Systematic creation and phenotyping of Mendelian disease models in *C. elegans*: towards large-scale drug repurposing

Thomas J. O'Brien[1,2*], Ida L. Barlow[1,2*], Luigi Feriani[1,2] and André E.X. Brown[1,2]

[1]Institute of Clinical Sciences, Imperial College London, London, UK
[2]MRC London Institute of Medical Sciences, London, UK
*These authors contributed equally
Correspondence: andre.brown@imperial.ac.uk

Abstract

There are thousands of Mendelian diseases with more being discovered weekly and the majority have no approved treatments. To address this need, we require scalable approaches that are relatively inexpensive compared to traditional drug development. In the absence of a validated drug target, phenotypic screening in model organisms provides a route for identifying candidate treatments. Success requires a screenable phenotype, however the right phenotype and assay may not be obvious for pleiotropic neuromuscular disorders. Here we show that high-throughput imaging and quantitative phenotyping can be conducted systematically on a panel of *C. elegans* disease model strains. We used CRISPR genome-editing to create 25 worm models of human Mendelian diseases and phenotyped them using a single standardised assay. All but two strains were significantly different from wild-type controls in at least one feature. The observed phenotypes were diverse, but mutations of genes predicted to have related functions in their human orthologs led to similar behavioural differences in worms. As a proof-of-concept, we performed a drug repurposing screen of an FDA approved compound library, and identified two compounds that rescued the behavioural phenotype of a model of UNC80 deficiency. Our results show that a single assay to measure multiple phenotypes can be applied systematically to diverse Mendelian disease models. The relatively short time and low cost associated with creating and phenotyping multiple strains suggests that high-throughput worm tracking could provide a scalable approach to drug repurposing commensurate with the number of Mendelian diseases.

Introduction

By definition, rare genetic diseases each affect a small number of patients. However, because there are many kinds of rare diseases, they are collectively common and affect more than 300 million people worldwide (Nguengang Wakap et al., 2020). Over 80% of rare genetic diseases affect children and many have severe symptoms (Bavisetty et al., 2013). Furthermore, 95% of rare diseases have no approved treatment (Nguengang Wakap et al., 2020). Given the severity and prevalence of rare diseases, the lack of treatments represents an important unmet need. Genome and exome sequencing has become faster and cheaper, enabling its widespread use in the diagnosis of genetic disease. In turn, this has led to the rapid identification of genetic lesions associated with rare genetic diseases (Gonzaga-Jauregui et al., 2012). Genetic diagnosis is an important first step, but often presents patients and physicians with a research problem rather than a treatment.

Genetic diagnoses rarely lead directly to a drug target hypothesis, especially for loss-of-function mutations. In the absence of a validated drug target, phenotypic screens in a model organism provide an alternative route to candidate treatments. Three things are required for a phenotypic screen to be successful: (1) a conserved gene; (2) a genetically tractable organism or culture system, in which the disease mutation can be introduced; and (3) a measurable phenotype compatible with high-

throughput screening. For the model nematode *C. elegans*, more than half of the human genes associated with rare genetic diseases in the OMIM database are conserved (Kropp et al., 2021) and CRISPR-based editing has made the generation of *in vivo* models containing disease-relevant mutations faster and easier. However, the third condition, a screenable phenotype, remains a bottleneck for many mutations and models.

If all three conditions are met, a phenotypic screen for candidate treatments can be performed. Because there is limited investment in drug development for rare diseases, lead optimisation and safety testing may not be possible. Hence, there is a particular focus on screening approved drugs that have already been shown to be safe and bioavailable in humans to reduce the cost and time it takes to translate screening hits to human trials (Roessler et al., 2021). Two recent examples of promising repurposing screens using multiple model organisms illustrate this potential.

Firstly, modelling amyotrophic lateral sclerosis (ALS) with disease-associated genetic variants of conserved genes in *C. elegans*, zebrafish and mice led to the discovery and clinical trials for repurposing of the antipsychotic drug, pimozide. By exploiting high throughput drug screening in *C. elegans* carrying point mutations in *TDP-43*, a panel of drugs were identified that rescued ALS movement and neurodevelopment phenotypes in all animal models. Early clinical trials found pimozide improved clinical and physiological outcomes in ALS patients (Patten et al., 2017). Secondly, the glycosylation disorder, PMM2-CDG, is caused by mutations in *PMM2* and presents clinically with developmental delay, psychomotor retardation and axial hypotonia, among other symptoms. *C. elegans* carrying the same disease-associated mutation in their endogenous *pmm-2* gene exhibited larval arrest upon pharmacological ER stress. *In vivo* screening identified two compounds that rescued development in *C. elegans* and enzymatic activity in PMM2-CDG patient fibroblasts (Iyer et al., 2019). One of the compounds, the aldose reductase inhibitor epalrestat, is currently being tested in clinical trials for PMM2-CDG patients (National Clinical Trial number: NCT04925960).

A screenable phenotype was critical for the success of these studies. In the ALS model, worms became paralysed after 2 hours in liquid and in the PMM2-CDG model, brood size was severely affected. The majority (~74%) of rare genetic diseases affect nervous system function (Lee et al., 2020), and many of these mutations will not cause strong effects on development or gross motility when mutated in worms. In this study, we use a combination of high-throughput imaging (Barlow et al., 2022) and quantitative phenotyping (Javer, Currie, et al., 2018; Javer, Ripoll-Sánchez, et al., 2018) to phenotype a diverse panel of 25 *C. elegans* strains with mutations in orthologs of human disease genes. All but two of the mutant strains had detectable phenotypes compared to controls in at least one aspect of their morphology, posture, or motion. No single feature was different in all strains highlighting the need for a multidimensional phenotype. Mutation of genes predicted to have similar functions in their human orthologs (e.g., *bbs-1*, *bbs-2,* and *tub-1*) led to similar phenotypes. As a proof-of-concept, we then performed a repurposing screen using a library of 743 FDA approved compounds to identify drugs that improved the behavioural phenotype of *unc-80* loss-of-function mutants. Liranaftate and atorvastatin rescued the core behavioural features associated with *unc-80* loss of function and did not cause a large number of detectable side effects.

The ability to detect phenotypic difference in diverse strains using a standardised 16-minute assay will make it possible to perform repurposing screens for existing and newly described rare diseases efficiently.

Results

Selection and molecular characterisation of a diverse panel of C. elegans disease models

The problem of large-scale disease modelling can be thought of as a problem of sampling and connecting two related genotype-phenotype (GP) maps (Fig. 1A). The first is the human disease map. The genomic space is sampled by mutation and mating in humans and the map is elucidated by sequencing and clinical phenotyping of patients with genetic diseases. The second is the model organism map. Here, the genomic space is sampled using mutagenesis and genome editing and the map is elucidated using laboratory experiments. Given the accelerated characterisation of the human GP map, we sought to accelerate the characterisation of the corresponding *C. elegans* GP map. We follow a phenolog-inspired approach to connecting the two maps. Provided there is a causal connection between a genetic variant and a disease in humans and that the causal gene is conserved in worms, the observed worm phenotype may be a useful disease model even if the connection to the human phenotype is non-obvious (McGary et al., 2010).

