


Matching variants for functional characterization of genetic variants

Sebiha Cevik,^{1,*} Pei Zhao,^{2,3,‡} Atiyye Zorluer,¹ Mustafa S. Pir,¹ Wenyin Bian,³ Oktay I. Kaplan ^{1,*}

¹Rare Disease Laboratory, School of Life and Natural Sciences, Abdullah Gul University, Kayseri 38080, Turkey

²School of Applied Science and Engineering, Fuzhou Institute of Technology, Fuzhou 350014, China

³SunyBiotech Co., Ltd., Fuzhou 350000, China

*Corresponding author: Email: sebiha.cevik@agu.edu.tr; *Corresponding author: Email: oktay.kaplan@agu.edu.tr

‡Both authors contributed equally to this work.

Rapid and low-cost sequencing, as well as computer analysis, have facilitated the diagnosis of many genetic diseases, resulting in a substantial rise in the number of disease-associated genes. However, genetic diagnosis of many disorders remains problematic due to the lack of interpretation for many genetic variants, especially missenses, the infeasibility of high-throughput experiments on mammals, and the shortcomings of computational prediction technologies. Additionally, the available mutant databases are not well-utilized. Toward this end, we used *Caenorhabditis elegans* mutant resources to delineate the functions of eight missense variants (V444I, V517D, E610K, L732F, E817K, H873P, R1105K, and G1205E) and two stop codons (W937stop and Q1434stop), including several matching variants (MatchVar) with human in ciliopathy associated IFT-140 (also called CHE-11)/IFT140 (intraflagellar transport protein 140). Moreover, MatchVars carrying *C. elegans* mutants, including IFT-140(G680S) and IFT-140(P702A) for the human (G704S) (dbSNP: rs150745099) and P726A (dbSNP: rs1057518064 and a conflicting variation) were created using CRISPR/Cas9. IFT140 is a key component of IFT complex A (IFT-A), which is involved in the retrograde transport of IFT along cilia and the entrance of G protein-coupled receptors into cilia. Functional analysis of all 10 variants revealed that P702A and W937stop, but not others phenocopied the ciliary phenotypes (short cilia, IFT accumulations, mislocalization of membrane proteins, and cilia entry of nonciliary proteins) of the IFT-140 null mutant, indicating that both P702A and W937stop are phenotypic in *C. elegans*. Our functional data offered experimental support for interpreting human variants, by using ready-to-use mutants carrying MatchVars and generating MatchVars with CRISPR/Cas9.

Keywords: matching variants; cilia; IFT140; ciliopathy; Mainzer–Saldino syndrome

Introduction

Rare diseases are classified as disorders that affect a small number of people. A recent estimation revealed that 3.5–5.9% of the world's population will be affected by nearly 6200 different rare diseases, 72% of which have genetic foundations (Nguengang Wakap et al. 2020). The genetic diagnosis of many rare diseases was hindered due to a lack of extensive human sequencing data that would have provided allele frequency estimates for each rare human variant. Because it is thought that rare genetic diseases are more likely to be caused by genetic variants with rare representation in human populations, researchers and clinical scientists sequenced a small sample of individuals with and without diseases to estimate the prevalence of a disease-causing variant. To help diagnosis of genetic diseases, the large genome consortiums, including the 1000 Genomes Project, Exome Sequencing Project, the Exome Aggregation Consortium, the genome aggregation database (gnomAD), and Trans-Omics for Precision Medicine (TOPMed) Program undertook the steps necessary to create the human variation databases that could provide the reference allele frequency (The 1000 Genomes Project Consortium 2010; Fu et al. 2013; Exome Aggregation Consortium et al. 2016; Karczewski et al. 2020; Taliun et al. 2021). Despite

substantial technological advancements in genomic sequencing and the availability of great databases for the reference allele frequency, the analysis of genomics data, the interpretation of each variant for that particular type of disease still presents a major challenge for patients and their families in the diagnosis of rare diseases. Large-scale human genome sequencing has revealed that there are a significant number of genetic variations (705,486,649 variants) in human populations. Furthermore, the underrepresentation of different ethnicities and populations in gnomAD and TOPMed makes it difficult to classify missense variants (Karczewski et al. 2020; Taliun et al. 2021). For this reason, a large percentage of human variants are classified as “variant of uncertain significance” (VUS) (Pir, Bilgin et al. 2022).

A variety of computer prediction methods are now available for analyzing the effects of genetic variants, but the classification of missense variations does not always correspond with the development of the disease (Ng 2003; Chun and Fay 2009; Adzhubei et al. 2010; Davydov et al. 2010; Reva et al. 2011; Shihab et al. 2013; Kircher et al. 2014; Schwarz et al. 2014). For instance, employing four different pathogenic prediction technologies, including Align-GVGD, SIFT, MutationTaster2, and PolyPhen-2, to classify 670 VUS in BRCA1 and BRCA2 revealed the insufficiency of

Received on 22 February 2023; accepted on 06 September 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of The Genetics Society of America.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

these tools for pathogenicity prediction (Ernst et al. 2018). Furthermore, high-throughput experiments have emerged as an alternate method of classifying variants, but scaling them to scan millions of variants is currently not feasible (Findlay et al. 2018; Giacomelli et al. 2018; Glazer et al. 2020).

Model organisms serve as a valuable testing system to discover the functional impacts of human genetic variants. Indeed, the ease of implementation of CRISPR/Cas9 genome editing has paved the way for such analysis, aiding to classify variants into functional categories, including phenotypic/pathogenic variants (McDiarmid et al. 2018; Wong et al. 2019). Furthermore, we recently unveiled a matching (equivalent) variants (MatchVars) search engine for *C. elegans*, mice, and human variants (Pir, Bilgin et al. 2022). Analysis of the Australian Phenome Bank (APB) and the Million Mutation Project databases revealed that they produced many mutants with missense variants, including MatchVars, whose potential had previously received less attention.

In the current study, we took advantage of available *C. elegans* mutants from Million Mutation Project database and concentrated on mutants bearing variants in intraflagellar transport 140 (IFT140). IFT is a cilia-specific bi-directional transport activity that has been preserved throughout evolution and is essential for the formation of cilia (Kozminski et al. 1993; Blacque et al. 2008). IFT140 is required for the return of IFT complex from the ciliary tip to the ciliary base. Furthermore, IFT140 is found to regulate the ciliary entry of G protein-coupled receptors (Absalon et al. 2008; Mukhopadhyay et al. 2010). IFT140 was implicated in several ciliopathies, including Short-Rib Thoracic Dysplasia 9 with or Without Polydactyly (SRTD9: 266920), [also known as Mainzer-Saldino syndrome (MSS), conorenal syndrome and Jeune asphyxiating thoracic dystrophy], nonsyndromic retinitis pigmentosa, syndromic congenital retinal dystrophy, and autosomal dominant polycystic kidney disease (Schmidts et al. 2013; Xu et al. 2015; Bifari et al. 2016; Bayat et al. 2017; Helm et al. 2017; Ali et al. 2023). However, the majority of variants (80%) in human IFT140 submitted to ClinVar are categorized as VUS (Landrum et al. 2018). We first identified MatchVars between *C. elegans* IFT-140 and human IFT140 using ConVarT, then we obtained 12 mutants harboring missense, including MatchVars and stop codons (V444I, V517D, E610K, L732F, E817K, H873P, R1105K, G1205E, R337stop, Q450stop, W937stop, and Q1434stop) to examine the functional implications of these variants (Pir, Bilgin et al. 2022; Pir, Cevik et al. 2022). Furthermore, we employed the CRISPR system to introduce precise amino acid substitutions at two different conserved amino acid positions (G680S and P702A) in *ift-140*, the ortholog of human IFT140 in *C. elegans*. Our functional analysis revealed that mutants bearing *ift-140*^{W937stop} or *ift-140*^{P702A} display null-mutant-like phenotypes, including lipophilic fluorescent dye uptake defects and IFT accumulations. Overall, our work shows that with the help of existing mutant resources and CRISPR/Cas9 systems, we can efficiently produce more data on human variants than we previously realized.

Materials and methods

Fluorescent lipophilic dye uptake assay

The fluorescent lipophilic dye was diluted with M9 (1:200 fluorescent lipophilic dye: M9 and fluorescent lipophilic dye were stored at -20°C) and incubated with mixed stages of healthy, well-fed *C. elegans* worms for 45 minutes. Wild types were always included in each Dye assay, while contaminated plates (bacteria or fungi) were not examined. At least 50 worms were scanned for each independent experiment. After it was confirmed that wild types

fully absorbed Dye, microscope images were collected. Following the *ift-140*(*syb1325*)*V*(*P702A*); *Ex*[*IFT-140::GFP + pRF4*] dye uptake assay, the dye rescue assay plate containing mutants with or without *Ex*[*IFT-140::GFP + pRF4*] was photographed to ensure that the dye uptake is fully restored. The same procedure was carried out for the plate carrying the mutation *che-11*(*gk925030*) (*R337stop*); *Ex*[*IFT-140::GFP + pRF4*]; however, there was no rescue for the dye uptake defect of *che-11*(*gk925030*) (*R337stop*). To better understand the cause of the dye uptake defect failure, *IFT-140::GFP* in *che-11*(*gk925030*) (*R337stop*) were photographed, which showed the ciliary accumulations of *IFT-140::GFP*.

