REFERENCES

- Liew G, Michaelides M, Bunce C. A comparison of the causes of blindness certifications in England and Wales in working age adults (16-64 years), 1999-2000 with 2009-2010. *BMJ Open* 2014;4:e004015.
- Morizane Y, Morimoto N, Fujiwara A, et al. Incidence and causes of visual impairment in Japan: the first nation-wide complete enumeration survey of newly certified visually impaired individuals. *Jpn J Ophthalmol* 2019;63:26-33.
- Solebo AL, Teoh L, Rahi J. Epidemiology of blindness in children. *Arch Dis Child* 2017;102:853-7.
- Pontikos N, Murphy C, Moghul I, et al. Phenogenon: gene to phenotype associations for rare genetic diseases. *PLoS One* 2020;15:e0230587.
- Smith J, Ward D, Michaelides M, et al. New and emerging technologies for the treatment of inherited retinal diseases: a horizon scanning review. *Eye* 2015;29:1131-40.
- Scholl HPN, Strauss RW, Singh MS, et al. Emerging therapies for inherited retinal degeneration. *Sci Transl Med* 2016;8:368rv6-366.
- Hardcastle AJ, Sieving PA, Sahel J-A, et al. Translational retinal research and therapies. *Transl Vis Sci Technol* 2018;7:8.
- De Fauw J, Ledsam JR, Romera-Paredes B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nat Med* 2018;24:1342-50.
- Ting DSW, Peng L, Varadarajan AV, et al. Deep learning in ophthalmology: the technical and clinical considerations. *Prog Retin Eye Res* 2019;72:100759.
- Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-8.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 2012;2.
- Abràmoff MD, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 2016;57:5200-6.
- Abràmoff MD, Lavin PT, Birch M, et al. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digit Med* 2018;1:39.
- Roberts CB, Hiratsuka Y, Yamada M, et al. Economic cost of visual impairment in Japan. *Arch Ophthalmol* 2010;128:766-71.
- Galvin O, Chi G, Brady L, et al. The impact of inherited retinal diseases in the Republic of Ireland (ROI) and the United Kingdom (UK) from a cost-of-illness perspective. *Clin Ophthalmol* 2020;14:707-19.
- Fujinami-Yokokawa Y, Pontikos N, Yang L, et al. Prediction of causative genes in inherited retinal disorders from spectral-domain optical coherence tomography utilizing deep learning techniques. *J Ophthalmol* 2019;2019:1-7.
- Kameya S, Fujinami K, Ueno S, et al. Phenotypical characteristics of POC1B-Associated retinopathy in Japanese cohort: cone dystrophy with normal fundus appearance. *Invest Ophthalmol Vis Sci* 2019;60:3432-46.
- Yang L, Fujinami K, Ueno S, et al. Genetic spectrum of EYS-associated retinal disease in a large Japanese cohort: identification of disease-associated variants with relatively high allele frequency. *Sci Rep* 2020;10:5497.
- Fujinami-Yokokawa Y, Fujinami K, Kuniyoshi K, et al. Clinical and genetic characteristics of 18 patients from 13 Japanese families with CRX-associated retinal disorder: identification of genotype-phenotype association. *Sci Rep* 2020;10:9531.
- Fujinami K, Kameya S, Kikuchi S, et al. Novel RP11L variants and Genotype-Photoreceptor microstructural phenotype associations in cohort of Japanese patients with occult macular dystrophy. *Invest Ophthalmol Vis Sci* 2016;57:4837-46.
- Miere A, Le Meur T, Bitton K, et al. Deep Learning-Based classification of inherited retinal diseases using fundus autofluorescence. *J Clin Med* 2020;9:3303.
- Georgiou M, Kane T, Tanna P, et al. Prospective cohort study of childhood-onset Stargardt disease: fundus autofluorescence imaging, progression, comparison with adult-onset disease, and disease symmetry. *Am J Ophthalmol* 2020;211:159-75.
- Fujinami K, Zernant J, Chana RK, et al. Clinical and molecular characteristics of childhood-onset Stargardt disease. *Ophthalmology* 2015;122:326-34.
- Fujinami K, Lois N, Davidson AE, et al. A longitudinal study of Stargardt disease: clinical and electrophysiologic assessment, progression, and genotype correlations. *Am J Ophthalmol* 2013;155:e1013:1075-88.
- Fujinami K, Lois N, Mukherjee R, et al. A longitudinal study of Stargardt disease: quantitative assessment of fundus autofluorescence, progression, and genotype correlations. *Invest Ophthalmol Vis Sci* 2013;54:8181-90.
- Fujinami K, Sergouniotis PI, Davidson AE, et al. The clinical effect of homozygous ABCA4 alleles in 18 patients. *Ophthalmology* 2013;120:2324-31.
- Khan KN, Kasilian M, Mahroo OAR, et al. Early patterns of macular degeneration in ABCA4-Associated retinopathy. *Ophthalmology* 2018;125:735-46.
- Tanna P, Georgiou M, Strauss RW, et al. Cross-Sectional and longitudinal assessment of the ellipsoid zone in childhood-onset Stargardt disease. *Transl Vis Sci Technol* 2019;8:1.
- Fujinami K, Singh R, Carroll J, et al. Fine central macular dots associated with childhood-onset Stargardt disease. *Acta Ophthalmol* 2014;92:e157-9.
- Robson AG, Egan CA, Luong VA, et al. Comparison of fundus autofluorescence with photopic and scotopic fine-matrix mapping in patients with retinitis pigmentosa and normal visual acuity. *Invest Ophthalmol Vis Sci* 2004;45:4119-25.
- Robson AG, Tufail A, Fitzke F, et al. Serial imaging and structure-function correlates of high-density rings of fundus autofluorescence in retinitis pigmentosa. *Retina* 2011;31:1670-9.
- Suto K, Hosono K, Takahashi M, et al. Clinical phenotype in ten unrelated Japanese patients with mutations in the EYS gene. *Ophthalmic Genet* 2014;35:25-34.
- Miyake Y, Ichikawa K, Shiose Y, et al. Hereditary macular dystrophy without visible fundus abnormality. *Am J Ophthalmol* 1989;108:292-9.
- Fujinami K, Yang L, Joo K, et al. Clinical and genetic characteristics of East Asian patients with occult macular dystrophy (Miyake disease): East Asia occult macular dystrophy studies report number 1. *Ophthalmology* 2019;126:1432-44.
- Fujinami K, Tsunoda K, Hanazono G, et al. Fundus autofluorescence in autosomal dominant occult macular dystrophy. *Arch Ophthalmol* 2011;129:597-602.
- Yang L, Joo K, Tsunoda K, et al. Spatial functional characteristics of East Asian patients with occult macular dystrophy (Miyake disease); EAOMD report No.2. *Am J Ophthalmol* 2021;221:169-80.