To identify a set of *C. elegans* genes that samples a broad genotypic space with connections to the human disease GP map, we first filtered the full database of human-*C. elegans* gene orthologs, Ortholist 2 (Kim et al., 2018), using three criteria: (1) at least two orthology prediction algorithms agree the human and worm genes are orthologs; (2) the WormBase (version WS270) (Harris et al., 2020) gene description includes either 'neuro' or 'musc' (this captures variants of neuronal, neural, muscle, muscular etc.); and (3) mutations in the human gene are associated with a Mendelian disease. A total of 767 *C. elegans* genes with 2558 orthologous human protein-coding genes met these criteria (Figure 1B). We then searched the list of high-confidence disease-associated orthologs using the following keyword filters to obtain 543 *C. elegans* genes well suited for the study of human Mendelian neuro-muscular diseases: neuron; neuronal; muscle; muscular; neurotransmission; ion; ion channel; dopamine; serotonin; glutamate; acetylcholine; 5-HT; behaviour; behaviour; disease; epilepsy; autism; Parkinson; schizophrenia; bipolar; ADHD; seizure; G protein coupled receptor; GPCR; coupled receptor; antidepressant; antipsychotic. In the present study, we prioritised genes associated with autosomal recessive disease in humans to increase the chance that a gene deletion would yield an appropriate model, and we focussed on those where loss-of-function mutants were known or likely to be viable. To get the list of 25 genes presented in this study, the final selection criterion was subjective interestingness based on our prior knowledge, WormBase gene descriptions, and/or brief literature searches.

In the final panel of 25 worm genes selected, 22 genes have >60% sequence similarity to their human ortholog and 11 of the genes share >90% sequence similarity to their human counterpart (Figure 1C). Furthermore, 24/25 genes are predicted to be orthologous to human genes across >3 orthology prediction algorithms (Figure 1C), and all selected genes have a BLAST E-value (McGinnis & Madden, 2004) smaller than $3x10^{-12}$ (Supplementary Table 1).

Large CRISPR-Cas9 deletions (mean 4.4 kb) were made in each of the target genes to generate 25 strains containing a >55% deletion (average 76%) of the chosen gene. We direct the reader to the strain-specific gene cards (Supplementary Information) for specific details on the size and position of genomic deletions for each of the individual mutants as well as a brief summary of the associated disease and worm phenotypes. Although we have focused on genes expected to have direct

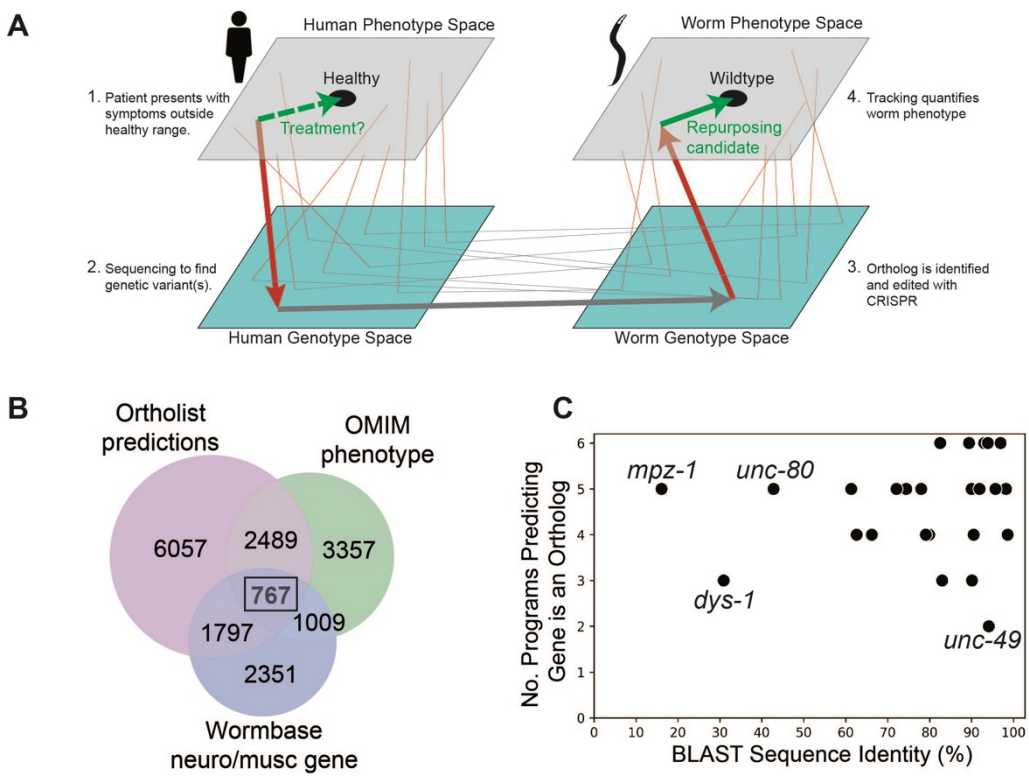

Figure 1: Overview of genotype-phenotype mapping and disease model panel. *Disease modelling as genotype-phenotype mapping in humans and a model organism. Arrows show progression from symptom identification (a phenotype outside the healthy range) to genotyping (placing a patient in genotype space by identifying a genetic variant), disease model creation (making a corresponding mutation in a model organism), and model organism phenotyping. Phenotypic drug screens identify compounds that move a disease model towards the wild-type phenotype (green arrow). These candidates can then be tested in humans (dashed green arrow). Thin lines show other symptom-gene-ortholog-phenotype connections.* ***(B)*** *Venn diagram showing number of conserved genes (Ortholist 2), those involved in neuron or muscle function (Wormbase), and those associated with human genetic disorders according to the Online Mendelian Inheritance in Man (OMIM) database.* ***(C)*** *Sequence similarity between human and C. elegans genes and the total number of orthology programs predicting that the gene is an ortholog.*

effects on neurons or muscles, even this small subset of disease-associated genes is predicted to affect diverse cellular processes including neurotransmission, excitability, development, and cellular structure (Table 1). According to OMIM, the mutated genes are associated with 31 rare genetic disorders in humans including intellectual disability, developmental delay, and disorders affecting the muscular and/or nervous systems, and are associated with >70 clinical presentations of disease according to the Human Disease Ontology database (Schriml et al., 2022) (Table 1).