Confocal microscopy imaging and subsequent image analysis

Before beginning the microscope analysis, 1 μL of 10 mM levamisole was applied to the freshly prepared microscope slides with a 2–3% agarose pad. These slides were placed into the microscope. Before the microscopy imaging, the worms were checked to determine whether worms were well-fed and healthy. Next, microscope images were collected with the Zeiss LSM900 confocal microscope equipped with Airyscan 2 and controlled by ZEN 3 Blue edition software. A Plan ApoChromat 63x/1.40 NA with 0.14 μm intervals was used to collect Z-stack images, and then the Blue edition software ZEN 3 generated Z-stack generation. ImageJ (NIH) and Fiji software were used for the subsequent image analysis (Schindelin et al. 2012).

Strains and genetic cross

The strains are maintained in a 20°C incubator on standard Nematode Growth Medium (NGM) plates and fed with *Escherichia coli* OP50 bacteria (Brenner 1974). Standard crossing techniques were utilized to generate transgenic worms with various mutant backgrounds. The rescue lines were generated by crossing *Ex*[*IFT-140::GFP + pRF4*] into the *ift-140*(*P702A*) mutant. PCR and dye-filling assay were used for confirmation of genotypes (genotyping *ift-140*(*P702A*) are available in Supplementary Fig. 1). *Ex*[*IFT-140::GFP + pRF4*] were crossed into VC41005 [(*ift-140*(*gk925030*) (*R337stop*))] mutants, the dye assay were used for confirmation. For the complete mutants they produced, the Million Mutation Project did whole-genome sequencing (WGS) for VC40309[(*ift-140*(*gk566606*) (*W937stop*))] and VC41005[(*ift-140*(*gk925030*) (*R337stop*))] revealed the details of mutations. The presence of mutations in the VC40309[(*che-11*(*gk566606*) (*W937stop*))] and VC41005[(*ift-140*(*gk925030*) (*R337stop*))] mutants was confirmed by Sanger sequencing. All strains and primers generated by the CRISPR/Cas9 system and genetic crossing were listed in Supplementary Tables 4 and 5.

CRISPR/Cas9 mediated generation of variants in IFT-140 in *C. elegans*

The Bristol N2 strain was used as the wild type background strain for CRISPR/Cas9 editing experiments. Genome editing was conducted as previously described but with small modifications (Arribere et al. 2014; Kim et al. 2014). In brief, sgRNA plasmid, Cas9 plasmid, reporter marker plasmid, repair template (oligo or plasmid), Co-CRISPR plasmid, or Co-conversion sgRNA plasmid (and Co-conversion repair oligo) were injected into N2 animals. F1 Animals with Co-CRISPR or Co-conversion phenotype were isolated and cultured in a single NGM plate. The F1 animals were lysed for screening of heterozygotes with targeted gene editing. F2 animals were cloned for homozygotes screening. The homozygotes were verified by PCR amplification and sequencing. The sgRNA target sites used in this study are listed

in [Supplementary Table 2](#). For the repair templates, we used the long oligos in IFT-140(G680S) editing and plasmid in IFT-140(P702A) editing. For the oligo templates, the homology arm is about 50 nt on either side of the mutated site. For the plasmid template, the homology arm is about 400 bp on either side of the mutated site. The nucleotide mutations generated in each gene editing experiment are listed in the [Supplementary Table 3](#). In IFT-140(P702A) editing experiments, precise gene editing introduced a restriction enzyme (Nru I) site. We thus amplify the genomic DNA sequences spanning each of the above four sites, and screen for the mutated animals through corresponding restriction enzyme digestion. For the IFT-140(G680S) editing, as there is no restriction site generated or destroyed, we used the allele-specific primers to screen for heterozygotes. The sequencing primers and genotyping primers used in each gene editing experiment are listed in the [Supplementary Table 4](#).

Multiple sequence alignments, conservation score, plot generation, and statistical analysis

Following the input of human IFT140 protein sequences (Human-NP_055529) into protein-protein Blast (BlastP), the following settings were used: Max target sequences: 5000, Max target sequences: 0.05, Matrix: BLOSUM62. In the BlastP results, we observed a total of 3232 proteins and FAST complete sequences were obtained from NCBI. Eventually, the query covers less than 40, and proteins from the same organisms were removed from the download file to eliminate duplicated and incorrect orthologs. Using the msa R package (packageVersion: "1.30.1"), multiple sequence alignments (MSA) of IFT140 orthologs from 1088 organisms were carried out to visualize the conservation score for each position ([Bodenhofer et al. 2015](#); [Winter 2017](#)). For a representative MSA and plot, the following organisms were chosen and the RefSeq Protein IDs of the following organisms were gathered: Human-NP_055529, Chimp-XP_016784643, Macaca-XP_001089057, Cow-XP_002697959, Mouse-NP_598887, Rat-XP_006246116, Frog-NP_001116497, Zebrafish-XP_695732, Fruit Fly-NP_995608, *C. elegans*-NP_506047, *C. reinhardtii*-XP_042921850, *T. thermophila*-XP_001020653, and *Paramecium tetraurelia*-XP_001347023. The rentrez package was used to download the protein sequences of these selected organisms after creating tables in R with accession numbers. MSA was carried out with the msa R package ([Bodenhofer et al. 2015](#); [Winter 2017](#)). Plots were generated using the following R packages, including trackViewer (packageVersion: "1.34.0"), ggmsa (packageVersion: "1.4.0"), and Welch's two-sample t-test (statistical analysis) were performed using R ([Ou and Zhu 2019](#); [Zhou et al. 2022](#)). R 4.2.2 was used throughout the R analysis.

Collection of ClinVar variants

IFT140 input was placed into the ClinVar website (<https://www.ncbi.nlm.nih.gov/clinvar/>). The numbers on the bottom left of the screen displayed the variation type, molecular consequences, and clinical significance. The variation type was plotted in [Fig. 2a](#). Selected missenses with molecular consequences were displayed in [Fig. 2b](#).

Results

Functional characterization of matching variants (MatchVars) in *C. elegans*

To leverage the already existing *C. elegans* mutants harboring variants, we initially gathered human and *C. elegans* variants for the

IFT140 gene from variant databases, including ClinVar, the gnomAD, the TOPMed, Wormbase and analyzed each position with ConVarT (Congruent clinical Variation Visualization Tool) to locate matching variants (MatchVar) between human IFT140 and *C. elegans* IFT-140 ([Fig. 1](#) and [Supplementary Table 1](#)). 80% of variation types in human IFT140 presented by ClinVar are single nucleotide alterations, including 761 missense variants (13th February 2023) ([Fig. 2, a and b](#)). The functional classifications of these variants in human IFT140 are far from complete because 80% of these missense variants are labeled as VUS ([Fig. 2b](#)).

Our analysis revealed that *C. elegans* ift-140 mutants contain a total of 39 variants ([Pir, Bilgin et al. 2022](#); [Pir, Cevik et al. 2022](#)). However, we restricted our analysis to 12 variants, primarily focusing on MatchVars and as well as stop codons. The stop codon-carrying mutants would reveal which domains are important for the function of IFT-140, while the MatchVars of human variants would be particularly relevant for clinical interpretation by physicians. The functional studies for each variant would be beneficial for medical geneticists requiring relevant functional data. As a result, the conclusions and findings drawn for each variant will be integrated into ConVarT (<https://convar.org/>), making them easily accessible to other researchers, medical geneticists, and healthcare professionals ([Pir, Bilgin et al. 2022](#); [Pir, Cevik et al. 2022](#)).

In light of this, human variants in IFT140 in the same position as those in *C. elegans* were listed together with the disease association, clinical significance, allele frequency, Evolutionary model of Variant Effect score (EVE), and conservation score for each position ([Fig. 2c](#)) ([Karczewski et al. 2020](#); [Frazer et al. 2021](#)). We employed a multiple sequence alignment of IFT140 proteins derived from 1088 distinct species to compute conservation scores for individual positions ([Supplementary Fasta](#)). A high conservation score indicates that a given position undergoes minimal variability, as the amino acid at that specific location likely plays a crucial role in the functionality of the respective proteins. ConVarT indicate that multiple variants, including *C. elegans* G1205E > human G1229R, are likely MatchVars ([Fig. 2d](#)). G1229R in human IFT140 is MSS-related variations of unknown significance (VUS). We obtained all these 12 mutants and a null ift-140 mutant from Caenorhabditis Genetics Center (CGC) ([Supplementary Table 1](#)). We found that several mutants, including VC20793 [IFT-140(Q450stop)], had delayed development and excluded them from further investigation.