Disease-associated mutations result in diverse phenotypes that are captured by multidimensional behavioural fingerprints

Given that there are thousands of rare diseases, methods to characterise the corresponding disease models should be as scalable as possible. Scalability is improved if a single standardised assay can be used to detect model phenotypes

| Worm Gene | Human Ortholog(s) | Functional Class | ADHD | Agenesis of the corpus callosum | Aggressive behaviour | Amyotrophic lateral sclerosis | Anxiety | Artery disease | Astigmatism | Ataxia | Autism | Bardet-Biedl syndrome | Becker muscular dystrophy | Bipolar disorder | Brachycephaly | Bradykinesia | Cerebral palsy | Charcot-Marie-Tooth disease | Contractures | Decreased nerve conduction velocity | Depression | Developmental delay | Diabetes | Duchenne muscular dystrophy | Dystonia | Encephalopathy | Epilepsy | Esotropia | Exotropia | Feeding difficulties | Fatigue | Fibrosis | Glaucoma | Hearing impairment | Hernia | Heterotopia | Huntington's Disease | Hydrocephalus | Hypertrophy | Hyperphenylalaninemia | Hyperreflexia | Hypertension | Hypertonia | Hypokinesia | Hypotonia | Intellectual disability | Joubert syndrome | Leber congenital amaurosis | Lennox-Gastaut syndrome | Macrocephaly | Meckel syndrome | Microcephaly | Movement abnormalities | Muscle atrophy | Myoclonus | Myopathy | Myopia | Neuropathy | Nystagmus | Obesity | Paralysis | Parkinson disease | Photophobia | Polydactyly | Polymicrogyria | Psychosis | Ptosis | Rheumatoid arthritis | Respiratory failure/distress | Retinal dystrophy | Retinitis pigmentosa | Schizophrenia | Scoliosis | Segawa syndrome | Seizures | Sheldon-Hall syndrome | Sleep disturbance | Spasticity | Speech disorder | Startle disorder | Strabismus | Syndactyly | Tremor | Visual impairment | Yunis-Varon syndrome |
|---|---|---

regardless of the mutation. High-resolution worm tracking and behavioural fingerprinting is a promising candidate for a widely applicable assay because it is both fast and multidimensional: a single minutes-long recording can capture differences in morphology, posture, and behaviour that can classify the effect of mutations and drug treatments (Baek et al., 2002; Barlow et al., 2022; Geng et al., 2004; Javer, Ripoll-Sánchez, et al., 2018; McDermott-Rouse et al., 2021; Perni et al., 2018; Ramot et al., 2008; Restif et al., 2014; Swierczek et al., 2011; Tsibidis & Tavernarakis, 2007; Wang & Wang, 2013; Yemini et al., 2013).

We tracked disease model worms on 96-well plates (Figure 2A) in three periods: (1) a 5-minute baseline recording (prestim); (2) a 6-minute video with three ten-second blue light stimuli separated by 90 seconds (blue light), and (3) a 5-minute post-stimulus recording (poststim). For each recording we extract 2763 features covering morphology, posture, and motion and concatenate the feature vectors together to represent the phenotype of each strain (8289 features in total). Many of these features are highly correlated and so for clustering we use a pre-selected subset of 256 features (Javer, Ripoll-Sánchez, et al., 2018) (768 across the three recordings). Using these reduced feature vectors, we performed hierarchical clustering across all strains which demonstrates phenotypic diversity (many of the strains are both different from the N2 wild-type strain and from each other) as well as some similarities. For example, *bbs-1, bbs-2* and *tub-1* mutants affect cilia formation and result in similar phenotypes that cluster together. The same is true for *cat-2*, *kcc*-2, *snf-11* and *unc-25* mutants, which all affect neurotransmitter biosynthesis, GABA transport or the regulation of GABAergic synaptic transmission. Two interactive heatmaps, one containing all behavioural feature vectors extracted for every disease model strain and one showing the Tierpsy256 feature set shown in Figure 2C, are also available for download (doi: zenodo.org/records/13911801).

Principal component analysis (PCA) of the same data, but now separated into prestimulus, blue light, and poststimlus recordings again shows the separation of many of the mutant strains from the wild-type N2 (Figure 2D). The effect of the blue light stimulus on behaviour is clear from the motion of the points through this phenotype space and also reveals new phenotypes. While most strains show a partial recovery in the post-stimulus recording, some strains including *bbs-1*, *bbs-2*, *glr-1*, and *unc-43* display a sustained photophobic response (Figure 2D).

From the panel of 25 disease model strains, 23 are significantly different from wild-type in at least one feature (Figure 2E). The strains that are most different from wild-type show significant differences across a large number of features. These strains are particularly useful for high-throughput drug screens because the phenotypes can be reliably detected with a small number of replicates. Strain-specific summaries (gene cards) of each disease model mutant, its associated phenotype, and individual molecular characteristics are available in the supplementary information. Here we provide more detailed characterisation of two classes of mutants modelling ciliopathies and channelopathies in humans.

Ciliopathies

Heritable mutations that affect cilia function are associated with a group of rare genetic disorders (ciliopathies) that share diverse symptoms including retinal dystrophy, developmental polydactyly, obesity, cognitive impairment, and renal dysfunction (Reiter & Leroux, 2017). The pleotropic nature of ciliopathies is highlighted by significant interfamilial and intrafamilial phenotypic variability (Shaheen et al., 2016), and the resulting poor gene-phenotype correlation significantly complicates the

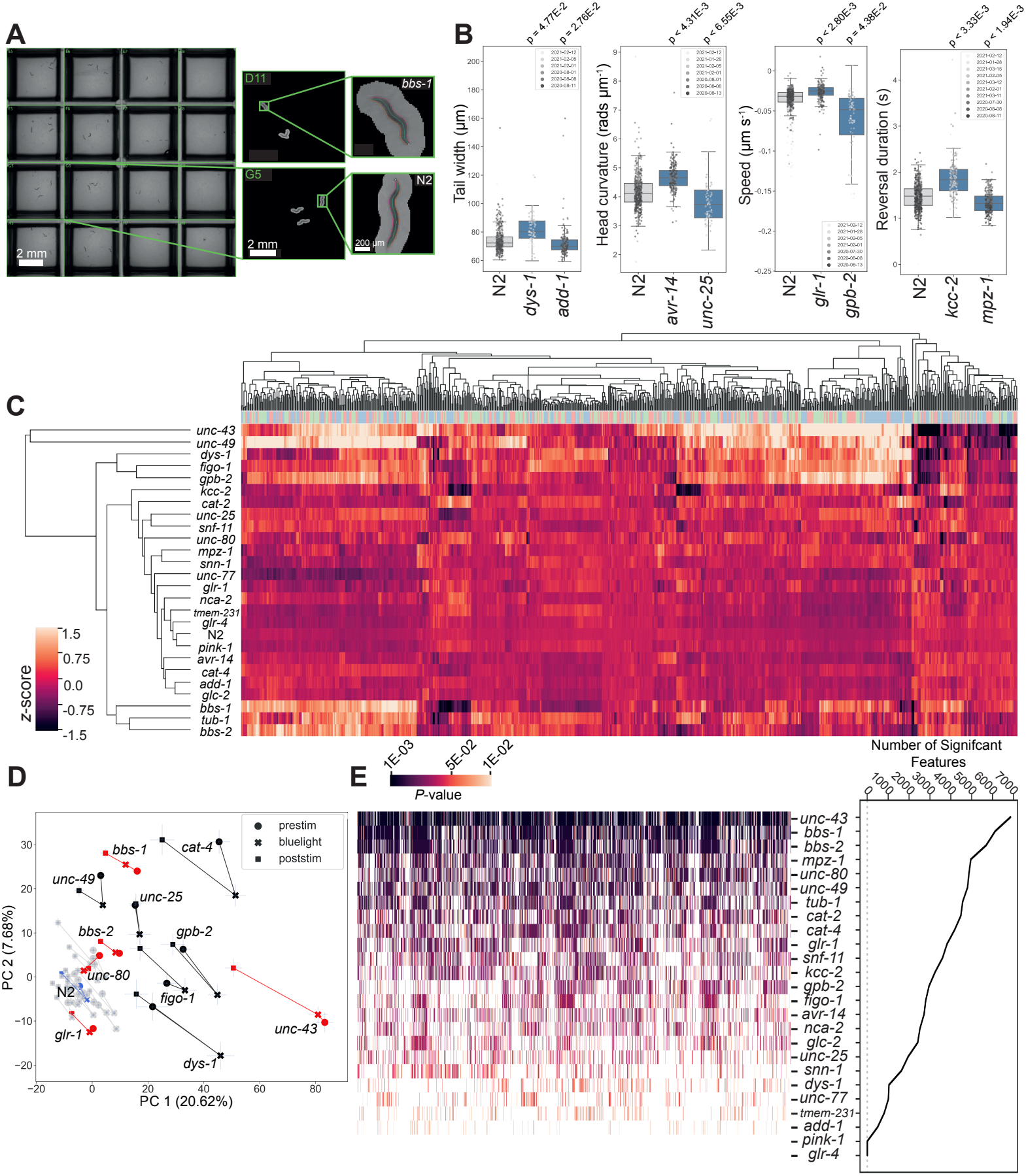

Figure 2: Diverse multidimensional behavioural phenotypes are obtained across the diverse panel of disease model mutants. **(A)** Representative field of view for a single camera channel with individual skeletonised worm images of *bbs-1(syb1588)* (top) and N2 (bottom) imaged on the same plate viewed using Tierpsy Tracker. **(B)** Representative behavioural phenotypes extracted by Tierpsy, representing changes in morphology, posture and locomotion of different disease mutant strains. *P*-values are for comparisons to wild-type N2 worms, block permutation t-test, corrected for multiple comparisons. **(C)** Hierarchical clustering of behavioural fingerprints. Features are *Z*-normalised. The top barcode shows the period of image acquisition where the behavioural feature was extracted by Tierpsy: pre-stimulation (pink), blue light (blue) and post-stimulation (green). **(D)** Principal component analysis of the disease model mutants and N2 reference (blue). Strains move in phenospace between pre-stimulus (circular points), blue light (crosses) and post-stimulation (squares) recordings. Strains with an aberrant blue light response are shown in red. Error bars represent standard error of the mean. **(E)** The total number of statistically significant behavioural features for each strain compared to N2 (from the total set of 8289 features extracted by Tierpsy). *P*-values for each feature were calculated using block permutation t-tests, using n = 100,000 permutations and $P < 0.05$ considered statistically significant after correcting for multiple comparisons using the Benjamini-Yekutieli method.

discovery of effective therapeutics. As a result, current treatment regimens are based on treating individual symptoms.

In the longlist of 543 disease genes, there were 35 genes (6.4% of the total list) predicted to be involved in cilia function. We selected four genes (*bbs-1*, *bbs-2*, *tmem-231* and *tub-1*) involved in regulating nonmotile (primary) cilia development and function that have confirmed expression in the cilia of amphid sensory neurons (Taylor et al., 2021; Ward et al., 2008; Zhang et al., 2022).

Of the genes affecting cilia function, two are orthologs of the Bardet-Biedl syndrome (BBS) family of genes (*bbs-1* and *bbs-2*). BBS is a pleiotropic syndrome often diagnosed in late childhood and is characterised by retinitis pigmentosa, cognitive impairment, obesity, renal dysfunction, hypogonadism, and polydactyly (Forsythe & Beales, 2013). The most prevalent BBS mutations are in *BBS1*, accounting for 23.4% of cases (Forsyth & Gunay-Aygun, 2023), and with 50% embryonic lethality and highly variable phenotypes, *Bbs1-/-* mouse models demonstrate model validity but have not yet yielded insights into potential therapies (Forsyth & Gunay-Aygun, 2023). Mutations in *BBS2* account for 8% of BBS cases (Forsythe & Beales, 2013) and are also implicated in other ciliopathies, such as Meckel syndrome (Karmous-Benailly et al., 2005).

Mutations in *bbs-1* and *bbs-2* led to strong phenotypes including changes in morphology, posture, and locomotion (Figure 3A-D). *bbs-1(syb1588)* and *bbs-2(syb1547)* mutants were shorter, wider and had decreased curvature compared to wild-type worms (Figures 3A, 3B and 3C). They are also hyperactive, with a greater fraction of worms moving during pre-stimulus (baseline) recordings (see individual gene cards), and had faster body bends (derivative of curvature) compared to N2 (Figure 3D and 3F). BBS proteins also regulate the *C. elegans* photoreceptor protein LITE1 in ASH photosensory neurons through a DLK-MAPK signalling pathway independent of their function in cilia (Zhang et al., 2022). Consistent with this, we find that *bbs-1(syb1588)* and *bbs-2(syb1547)* mutants have attenuated blue light sensitivity, with both strains exhibiting a delayed forward, but enhance backward photophobic escape response upon 10-second stimulation with blue light (Figures 3F-H).

The vertebrate family of tubby-like proteins (TUB, TULP) are required for correct GPCR localisation and intra-flagellar transport in primary cilia (Mukhopadhyay & Jackson, 2011). In humans, mutations in this family are associated with rod-cone dystrophy, obesity, and retinitis pigmentosa (North et al., 1997). Furthermore, *Tub1-/-* mice reflect the associated human disease pleiotropy, with insulin resistance, obesity, and cochlea and visual degeneration (Sun et al., 2012). *C. elegans* have a single ortholog of tubby-like proteins, *tub-1*, and mutants have defects in chemotaxis and insulin signalling as well as increased lipid accumulation (Mak et al., 2006).

tub-1(syb1562) mutants are shorter, wider, and hyperactive compared to wild-type worms, similar to *bbs-1* and *bbs-2* mutants (Figures 3A, 3B and 3E). They also have a defect in their response to blue light (Figures 3F-H). Unlike BBS mutants, cilia integrity is maintained in *tub-1* loss-of-function mutants (Mak et al., 2006), which could explain the similar but less severe phenotype of *tub-1* mutants compared to *bbs-1* and *bbs-2* mutants.