We next went on to perform a functional investigation of each mutant. The structural intactness of the cilia in *C. elegans* is frequently examined using the fluorescent lipophilic dye (DiO) uptake assay. The sensory neurons in the head (amphid) and tail (phasmid) of the wild type completely take in the dye via their cilia, whereas the failure of DiO uptake in the head (amphid) and tail (phasmid) frequently indicates structurally compromised cilia as observed with ift-140 loss-of-function (lf) mutants ([Fig. 2e](#)). We anticipate finding dye absorption failure if mutants exhibit structural ciliary defects comparable to the null ift-140 mutant. Consistent with expectation, two different mutants (R337stop and W937stop), but not others display dye uptake failure, suggesting they might have defective cilia structures ([Fig. 2e](#)). Furthermore, we performed the dye-filling uptake assay following a shift to 25°C for 16 hours to identify the temperature-sensitive mutants, and mutants bearing the V444I mutation exhibit a minor dye uptake deficiency, but further examination found no significant structural abnormalities in two distinct cilia (data not shown). Both R337stop and W937stop mutants were independently crossed into CRISPR knock-in IFT-74::GFP allele and they display

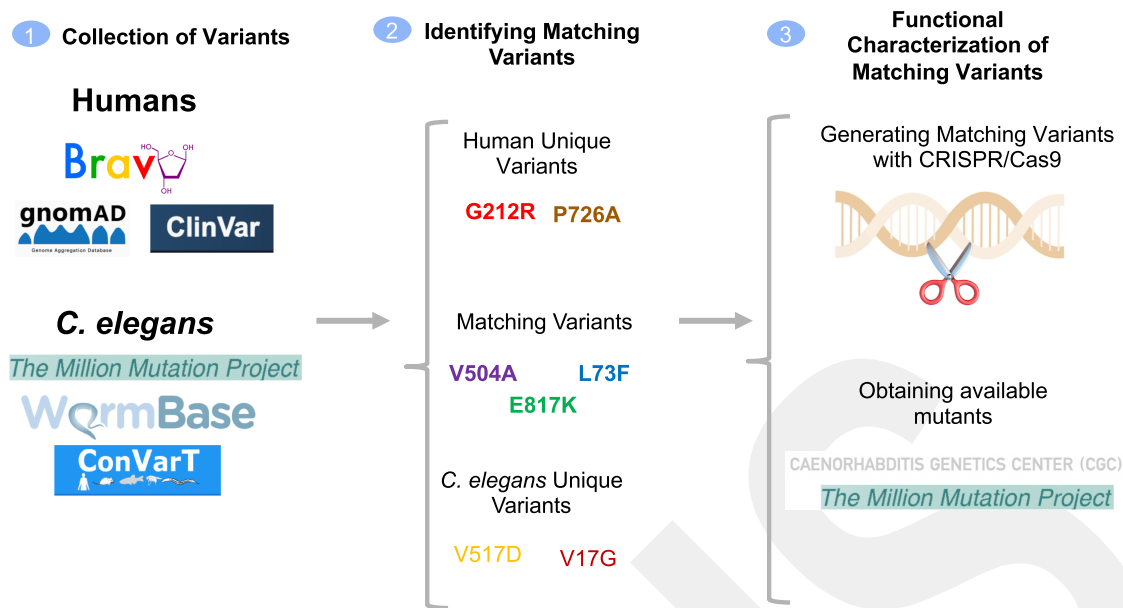


Fig. 1. The workflow for variant assessment. The workflow of the current study is displayed. *Caenorhabditis elegans* and human variants for *IFT140* were gathered from the indicated resources along with relevant info, such as allele frequency and clinical significance. Following the discovery of matching (equivalent) variants (MatchVars) and distinct variants, mutants bearing a missense or a stop codon were obtained from the CGC, or MatchVars were produced using the CRISPR/Cas9 genome editing tool.

cilia IFT accumulations resembling the null *ift-140* mutant, but the dye uptake defect of R337stop mutants could not be rescued with a functional IFT-140::GFP transgene, we, therefore, did not include the R337stop for further analysis and our predictions (Fig. 2f and data not shown). Introduction of IFT-140::GFP rescued the dye uptake defects observed in *ift-140*^{W937stop} mutants (Supplementary Fig. 2a). Similar to *ift-140(lf)*, our confocal microscopy analysis demonstrates that *ift-140*^{W937stop} mutants are shorter than those of the wild type (Fig. 2, f and g). The phenotype of truncated cilia was successfully rescued by the expression of functional IFT-140::GFP in *ift-140*^{W937stop} mutants (Fig. 2g and Supplementary Fig. 2b). Taken together, our strategy reveals W937stop [removing the last 599 amino acids, including the tetrapeptide repeat (TPR) domain] is likely a robust hypomorph or a functionally null allele of *ift-140*, but *ift-140*^{V444I}, *ift-140*^{V517D}, *ift-140*^{E610K}, *ift-140*^{L732F}, *ift-140*^{E817K}, *ift-140*^{H873P}, *ift-140*^{R1105K}, and *ift-140*^{G1205E} variants are not phenotypic, meaning that they are more likely benign.

P726A variant associated with short-rib thoracic dysplasia, but not G704S in IFT-140, is a loss of function mutation and leads to the ciliary accumulations of the IFT

As a part of our ongoing efforts to investigate the functional impact of MatchVars, we selected two ClinVar variants, including a conflicting variant interpretation (p.Pro726Ala) and a likely benign variant (p.Gly704Ser), to offer functional evidence for these two specific variants. Subsequently, we employed the CRISPR/Cas9 technology specifically to introduce the corresponding amino acid substitutions [Human IFT140(P726A) → *C. elegans* IFT-140(P702A) and human IFT140(G704S) → *C. elegans* IFT-140(G680S)] into the *C. elegans* IFT-140 (GenBank: NM_014714.3) (Fig. 3a). The human G704S in the IFT140 gene has not been implicated in IFT140-related disease whereas G212R and P726A (*C. elegans* P702A) compound heterozygosity in IFT140 resulted in short-rib thoracic dysplasia in a patient, therefore, P726A is

probably pathogenic, despite there has not yet been any functional data to classify the P726A variant in the IFT140 gene as a pathogenic variant (Fig. 3a) (Forbes et al. 2018). Additionally, the ClinVar put the P726A variant in the IFT140 gene in the category of conflicting pathogenicity interpretations. The position of P726 and G704 are highly conserved among IFT140 orthologs from the 1088 different species (Conservation scores: 94 for G704 and 97 for P726; Fig. 3b and Supplementary Fasta). The *C. elegans* IFT140(G680S) and IFT140(P702A) are MatchVars of human IFT140(G704S) (dbSNP: rs150745099) and IFT140(P726A) (dbSNP: rs1057518064), respectively (Fig. 3, c and d). We, therefore, employed *C. elegans* to experimentally investigate the functional effects of these two MatchVars: G680S (human G704S) and P702A (human P726A) in *C. elegans*.

The DiO uptake assay results show that *ift-140(lf)* mutants and mutants carrying homozygous P702A but not homozygous G680S fail to take up the fluorescent dye, suggesting the P to A change at the position of 702 likely results in defects in cilia structure, which is a strong indicator of disruption of IFT140 function. Importantly, dye-filling defect (Dyf) of *ift-140*^{P702A} mutants was fully rescued by introducing a wild-type copy of *ift-140/IFT140* into homozygous *ift-140*^{P702A} mutants, suggesting Dyf is due to P702A variant in IFT-140 (Fig. 3e).

IFT-140 is a critical part of IFT-A, hence its absence should cause IFT to be defective. We, therefore, crossed GFP-tagged IFT markers (endogenous IFT-74::GFP and OSM-3/KIF17::GFP) into *ift-140(lf)*, *ift-140*^{P702A} and *ift-140*^{G680S} mutants to visualize the IFT. Our confocal microscopy analysis revealed that both *ift-140(lf)* and *ift-140*^{P702A} mutants but not *ift-140*^{G680S} display remarkably similar phenotypes with the ciliary accumulation of a GFP-tagged IFT core machinery components (Fig. 4a, 100% ciliary accumulations in *ift-140*^{P702A}, *n* = 38). Furthermore, the PHA/PHB cilia are shorter in both *ift-140(lf)* and *ift-140*^{P702A} as compared to wild type and *ift-140*^{G680S} (Fig. 4b). We next went on investigating other cilia types, including AWA and AWB. The AWA olfactory neurons possess complex cilia with multiple branches, whereas

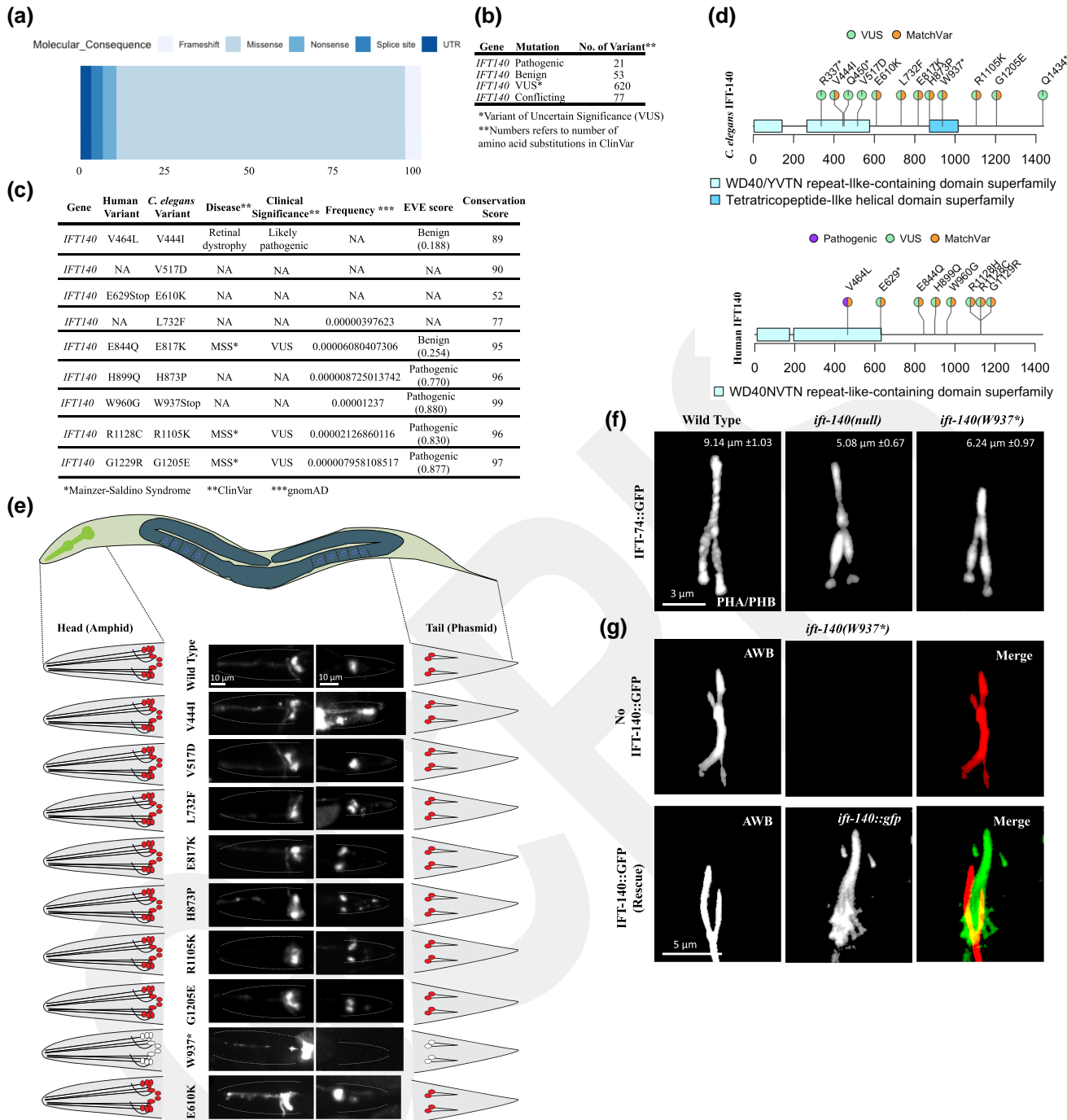


Fig. 2. The *ift-140*^{W937stop} mutants phenocopy defects of loss-of function *ift-140* mutants. a) Shown are the percentage distribution of the variation types for human *IFT140* in the ClinVar. b) Number of missense variants in human *IFT140* from ClinVar is displayed with clinical significance. c) *Caenorhabditis elegans* variants in *IFT-140* and human variants in *IFT140* corresponding to the human positions of *C. elegans* variants are presented. Clinical significance, disease and frequency (downloaded from gnomAD), EVE score, and conservation score (see Materials and Methods) were presented for human variants. d) The variants, including matching variants (MatchVars) from humans and *C. elegans* *IFT140* were displayed in their corresponding positions. VUS means variant of uncertain significance. e) Fluorescence images from the head and tail were displayed with representative drawings following the dye-filling assay. *C. elegans* cartoon shows the whole worm. The intermittent line in fluorescence images marks the edge of worms in the head and tail. With exception of mutants carrying W937stop, wild type and other worms absorb the lipophilic fluorescent dye in the head and tail. Asterisk (*) refers to the stop codon. Scale bar: 10 μm . f) Shown are the localization of *IFT-74::GFP* in the tail cilia (phasmid) of wild type, *ift-140(lf)*, and *ift-140*^{W937stop}. Scale bar of 3 μm . The average tail cilia lengths for wild type, *ift-140(lf)*, and *ift-140*^{W937stop} were presented. g) Shown are images depicting the AWB cilia in *ift-140*(W937*) mutants and the rescued line. The rescued line consists of *ift-140*(W937*) mutants expressing extrachromosomal *IFT-140::GFP*.

the AWB cilia are Y-type cilia, with short and long branches. Expectedly, the complexity of AWA drastically decreases in both *ift-140(lf)* and *ift-140*^{P702A} mutants, and the AWB cilia in both mutants display dramatically altered Y-type shape (Fig. 4c and

Supplementary Fig. 2c; 100% animals with phenotypes in AWB ($n = 28$) and AWA ($n = 12$) in *ift-140*^{P702A}).

Both GFP-tagged odorant response abnormal protein 10 (*ODR-10::GFP*, AWA cilia) and *OSM-9::GFP* (QLQ cilia, *C. elegans* ortholog

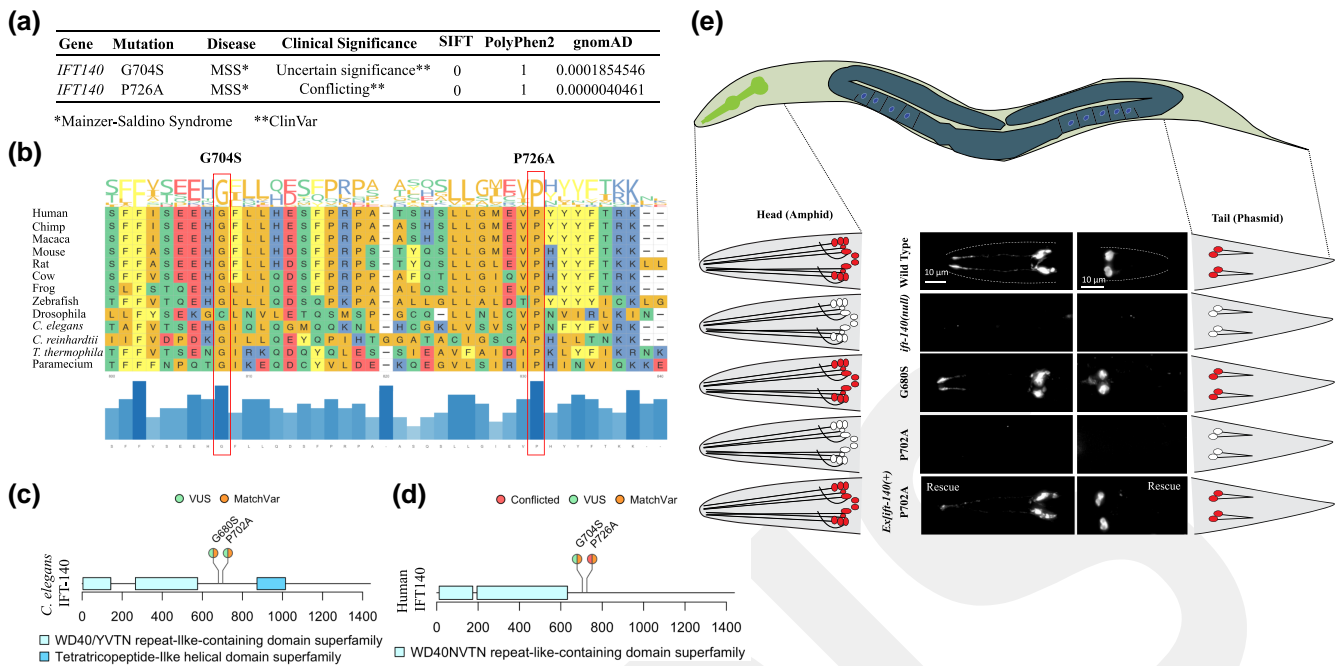


Fig. 3. The *C. elegans* MatchVars for human *IFT140*^{P726A} but not *IFT140*^{G704S} displays defects in the lipophilic fluorescent dye uptake in the head and tail. a) Clinical significance and disease relevance from ClinVar, frequency from gnomAD, SIFT, and PolyPhen2 score were presented for two human variants, P726A and G704S. MSS stands for Mainzer–Saldino syndrome. b) The positions of these two variants were shown on the MSA of *IFT140* orthologs from 13 different species. The conservation of the corresponding amino acid position across the 13 species is depicted by the blue bars at the bottom of the plot. The extended versions of MSAs were provided as supplementary files. c) The positions of human P726A and G704S variants and *C. elegans* P702A and G680S variants were shown in a lollipop plot of *IFT140* proteins from humans and *C. elegans*, respectively. The P702A and G680S are MatchVars of the human P726A and G704S in human *IFT140*, respectively. Clinical significance (VUS and conflicting) for human variants from ClinVar were presented. d) The entire worms, including heads and tails, were depicted in the *C. elegans* representative drawing. The fluorescence images were displayed alongside representative sketches of the head and tail cells. White indicates no dye uptake in the head and tail, while red indicates dye uptake in the cells. After the fluorescent dye assay, fluorescent images of the wild type and the indicated mutants were shown. *ift-140(lf)* and *ift-140*^{P702A} mutants were Dye negative in the head and tails. The expression of *ift-140* completely restored the dye uptake failure of *ift-140*^{P702A} mutants. Scale bar: 10 μ m.

of human TRPV4) were independently crossed into both *ift-140(lf)* and *ift-140*^{P702A} mutants, and following confirmation, wild type and mutants expressing *ODR-10::GFP* or *OSM-9::GFP* were imaged using the confocal microscopy. Microscopy analysis confirms that AWA cilia are severely impacted in both *ift-140(lf)* and *ift-140*^{P702A} mutants. The *ODR-10* protein was very low in the cilia in both mutants, but it was difficult to determine whether this was due to significantly altered cilia structure or a decrease in *ODR-10* staining in cilia (Fig. 5a). However, *OSM-9::GFP* stains the whole QLQ cilia in both mutants, indicating that *OSM-9::GFP* enters cilia, and both mutants exhibit *OSM-9* accumulations at the base of the cilia as compared to the wild type (Fig. 5a). Measuring fluorescence intensity along the distal dendrite and cilia provides supporting evidence for the accumulation of *OSM-9* (Fig. 5, b and c). Our analysis reveals that both mutants appear to have altered QLQ cilia morphology. Taken together, *ift-140*^{P702A} affects the localization of membrane proteins.