Compartmentalisation of cilia-specific signalling components is regulated by a complex of transmembrane proteins including TMEM231 and mutations in this compartmentalisation complex are associated with neurodevelopmental limb defects and pathologies of the brain and kidneys found in Joubert and Meckel syndrome (Chih et al., 2012). Mouse *Tmem231-/-* mutations are embryonic lethal and the *C. elegans*

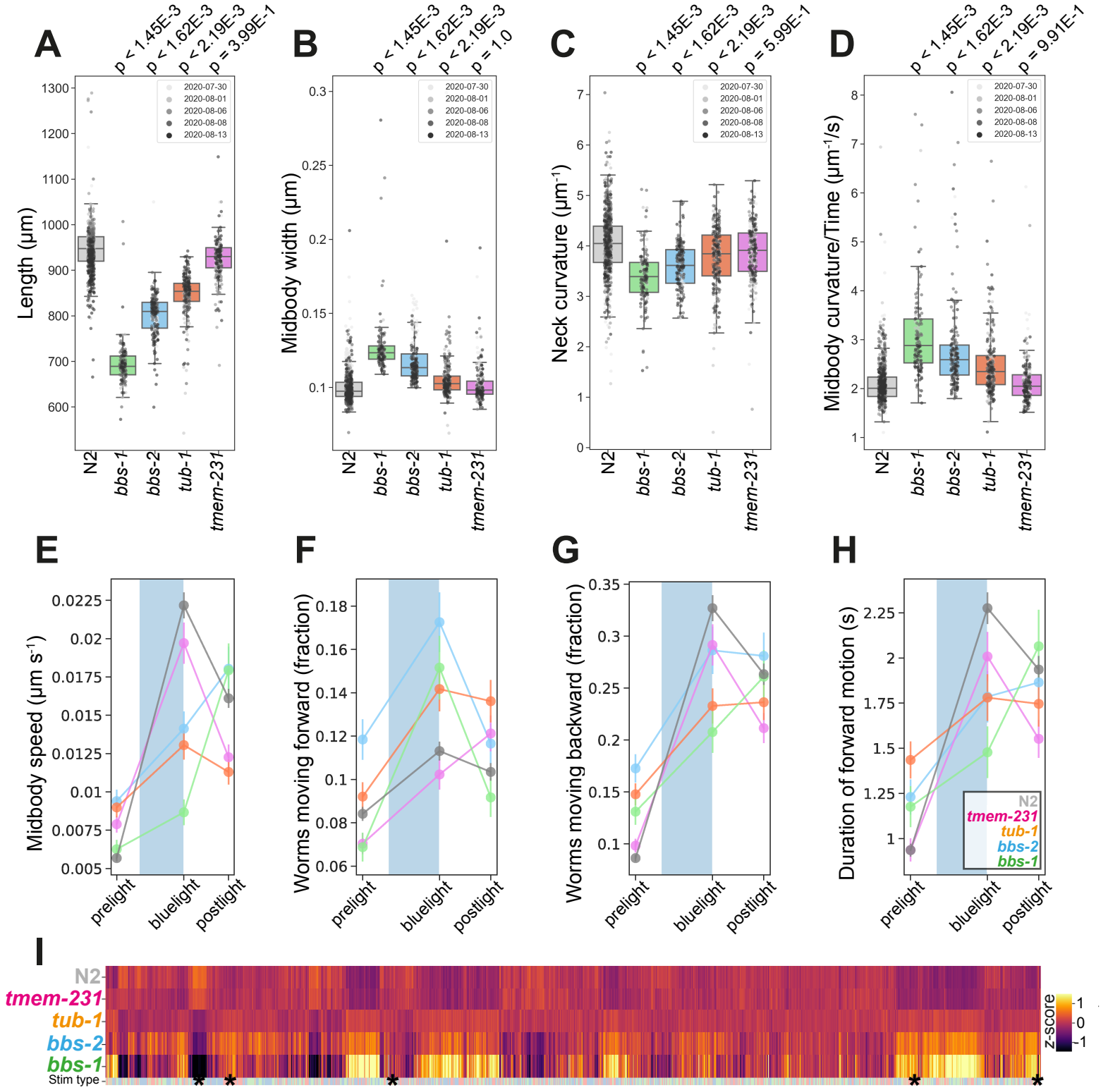

Figure 3: Ciliopathies disease model phenologs.
(A-D) Key behavioural features altered in loss-of-function mutant strains associated with ciliopathies: *bbs-1(syb1588)*, *bbs-2(syb1547)*, *tub-1(syb1562)* and *tmem-231(syb1575)*, under baseline (pre-stimulus) imaging conditions. Individual points marked on the box plots are well averaged values (three worms per well) for each feature across the independent days of tracking. *P*-values are for comparisons to wild-type N2 worms using block permutation t-tests (*n* = 100,000 permutations, correcting for multiple comparisons using the Benjamini-Yekutieli method). **(E-H)** Changes in selected features in response to stimulation with a single 10-second blue light pulse (blue shaded region). Feature values were calculated using 10 second windows centred around 5 seconds before, 10 seconds after, and 20 seconds after the beginning of each blue light pulse. **(I)** Heatmap of the entire set of 8289 behavioural features extracted by Tierpsy for the disease model strains associated with ciliopathies and N2. The 'stim type' barcode denotes when during image acquisition the feature was extracted: pre-stimulation (pink), blue light stimulation (blue) and post-stimulation (green). Asterisks show the location of selected features present in A-E.

ortholog *tmem-231* has been shown to have conserved molecular and cellular function (Roberson et al., 2015). However, unlike the other cilia-related mutants, *tmem-231* mutants do not have significant differences in body morphology or posture compared to N2 (Figure 3). We do note that *tmem-231(syb1575)* is slightly more active during baseline tracking (see strain-specific gene card).

Channelopathies

The Na+ cation leak channel (NALCN) is expressed throughout the central nervous system, and in parts of the endocrine (pancreas, adrenal, thyroid gland), respiratory and cardiac systems (Cochet-Bissuel et al., 2014). NALCN is a voltage-independent, nonselective cation channel that regulates resting potential (Senatore et al., 2013), and plays a role in neuromodulation by neurotransmitters (Cochet-Bissuel et al., 2014). Mutations in NALCN are associated with neuromuscular disorders including severe hypotonia, infantile neuroaxonal dystrophy (INAD), congenital contractures, cognitive delay, autism, epilepsy, bipolar disorder, and cardiac/respiratory problems (Aoyagi et al., 2015; Cochet-Bissuel et al., 2014; Gal et al., 2016). NALCN-knock-out mice die 12 hours after birth (Lu et al., 2007), so development of appropriate non-mammalian animal models to understand the associated human diseases is essential.

C. elegans encodes two functionally redundant, but differentially expressed, NALCN homologs, NCA-2 and UNC-77, whose proper expression and axonal localization in cholinergic neurons is regulated by UNC-80 (Zhou et al., 2019). Unlike murine models, *nca-2*, *unc-77*, and *unc-80* loss-of-function mutants are viable. Similar to *Drosophila melanogaster* or mouse neonates lacking the cation leak channel (Cochet-Bissuel et al., 2014), loss-of-function mutations in *nca-2* and *unc-80* result in morphological changes in *C. elegans*: both mutants are significantly shorter than wild-type N2 (Figure 4A).