A previous study revealed that *IFT140* regulates the ciliary gate, thus we investigated the impact of P702A variant on the ciliary gate and protein trafficking (Scheidel and Blacque 2018). Translocation Associated Membrane Protein 1 ortholog (*TRAM-1*) was not inside cilia in the wild type, but *TRAM-1* was not the transition zone protein *MKS-2* (the human *TMEM216* orthologue) enters cilia in both *ift-140(lf)* and *ift-140*^{P702} mutants (Fig. 5e and Supplementary Fig. 3a). In contrast, the localization of transition zone protein *NPHP-1* (nephrocystin-1 orthologue) and the transition fiber protein *DYF-19* (the human *FBF1* orthologue) remains unaltered in these mutants (Fig. 5d, and Supplementary Fig. 3a).

Taken together, multiple lines of evidence suggest that *C. elegans* MatchVars for human *IFT140*^{P726A} but not *IFT140*^{G704S} disrupts the functions of *IFT140*, and *C. elegans ift-140*^{P702A} change likely represents a loss of function variant of *IFT-140*.

Discussion

Determining the consequence of human variants is important for diagnosing genetic diseases and for guiding the development of drug response regimes. It has been widely accepted that analyzing MatchVars from model organisms might shed light on the consequences of human MatchVars, including understanding the association between human variants and diseases (Wang et al. 2017, 2022; Platzer et al. 2019; The Alliance of Genome Resources Consortium et al. 2020; Zhu et al. 2020; Morbidoni et al. 2021; Di Rocco et al. 2022; Lange et al. 2022; Macaisne et al. 2022; AlAbdi et al. 2023). Recently, we published ConVarT, a search engine that displays disease and phenotypic data related to variants, including MatchVars from humans, mice, and *C. elegans* on MSA of orthologous genes from these three species (Pir, Bilgin et al. 2022; Pir, Cevik et al. 2022). The APB is a database of mouse strains carrying variations, including MatchVars in different genes, while the Million Mutation Project generated 2000 *C. elegans* mutants with many missense mutations in various genes (Thompson et al. 2013). There are currently 2359 mice mutant strains (5th February 2023) available through APB. The capability to explore the consequence of a single missense variant offers a great advantage for modeling the corresponding human MatchVars. Thus,

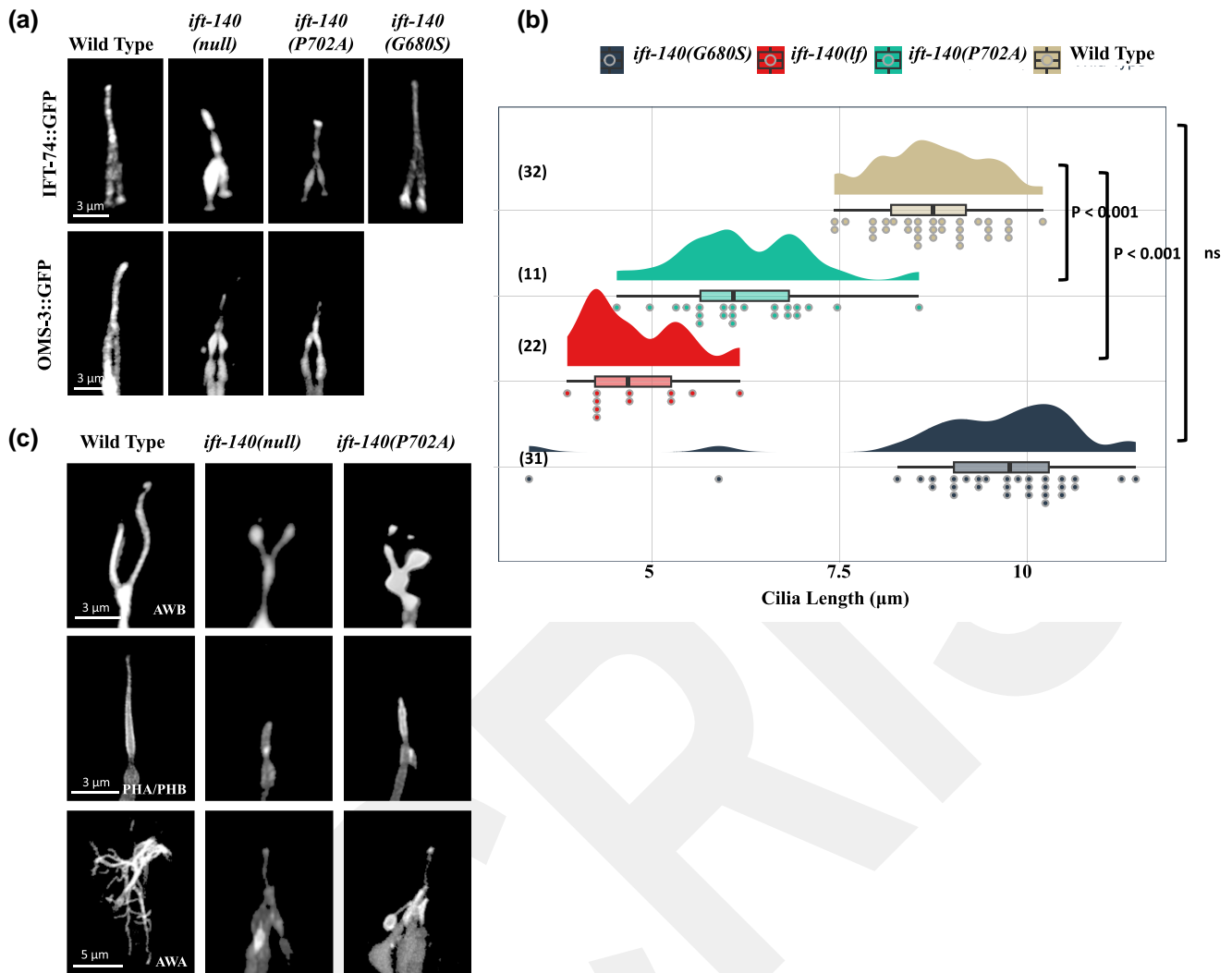


Fig. 4. The sensory cilia in the *C. elegans* IFT-140^{P702A} mutants are shortened. a) Shown are confocal images displaying the localization of fluorescently tagged IFT proteins in the tail (phasmid) of wild-type and the indicated mutants. Scale bar: 3 μ m. b) Plots of the PHA/PHB cilia length for wild-type and indicated mutants were shown. Not significant is abbreviated as ns. P values between the designated mutants and the wild type are also presented. Numbers in parentheses indicate the number of cilia used to calculate cilia length. c) Fluorescent markers display the following cilia: AWB cilia (Y-shaped cilia), PHA/PHB cilia (rod-shaped cilia), and AWA cilia (multiple complex branches). Shown are representative fluorescent images from wild type, *ift-140*(lf), and *ift-140*^{P702A} mutants. Scale bar (AWB): 3 μ m, scale bar (PHA/PHB): 3 μ m, and scale bar (AWA): 5 μ m.

these available *C. elegans* and mice mutant databases are valuable resources when searching for MatchVars for functional analysis. Indeed, we take advantage of the availability of *C. elegans* mutants bearing MatchVars and demonstrate that the use of these mutants can provide experimental evidence for the classification of human MatchVars. Our findings support the assertion that the MatchVars are useful tools for systematically assessing the molecular consequence of human variants and they can help provide functional proof for the clinically relevant variants. Moreover, CRISPR-mediated generation of MatchVars contributes to increasing the pools of MatchVars for functional evaluations, and the *C. elegans* community continuously creates MatchVars to evaluate their magnitude. This not only contributes to the functional characterization of human MatchVars in *C. elegans* but also provides valuable and independent resources for clinical scientists. This can be complementary to other findings for human variants. For example, the availability of phenotypic MatchVars from model organisms can be a great resource for clinical scientists when they look for evidence for a variant to decide on classifying it as a potential disease-causing variant.

In the current study, we chose *ift-140* for further evaluation because the phenotypic characterization of *ift-140* null mutants is straightforward. We first used ConVartT (<https://convalt.org>), which provided MSA along with the human, mouse, and *C. elegans* variants, to identify variants of interest (Pir, Bilgin et al. 2022; Pir, Cevik et al. 2022). We focused on 12 mutants, primarily studying MatchVars, and added two additional MatchVars using the CRISPR/Cas9 system. Expectedly, the Q1434stop variant of IFT140 (NCBI Reference Sequence: NP 506047 and 1437 amino acids) did not result in any anomalies in cilia because it just deletes the last three amino acids (data not shown). We are unable to come to any firm conclusions for R337Stop because the ciliary abnormalities (dye uptake defects) in the R337Stop variant mutants cannot be recovered with introduction of wild type copy of IFT-140. There are two plausible explanations for this unsuccessful rescue effort. First, the R337Stop variant may not be a null mutation; instead, it could exhibit antimorphic properties. Supporting this notion, we observed accumulations of GFP-tagged IFT-140 in the distal ciliary segment of homozygous R337Stop mutants; however, the absence of ciliary IFT-74 accumulations in

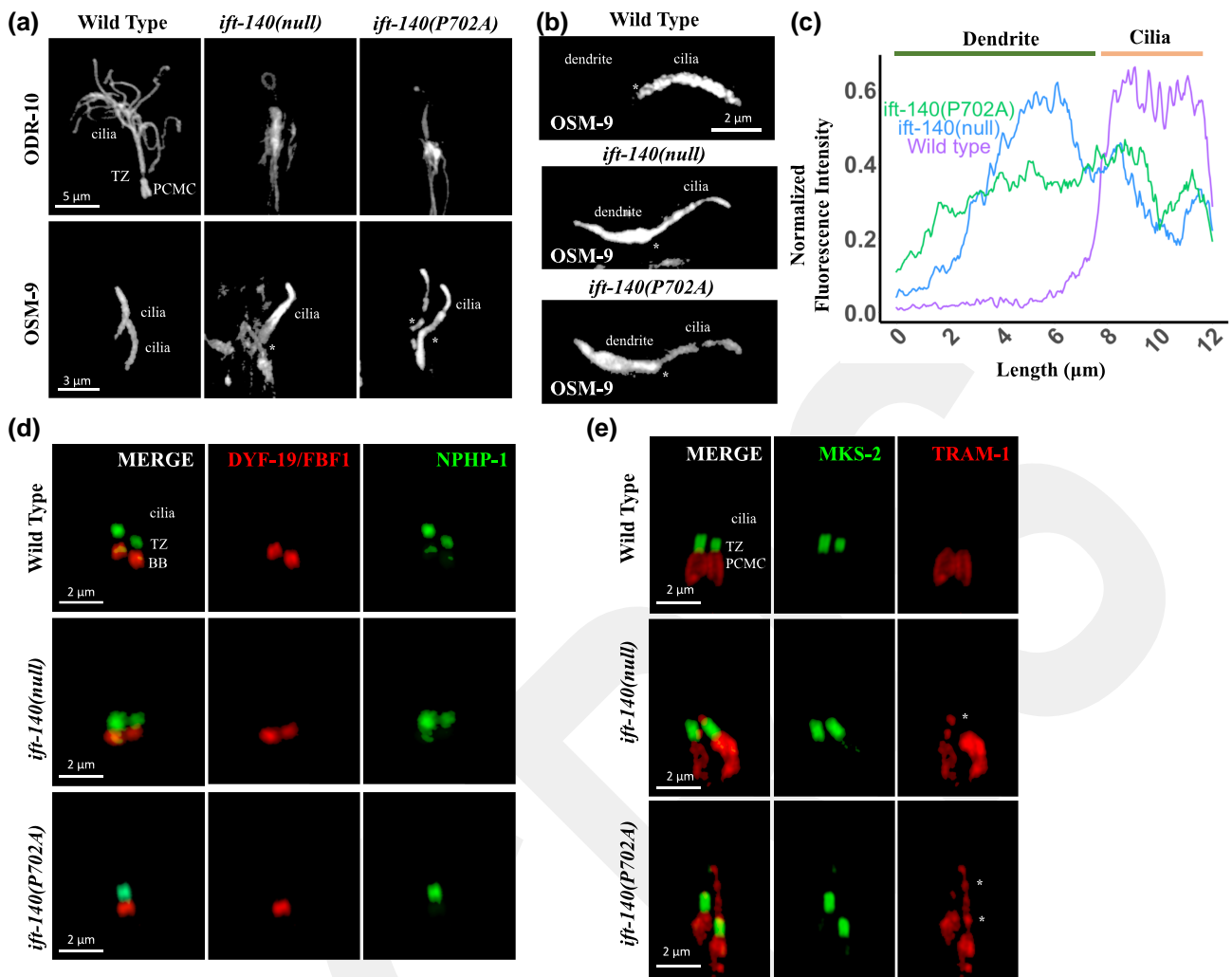


Fig. 5. The ciliary gate is altered in *C. elegans* IFT-140^{P702A} mutants. a) Localization of fluorescently tagged membrane proteins (ODR-10 and OSM-9) were shown in wild type, *ift-140(lf)*, and *ift-140*^{P702A} mutants. Scale bar (ODR-10): 5 μ m and scale bar (OSM-9): 3 μ m. Asterisk (*) shows accumulation of OSM-9 at the base of cilia. b) The images illustrate the distributions of OSM-9 in wild-type, *ift-140(lf)*, and *ift-140*^{P702A} mutants. The scale bar for OSM-9 is 2 μ m, and the asterisk (*) marks the ciliary base. c) The line plot displays the distributions of normalized fluorescence intensity measurements for the respective strains. The fluorescence intensity was measured and normalized in both the distal dendrite and cilia. d) Confocal images showing both the red-tagged DYF-19 (FBF-1, basal body protein) and the GFP-tagged NPHP-1 (a transition zone protein) were presented in wild type, *ift-140(lf)*, and *ift-140*^{P702A} mutants. BB and TZ stand for basal body and transition zone, respectively. Scale bar: 2 μ m. e) Shown are confocal images from nonciliary transmembrane protein TRAM-1 (labeled with red) and MKS-2 (green, a transition zone protein) in wild type, *ift-140(lf)*, and *ift-140*^{P702A} mutants. Asterisk (*) indicates the ciliary entry of TRAM-1. Scale bar: 2 μ m.

heterozygous R337Stop mutants contradicts this notion. If the R337Stop variant genuinely impacts the functioning of the wild-type IFT-140 copy, we would expect to observe ciliary accumulations of IFT-74 in heterozygous R337Stop mutants. An alternative hypothesis is that either a linked or unlinked second mutation could be responsible for these ciliary phenotypes. However, removing the last 500 amino acids (W937stop) containing half of the TPR causes ciliary IFT accumulation, with short cilia, indicating the functional importance of a tetratricopeptide-like helical domain. Consistent with this view, a recent study revealed that the TPR domain of IFT140 is important for its interaction with the TPR domains of IFT144 (Hesketh et al. 2022).

Our *C. elegans* work experimentally measured the functional consequence of eight missense variants in *C. elegans* IFT-140 (V444I, V517D, E610K, L732F, E817K, H873P, R1105K, and G1205E). All of these mutations were previously thought to be VUS; nevertheless, our findings now reclassify them as “likely

benign”. Despite this reclassification, it is critical to emphasize the shortcomings of the Dyf assay in classifying these variant mutants as benign because our assay might not detect modest defects present in these variant mutants.

Interestingly, the evolutionary conservation scores from MSA of 1088 IFT140 orthologs revealed that several residues (V517, E817, H873, R1105, and G1205) are highly evolutionarily conserved positions (conservation score ≥ 90) (Supplementary Fasta). However, none of these variations are situated in the amino acid position necessary for interaction between human IFT140 and IFT144 as the recent study revealed that in the human IFT140, many positions, including D at position 789, F at position 792, K at position 796, V at position 822, N at position 826, A at position 830, A at position 833, and R at position 837, are crucial for the expected interface between IFT140 and IFT144 (Hesketh et al. 2022).

The V464L missense variant in human IFT140 (dbSNP: rs2034681207 and NP 055529.2) was submitted to ClinVar as a likely pathogenic for Retinal dystrophy, however, there is no

functional evidence to support this claim. Our functional analysis suggests the V444I missense variant in *C. elegans* IFT-140 to be a benign mutation. Does the functional data with the V444I missense accurately reflect the V464L missense variant in the human IFT140 gene because the alteration is not the same? It could be due to the nonpolar and hydrophobic nature of the amino acids leucine (Leu) and isoleucine (Ile). Furthermore, consistent with our suggestion, the EVE (evolutionary model of variant effect) online tool predicts that the V464L missense variant in human IFT140 is likely benign (Frazer et al. 2021).

In addition, we utilized the CRISPR/Cas9 system to produce two missense variants (P702A and G680S) in IFT-140 in *C. elegans*, which are the MatchVars of P726A and G704S in human IFT140, respectively. The ClinVar database classifies the human IFT140^{G704S} missense variant (rs150745099) as likely benign, and our functional analysis supports this interpretation, as the *C. elegans* MatchVar of this variant does not result in a severe ciliary phenotype (Landrum et al. 2018). In contrast to human IFT140^{G704S}, the MatchVar of the human IFT140^{P726A} missense variant represents a loss of IFT-140 function in *C. elegans*. It is noteworthy that the ciliary phenotypes become apparent when the variant is present in a homozygous state, indicating the recessive nature of the IFT140^{P726A} variant. The human IFT140^{P726A} missense variant is a conflicting variant for clinical significance, and our findings provide functional evidence for the variant in favor of pathogenicity for the first time. Taken together by employing available mutant resources from model organisms and CRISPR/Cas9 systems in model organisms in determining the consequence of MatchVars, we can systematically generate more knowledge about human variants than previously realized.