Consistent with their differential expression (Jospin et al., 2007; Yeh et al., 2008), we observe phenotypic differences between *nca-2(syb1612)* and *unc-77(syb1688)* deletion mutants (Figure 4). Gain-of-function mutations in *unc-77* have previously been reported to cause deeper body bends (Topalidou et al., 2017) whereas *unc-77(syb1688)* deletion mutants have decreased curvature (Figure 4B). In contrast, *nca-2(syb1612)* has no significant change in curvature, but is slower than the wild-type strain (Figures 4B-C). We find that *unc-80(syb1531)* mutants, which affect the localisation of NALCN channel subunits, have the most severe phenotype exhibiting a decrease in both curvature and speed (Figures 4B-4C).

Mutations in both NALCN channel subunits in *C. elegans* cause a ‘fainter’ phenotype when stimulated mechanically or immersed in liquid (Pierce-Shimomura et al., 2008). Similarly, we found that a ten second blue light pulse resulted in increased post-exposure pausing and a decreased forward escape response in *nca-2(syb1612)* and *unc-80(syb1531)* single mutants (Figures 4E-H). *unc-77(syb1688)* mutants did not ‘faint’ after blue light exposure, but consistent with findings that the NALCN channelsome regulates reversal behaviour (Zhou et al., 2019) both *unc-77(syb1688)* and *unc-80(syb1531)* fail to initiate a backwards escape response. Pausing and ‘fainting’ after blue light exposure represents a novel screenable phenotype for NALCN-channelopathies.

Drug repurposing screen in unc-80 mutants

Because *unc-80* mutants have a clear behavioural phenotype without a strong developmental phenotype, we reasoned that they were a useful test case for a drug repurposing screen using an acute 4-hour treatment which might rescue behavioural

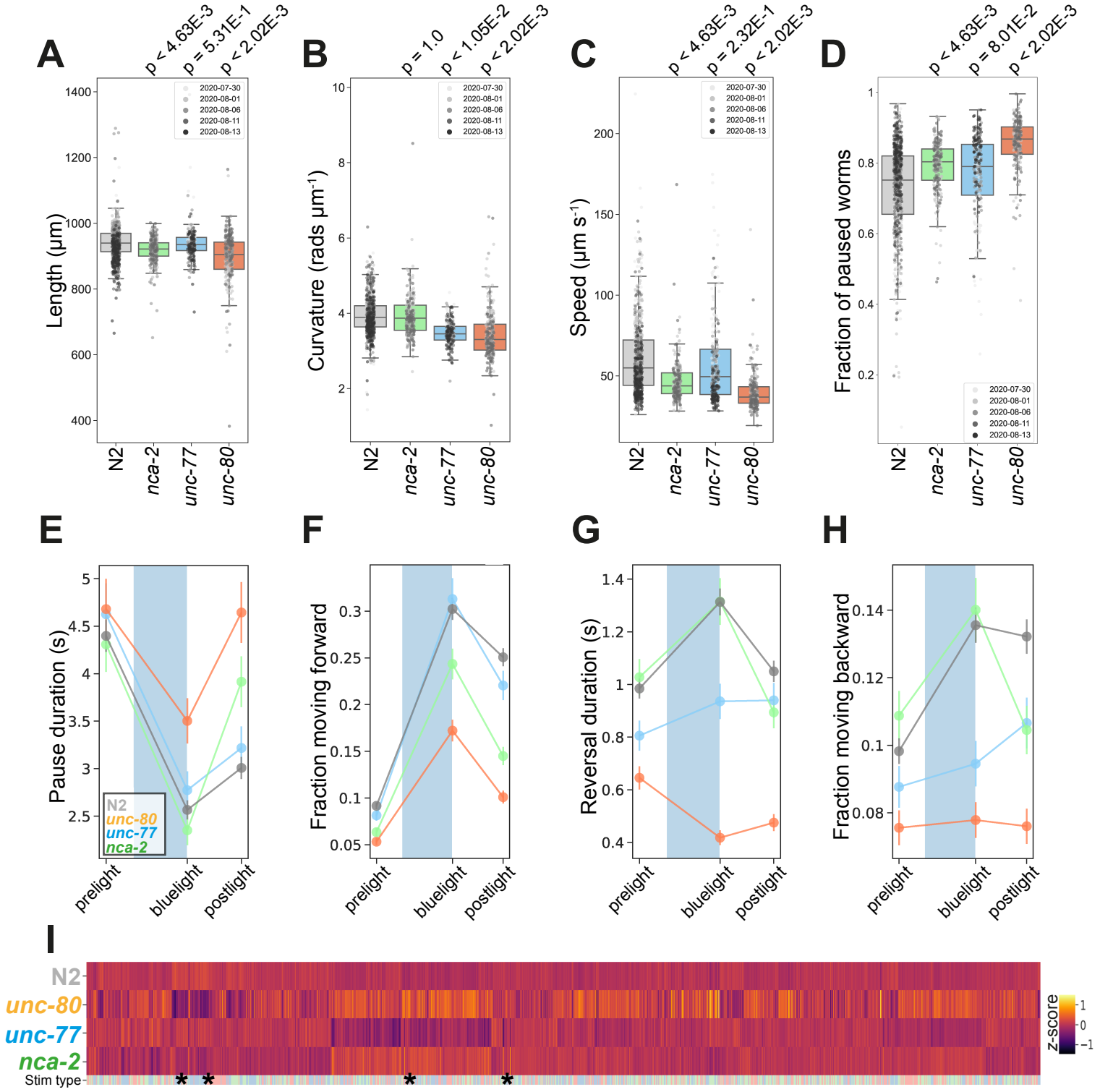

Figure 4: NALCN disease model phenologs.
(A-D) Key behavioural and postural features altered in loss-of-function mutant strains associated with NALCN mutants: *nca-2(syb1612)*, *unc-77(syb1688)* and *unc-80(syb1531)*, under baseline (pre-stimulus) imaging conditions. Individual points marked on the box plots are well averaged values (three worms per well) for each feature across the independent days of tracking. *P*-values are for comparisons to wild-type N2 worms using block permutation t-tests (*n* = 100,000 permutations correcting for multiple comparisons using the Benjamini-Yekutieli method). **(E-H)** Changes in selected features in response to stimulation with a single 10-second blue light pulse (blue shaded region). Feature values were calculated using 10 second windows centred around 5 seconds before, 10 seconds after, and 20 seconds after the beginning of each blue light pulse. **(E)** A representative 'fainting phenotype' for *unc-80(syb1531)* and *nca-2(syb1612)*, characterised by an increase in pausing following the cessation of stimulation with blue light. **(I)** Heatmap of the entire set of 8289 behavioural features extracted by Tierpsy for the disease model strains associated with NALCN disease and N2. The 'stim type' barcode denotes when during image acquisition the feature was extracted: pre-stimulation (pink), blue light stimulation (blue) and post-stimulation (green). Asterisks show the location of selected features present in A-D.