Data availability

A comprehensive list of the recently created mutant and transgenic strains utilized in the figure generation process is available in [Supplementary Table 5](#). For those interested, these strains can be obtained by reaching out to Sebiha Cevik at sebiha.cevik@agu.edu.tr.

Acknowledgments

We thank Furkan Torun for helping to initiate the project, and Furkan Kepenek for MSA. Lastly, we would be happy to share all reagents generated as a part of this study.

Funding

Although we would love to thank a funding agency for their support, we regrettably did not receive any financial support for this project. Thus, we collaborated with SunyBiotech.

Conflicts of interest

The author(s) declare no conflict of interest.

Author contributions

PZ and WB generated two mutants used in [Figs. 3, 4, and 5](#). AZ generated many strains, but due to her pregnancy and the subsequent delivery of her gorgeous child, she was abruptly forced to leave her work. Many congrats! Sadly, all of her strains—with the exception of one—were discovered to be lost and not frozen. The project was taken over by SC, who expanded it by creating

all missense strains and additional strains to complete the work. SC carried out the all microscopy analyses. SC and OIK designed, and supervised the project, and prepared the figures. OIK and SC wrote the article.

Literature cited

- Absalon S, Blisnick T, Kohl L, Toutirais G, Doré G, Julkowska D, Tavenet A, Bastin P. 2008. Intraflagellar transport and functional analysis of genes required for Flagellum formation in trypanosomes. *Mol Biol Cell*. 19(3):929–944. doi:[10.1091/mbc.e07-08-0749](#).
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods*. 7(4):248–249. doi:[10.1038/nmeth0410-248](#).
- AlAbdi L, Desbois M, Rusnac D-V, Sulaiman RA, Rosenfeld JA, Lalani S, Murdock DR, Burrage LC, Network UD, Au PYB, et al. 2023. Loss-of-function variants in MYCBP2 cause neurobehavioural phenotypes and corpus callosum defects. *Brain*. 146(4):1373–1387. doi:[10.1093/brain/awac364](#).
- Ali H, Naim M, Senum SR, AlSahow A, Bahbahani Y, Abu-Farha M, Abubaker J, Mohammad A, Al-Hunayan A, Asbeutah AM, et al. 2023. The genetic landscape of autosomal dominant polycystic kidney disease in Kuwait. *Clin Kidney J*. 16(2):355–366. doi:[10.1093/cjk/sfac236](#).
- Arribere JA, Bell RT, Fu BXH, Artiles KL, Hartman PS, Fire AZ. 2014. Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics*. 198(3):837–846. doi:[10.1534/genetics.114.169730](#).
- Bayat A, Kerr B, Douzgou S. 2017. The evolving craniofacial phenotype of a patient with Sensenbrenner syndrome caused by IFT140 compound heterozygous mutations. *Clin Dysmorphol*. 26(4):247–251. doi:[10.1097/MCD.000000000000169](#).
- Bifari IN, Elkharny SM, Bolz HJ, Khan AO. 2016. The ophthalmic phenotype of IFT140-related ciliopathy ranges from isolated to syndromic congenital retinal dystrophy. *Br J Ophthalmol*. 100(6):829–833. doi:[10.1136/bjophthalmol-2015-307555](#).
- Blacque OE, Cevik S, Kaplan OI. 2008. Intraflagellar transport: from molecular characterisation to mechanism. *Front Biosci* 13(13), 2633. doi:[10.2741/2871](#).
- Bodenhofer U, Bonatesta E, Horejš-Kainrath C, Hochreiter S. 2015. Msa: an R package for multiple sequence alignment. *Bioinformatics*. 31(24):3997–3999. doi:[10.1093/bioinformatics/btv494](#).
- Brenner S. 1974. The genetics of *Caenorhabditis elegans*. *Genetics*. 77(1):71–94. doi:[10.1093/genetics/77.1.71](#).
- Chun S, Fay JC. 2009. Identification of deleterious mutations within three human genomes. *Genome Res*. 19(9):1553–1561. doi:[10.1101/gr.092619.109](#).
- Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. 2010. Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol*. 6(12):e1001025. doi:[10.1371/journal.pcbi.1001025](#).
- Di Rocco M, Galosi S, Lanza E, Tosato F, Caprini D, Folli V, Friedman J, Bocchinfuso G, Martire A, Di Schiavi E, et al. 2022. *Caenorhabditis elegans* provides an efficient drug screening platform for GNAO1-related disorders and highlights the potential role of caffeine in controlling dyskinesia. *Hum Mol Genet*. 31(6):929–941. doi:[10.1093/hmg/ddab296](#).
- Ernst C, Hahnen E, Engel C, Nothnagel M, Weber J, Schmutzler RK, Hauke J. 2018. Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics. *BMC Med Genomics*. 11(1):35. doi:[10.1186/s12920-018-0353-y](#).
- Exome Aggregation Consortium; Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS,