319 ‘symptoms’ in fully developed animals. When screening compound libraries, it is not
320 practical to perform a large number of replicates which would be necessary to detect
321 subtle phonotypes and overcome the reduced statistical power that comes from
322 correcting for multiple comparisons in a large feature set. We therefore defined a
323 reduced set of core features to capture the *unc-80* behavioural phenotype consisting
324 of curvature, speed, and fraction of paused worms all during blue light stimulation
325 (Figure 4B-D).
326 For the repurposing screen, we used a library of 743 FDA approved drugs. We
327 prepared tracking plates where each well contained a drug at a final concentration of
328 100 μM with 1% DMSO. Three adult worms were added to each well, incubated for 4
329 hours, and then tracked. The screen was repeated across three independent tracking
330 days with three independent wells per day (9 total well replicates). Hits were defined
331 as compounds that significantly improved all three of the core *unc-80* features,
332 meaning that treated worms were significantly different from *unc-80* DMSO controls
333 and that the direction of the effect was towards wild-type controls. These hit
334 compounds shift *unc-80* towards N2 in a three-dimensional phenotype space (Figure
335 5A). To empirically test our ability to detect hits when including more features in
336 subsequent screens, we also repeated this analysis using pre-defined Tierpsy feature
337 sets (Javer, Ripoll-Sánchez, et al., 2018) containing an increasing number of
338 behavioural features. We found a clear reduction in statistical power when correcting
339 for multiple comparisons across more behavioural features, with no hits being
340 detected using >16 behavioural features (Supplementary Figure 1).
341 We then performed a confirmation screen on 30 of the hit compounds from the
342 initial screen with a larger number of replicates (24 wells per day over three
343 independent days). 13 of the 30 hits were reproduced in the sense of showing
344 significant differences from *unc-80* DMSO controls in the direction of wild-type,
345 however several showed a reduced effect size and were clustered around the *unc-80*
346 DMSO controls (Figure 5B). Having a larger number of replicates, we also tested for
347 ‘side effects’ in treated worms which we define as features where there was no
348 significant difference between *unc-80* mutants and wild-type N2 animals but where
349 there is a significant difference between *unc-80* mutants treated with a drug compared
350 to wild-type N2 animals. About half of the confirmed hit compounds lead to more than
351 1000 side effects (Figure 5C) and are coloured yellow in the core *unc-80* phenotype
352 space (Figure 5B), we also note that the same trend persists when looking at a
353 reduced behavioural feature set of 256 predefined key Tierpsy features: the “Tierpsy
354 256” set (Javer, Ripoll-Sánchez, et al., 2018) (Supplementary Figure 2).
355 Of the confirmed hit compounds with fewer than 1000 side effects, liranaftate
356 and atorvastatin (Lipitor) lead to good rescue of the core *unc-80* behavioural
357 phenotypes. Interestingly, both compounds act upon the melavonate biosynthesis
358 pathway. The main trunk of this pathway is conserved across species and converts
359 acetyl-CoA to farnesyl diphosphate (Rauthan & Pilon, 2011). Atorvastatin is an
360 inhibitor of HMG-CoA reductase that regulates lipid homeostasis (Goldstein et al.,
361 2006) and catalyses the rate limiting step of the melavonate pathway. Liranaftate, an
362 antifungal inhibitor of squalene epoxidase, shows the greatest behavioural
363 improvement, with the fewest number of side effects. However, *C. elegans* lack the
364 branch of the mevalonate pathway responsible for converting squalene to cholesterol,
365 instead relying on dietary cholesterol. In this case the effect may be due to an off-
366 target interaction on one of the other enzymes in the pathway.