- Hill AJ, et al. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 536(7616):285–291. doi:[10.1038/nature19057](https://doi.org/10.1038/nature19057).
- Findlay GM, Daza RM, Martin B, Zhang MD, Leith AP, Gasperini M, Janizek JD, Huang X, Starita LM, Shendure J. 2018. Accurate classification of BRCA1 variants with saturation genome editing. *Nature*. 562(7726):217–222. doi:[10.1038/s41586-018-0461-z](https://doi.org/10.1038/s41586-018-0461-z).
- Forbes TA, Howden SE, Lawlor K, Phipson B, Maksimovic J, Hale L, Wilson S, Quinlan C, Ho G, Holman K, et al. 2018. Patient-iPSC-derived kidney organoids show functional validation of a ciliopathic renal phenotype and reveal underlying pathogenetic mechanisms. *Am J Hum Genet*. 102(5):816–831. doi:[10.1016/j.ajhg.2018.03.014](https://doi.org/10.1016/j.ajhg.2018.03.014).
- Frazer J, Notin P, Dias M, Gomez A, Min JK, Brock K, Gal Y, Marks DS. 2021. Disease variant prediction with deep generative models of evolutionary data. *Nature*. 599(7883):91–95. doi:[10.1038/s41586-021-04043-8](https://doi.org/10.1038/s41586-021-04043-8).
- Fu W, O'Connor TD, Jun G, Kang HM, Abecasis G, Leal SM, Gabriel S, Rieder MJ, Altshuler D, Shendure J, et al. 2013. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature*. 493(7431):216–220. doi:[10.1038/nature11690](https://doi.org/10.1038/nature11690).
- Giacomelli AO, Yang X, Lintner RE, McFarland JM, Duby M, Kim J, Howard TP, Takeda DY, Ly SH, Kim E, et al. 2018. Mutational processes shape the landscape of TP53 mutations in human cancer. *Nat Genet*. 50(10):1381–1387. doi:[10.1038/s41588-018-0204-y](https://doi.org/10.1038/s41588-018-0204-y).
- Glazer AM, Wada Y, Li B, Muhammad A, Kalash OR, O'Neill MJ, Shields T, Hall L, Short L, Blair MA, et al. 2020. High-throughput reclassification of SCN5A variants. *Am J Hum Genet*. 107(1):111–123. doi:[10.1016/j.ajhg.2020.05.015](https://doi.org/10.1016/j.ajhg.2020.05.015).
- Helm BM, Willer JR, Sadeghpour A, Golzio C, Crouch E, Vergano SS, Katsanis N, Davis EE. 2017. Partial uniparental isodisomy of chromosome 16 unmasks a deleterious biallelic mutation in IFT140 that causes Mainzer-Saldino syndrome. *Hum Genomics*. 11(1):16. doi:[10.1186/s40246-017-0111-9](https://doi.org/10.1186/s40246-017-0111-9).
- Hesketh SJ, Mukhopadhyay AG, Nakamura D, Toropova K, Roberts AJ. 2022. IFT-A structure reveals carriages for membrane protein transport into cilia. *Cell*. 185(26):4971–4985.e16. doi:[10.1016/j.cell.2022.11.010](https://doi.org/10.1016/j.cell.2022.11.010).
- Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, Collins RL, Laricchia KM, Ganna A, Birnbaum DP, et al. 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 581(7809):434–443. doi:[10.1038/s41586-020-2308-7](https://doi.org/10.1038/s41586-020-2308-7).
- Kim H, Ishidate T, Ghanta KS, Seth M, Conte D, Shirayama M, Mello CC. 2014. A co-CRISPR strategy for efficient genome editing in *Caenorhabditis elegans*. *Genetics*. 197(4):1069–1080. doi:[10.1534/genetics.114.166389](https://doi.org/10.1534/genetics.114.166389).
- Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 46(3):310–315. doi:[10.1038/ng.2892](https://doi.org/10.1038/ng.2892).
- Kozminski KG, Johnson KA, Forscher P, Rosenbaum JL. 1993. A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proc Natl Acad Sci USA*. 90(12):5519–5523. doi:[10.1073/pnas.90.12.5519](https://doi.org/10.1073/pnas.90.12.5519).
- Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, Gu B, Hart J, Hoffman D, Jang W, et al. 2018. Clinvar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 46(D1):D1062–D1067. doi:[10.1093/nar/gkx1153](https://doi.org/10.1093/nar/gkx1153).
- Lange KI, Best S, Tsiropoulou S, Berry I, Johnson CA, Blacque OE. 2022. Interpreting ciliopathy-associated missense variants of uncertain significance (VUS) in *Caenorhabditis elegans*. *Hum Mol Genet*. 31(10):1574–1587. doi:[10.1093/hmg/ddab344](https://doi.org/10.1093/hmg/ddab344).
- Macaisne N, Touzon MS, Rajkovic A, Yanowitz JL. 2022. Modeling primary ovarian insufficiency-associated loci in *C. elegans* identifies novel pathogenic allele of MSH5. *J Assist Reprod Genet*. 39(6):1255–1260. doi:[10.1007/s10815-022-02494-0](https://doi.org/10.1007/s10815-022-02494-0).
- McDiarmid TA, Au V, Loewen AD, Liang J, Mizumoto K, Moerman DG, Rankin CH. 2018. CRISPR-Cas9 human gene replacement and phenomic characterization in *Caenorhabditis elegans* to understand the functional conservation of human genes and decipher variants of uncertain significance. *Dis Model Mech*. 11(12):dmm036517. doi:[10.1242/dmm.036517](https://doi.org/10.1242/dmm.036517).
- Morbidity V, Agolini E, Slep KC, Pannone L, Zuccarello D, Cassina M, Grosso E, Gai G, Salviati L, Dallapiccola B, et al. 2021. Biallelic mutations in the TOGARAM1 gene cause a novel primary ciliopathy. *J Med Genet*. 58(8):526–533. doi:[10.1136/jmedgenet-2020-106833](https://doi.org/10.1136/jmedgenet-2020-106833).
- Mukhopadhyay S, Wen X, Chih B, Nelson CD, Lane WS, Scales SJ, Jackson PK. 2010. TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia. *Genes Dev*. 24(19):2180–2193. doi:[10.1101/gad.1966210](https://doi.org/10.1101/gad.1966210).
- Ng PC. 2003. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res*. 31(13):3812–3814. doi:[10.1093/nar/gkg509](https://doi.org/10.1093/nar/gkg509).
- Nguengang Wakap S, Lambert DM, Olry A, Rodwell C, Gueydan C, Lanneau V, Murphy D, Le Cam Y, Rath A. 2020. Estimating cumulative point prevalence of rare diseases: analysis of the Orphanet database. *Eur J Hum Genet*. 28(2):165–173. doi:[10.1038/s41431-019-0508-0](https://doi.org/10.1038/s41431-019-0508-0).
- Ou J, Zhu LJ. 2019. Trackviewer: a bioconductor package for interactive and integrative visualization of multi-omics data. *Nat Methods*. 16(6):453–454. doi:[10.1038/s41592-019-0430-y](https://doi.org/10.1038/s41592-019-0430-y).
- Pir MS, Bilgin HI, Sayici A, Coşkun F, Torun FM, Zhao P, Kang Y, Cevik S, Kaplan OI. 2022a. Convart: a search engine for matching human genetic variants with variants from non-human species. *Nucleic Acids Res*. 50(D1):D1172–D1178. doi:[10.1093/nar/gkab939](https://doi.org/10.1093/nar/gkab939).
- Pir MS, Cevik S, Kaplan OI. 2022b. Convart: search engine for missense variants between humans and other organisms. *Curr Protoc*. 2(11):e619. doi:[10.1002/cpz1.619](https://doi.org/10.1002/cpz1.619).
- Platzer K, Sticht H, Edwards SL, Allen W, Angione KM, Bonati MT, Brasington C, Cho MT, Demmer LA, Falik-Zaccari T, et al. 2019. De novo variants in MAPK8IP3 cause intellectual disability with variable brain anomalies. *Am J Hum Genet*. 104(2):203–212. doi:[10.1016/j.ajhg.2018.12.008](https://doi.org/10.1016/j.ajhg.2018.12.008).
- Reva B, Antipin Y, Sander C. 2011. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res*. 39(17):e118–e118. doi:[10.1093/nar/gkr407](https://doi.org/10.1093/nar/gkr407).
- Scheidel N, Blacque OE. 2018. Intraflagellar transport complex A genes differentially regulate cilium formation and transition zone gating. *Curr Biol*. 28(20):3279–3287.e2. doi:[10.1016/j.cub.2018.08.017](https://doi.org/10.1016/j.cub.2018.08.017).
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 9(7):676–682. doi:[10.1038/nmeth.2019](https://doi.org/10.1038/nmeth.2019).
- Schmidts M, Frank V, Eisenberger T, al Turki S, Bizet AA, Antony D, Rix S, Decker C, Bachmann N, Bald M, et al. 2013. Combined NGS approaches identify mutations in the intraflagellar transport gene IFT140 in skeletal ciliopathies with early progressive kidney disease. *Hum Mutat*. 34(5):714–724. doi:[10.1002/humu.22294](https://doi.org/10.1002/humu.22294).
- Schwarz JM, Cooper DN, Schuelke M, Seelow D. 2014. Mutationtaster2: mutation prediction for the deep-sequencing age. *Nat Methods*. 11(4):361–362. doi:[10.1038/nmeth.2890](https://doi.org/10.1038/nmeth.2890).
- Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GLA, Edwards KJ, Day INM, Gaunt TR. 2013. Predicting the functional, molecular,

- and phenotypic consequences of amino acid substitutions using hidden Markov models. *Hum Mutat.* 34(1):57–65. doi:[10.1002/humu.22225](https://doi.org/10.1002/humu.22225).
- Taliun D, Harris DN, Kessler MD, Carlson J, Szpiech ZA, Torres R, Taliun SAG, Corvelo A, Gogarten SM, Kang HM, et al. 2021. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed program. *Nature.* 590(7845):290–299. doi:[10.1038/s41586-021-03205-y](https://doi.org/10.1038/s41586-021-03205-y).
- The 1000 Genomes Project Consortium; Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. 2010. A map of human genome variation from population-scale sequencing. *Nature* 467(7319), 1061–1073. doi:[10.1038/nature09534](https://doi.org/10.1038/nature09534).
- The Alliance of Genome Resources Consortium; Agapite J, Albou L-P, Aleksander S, Argasinska J, Arnaboldi V, Attrill H, Bello SM, Blake JA, Blodgett O, et al. 2020. Alliance of genome resources portal: unified model organism research platform. *Nucleic Acids Res* 48(D1), D650–D658. doi:[10.1093/nar/gkz813](https://doi.org/10.1093/nar/gkz813).
- Thompson O, Edgley M, Strasbourger P, Flibotte S, Ewing B, Adair R, Au V, Chaudhry I, Fernando L, Hutter H, et al. 2013. The million mutation project: a new approach to genetics in *Caenorhabditis elegans*. *Genome Res.* 23(10):1749–1762. doi:[10.1101/gr.157651.113](https://doi.org/10.1101/gr.157651.113).
- Wang D, Dao M, Muntean BS, Giles AC, Martemyanov KA, Grill B. 2022. Genetic modeling of GNAO1 disorder delineates mechanisms of G α dysfunction. *Hum Mol Genet.* 31(4):510–522. doi:[10.1093/hmg/ddab235](https://doi.org/10.1093/hmg/ddab235).
- Wang J, Al-Ouran R, Hu Y, Kim S-Y, Wan Y-W, Wangler MF, Yamamoto S, Chao H-T, Comjean A, Mohr SE, et al. 2017. MARRVEL: integration of human and model organism genetic resources to facilitate functional annotation of the human genome. *Am J Hum Genet.* 100(6):843–853. doi:[10.1016/j.ajhg.2017.04.010](https://doi.org/10.1016/j.ajhg.2017.04.010).
- Winter DJ. 2017. Rentrez: an R package for the NCBI eUtils API. *R J.* 9(2):520. doi:[10.32614/RJ-2017-058](https://doi.org/10.32614/RJ-2017-058).
- Wong W-R, Brugman KI, Maher S, Oh JY, Howe K, Kato M, Sternberg PW. 2019. Autism-associated missense genetic variants impact locomotion and neurodevelopment in *Caenorhabditis elegans*. *Hum Mol Genet.* 28(13):2271–2281. doi:[10.1093/hmg/ddz051](https://doi.org/10.1093/hmg/ddz051).
- Xu M, Yang L, Wang F, Li H, Wang X, Wang W, Ge Z, Wang K, Zhao L, Li H, et al. 2015. Mutations in human IFT140 cause non-syndromic retinal degeneration. *Hum Genet.* 134(10):1069–1078. doi:[10.1007/s00439-015-1586-x](https://doi.org/10.1007/s00439-015-1586-x).
- Zhou L, Feng T, Xu S, Gao F, Lam TT, Wang Q, Wu T, Huang H, Zhan L, Li L, et al. 2022. ggmsa: a visual exploration tool for multiple sequence alignment and associated data. *Brief Bioinform.* 23(4):bbac222. doi:[10.1093/bib/bbac222](https://doi.org/10.1093/bib/bbac222).
- Zhu B, Mak JCH, Morris AP, Marson AG, Barclay JW, Sills GJ, Morgan A. 2020. Functional analysis of epilepsy-associated variants in STXBP1/Munc18-1 using humanized *Caenorhabditis elegans*. *Epilepsia.* 61(4):810–821. doi:[10.1111/epi.16464](https://doi.org/10.1111/epi.16464).

Editor: M.-A. Félix