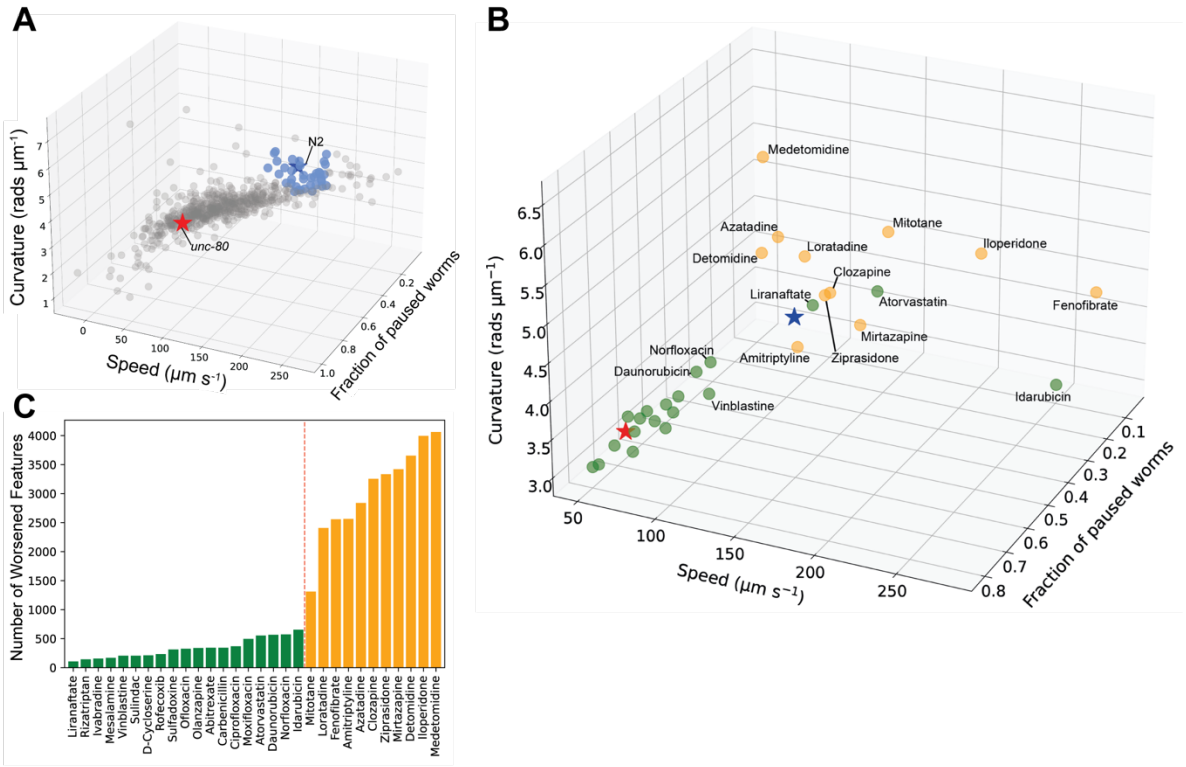

Figure 5: Drug repurposing screening

(A) *Phenotypes of unc-80(syb1531) mutant (red star) and N2 (blue star) worms treated with 1% DMSO, and unc-80(syb1531) mutants treated with a library of 743 FDA approved drugs at a concentration of 100 µM for 4 hours (circles). Each point represents an average of 3 well replicates, across 3 independent days of tracking (n = 9 total). Blue points are the top 30 compounds that significantly improved all three of the core behavioural features, pushing the unc-80 mutant strain towards the control in phenospace.* ***(B)*** *Confirmation screen of the top 30 compounds identified in the initial library screen. Again unc-80(syb1531) and N2 DMSO treated controls are represented by red and blue stars, respectively, and each circular point represents unc-80(syb1531) treated with 100 µM compound for 4 hours. The 13 compounds coloured in yellow lead to the worsening of >1000 behavioural features (see below). Liranaftate and atorvastatin both lead to a rescue of the core mutant phenotype with a low number of side effects.* ***(C)*** *Total number of behavioural 'side effects' following treatment of unc-80(syb1531) with the 30 compounds in the confirmation screen. Side effects are defined as features that are not significant between unc-80 mutants and wild-type N2 worms treated with 1% DMSO but where there is a significant difference between unc-80 mutants treated with a drug compared to N2. Red dashed line separates drug treatments that lead to a worsening of >1000 behavioural features that correspond to the points coloured in yellow in the 3D scatterplot.*

Fenofibrate, a PPAR-alpha activator that increases lipolysis, is also a hit but has a large number of side effects. Given that it overshoots perfect rescue by shifting *unc-80* animals past wild-type controls in core phenospace, it could be that at a lower dose it would provide good rescue with fewer side effects. Since UNC-80 is involved in localisation of the NACLN+ channel complex, we hypothesise that changes in

membrane composition might be indirectly affecting the localisation or function of channel complex.

Discussion

We framed the problem of disease modelling as one of sampling and elucidating a model organism genotype-phenotype map that is connected to a human disease gene-phenotype map through genetic conservation. However, both the genomic and phenomic spaces are large and high-dimensional and so it was not clear that a systematic approach would identify useful phenologs for disease modelling. Using an initial set of genes that are diverse in function (but biased towards roles in neurons and muscles) and that lead to diverse symptoms when mutated in humans, we found that it is possible to systematically create and phenotype worm models of rare diseases using a uniform assay and protocol. Without adding additional perturbations or experimental conditions, we found that 23 out of the 25 disease models we made had detectable differences compared to wild-type.

Of the disease models with detectable phenotypes, approximately half had strong phenotypes that would support high throughput phenotypic screens for candidate treatments. To test the feasibility of performing drug repurposing screens using the same uniform assay, we chose a worm model of UNC80 deficiency and, focussing on a reduced core phenotype, identified an initial list of 30 FDA-approved drugs that rescued the mutant phenotype. A confirmation screen with more replicates identified several compounds that still rescued the core disease model phenotype and did not have a large number of side effects. The detection of drug side effects will be useful for prioritising hits for further follow-up by identifying nuisance compounds. For example, if a disease model is hyperactive and gross motility is the only measure used in a screen, compounds that are broadly toxic may be detected as hits but not be good drug candidates.

Our treatment approach of a short exposure time (4h) was chosen to identify compounds with a potential to rescue neural phenotypes once the developmental effects of a mutation have already occurred, mimicking clinical intervention of a pre-existing genetic disorder. However, the flexibility afforded by our standardised phenotyping methodology ensures it is possible to test longer, or shorter, periods of drug treatment. For example, L1s could be reared on drug-containing plates for their entire life before phenotyping. We note that this approach will be particularly well suited to identify compounds that rescue developmental delayed strains, such as of some of the disease model mutants used in this study (e.g., *cat-4*, *dys-1*, *figo-1*, *gpb-2* and *kcc-2* LoF mutants).

With a current discovery rate of >100 new genetic diseases each year, the number of diseases with little aetiological understanding is going to grow at a faster rate than biomedical and pharmaceutical research will find therapies (Boycott et al., 2017). Therefore, novel approaches that can systematically screen the effects of large numbers of genetic variants associated with different diseases are required in order to scale up drug discovery across all rare diseases. CRISPR genome editing in model organisms including *C. elegans* allows the creation of new disease models at the required rate and cost. For this proof-of-principal study, we focused on the generation of large deletion alleles to ensure putative LoF of the target gene. However, we note that a limitation of this approach may be the removal of regulatory genetic elements, e.g., non-coding RNAs or introns, as part of the deletion. For future screening efforts, particularly those against a single/smaller panel of genes, we suggest the generation and testing of additional LoF mutants through targeting functional domains or the

423 introduction of early stop codons within a target allele. Better yet, the generation of
424 specifically engineered alleles that match to a corresponding human disease condition
425 will permit the generation of individual *C. elegans* "patient avatars" of diseases that
426 will enable the development and testing of personalised therapeutics against patient-
427 specific variants of genetic disorders. We have recently proven that our tracking
428 approach is capable of separating individual *sod-1* mutations, associated with the
429 onset of ALS in humans, in phenospace (Barlow et al., 2022). In summary, high-
430 throughput behaviour recording with multidimensional phenotyping brings an *in vivo*
431 drug repurposing screen within reach for every Mendelian disease caused by
432 mutations in a conserved gene.
433

434 ## Materials and Methods

435 ### *Mutant generation*

436 CRISPR guide RNAs were designed to target large deletions (>1000bp) that start
437 close to the start codon and excise several exons from the gene in order to give high
438 confidence of loss of function. For exceptionally large genes (e.g. *dys-1* or *mpz-1*) the
439 entire protein-coding region was excised. Mutants were designed and made by
440 SunyBiotech in an N2 background.
441

442 ### *Worm preparation*

443 All strains were cultured on Nematode Growth Medium at 20°C and fed with *E. coli*
444 (OP50) following standard procedure (Stiernagle, 2006). Synchronised populations of
445 young adult worms for imaging were cultured by bleaching unsynchronised gravid
446 adults, and allowing L1 diapause progeny to develop for 2.5 days at 20ºC (detailed
447 protocol: https://dx.doi.org/10.17504/protocols.io.2bzgap6). Several strains were
448 developmentally delayed and were allowed to grow for longer before imaging. *cat-*
449 *4*(*syb1591*), *gpb-2*(*syb1577*), *kcc-2*(*syb2673*), and *unc-25*(*syb1651*) were allowed to
450 develop for 3.5 and *dys-1*(*syb1688*), *figo-1*(*syb1562*), and *pink-1*(*syb1546*) were
451 allowed to develop for 5.5 days prior to imaging. On the day of imaging, young adults
452 were washed in M9 (detailed protocol:
453 https://dx.doi.org/10.17504/protocols.io.bfqbjmsn), transferred to the imaging plates
454 (3 worms per well) using a COPAS 500 Flow Pilot (detailed protocol:
455 https://dx.doi.org/10.17504/protocols.io.bfc9jiz6), and returned to a 20°C incubator for
456 3.5 hours. Plates were then transferred onto the multi-camera tracker for another 30
457 minutes to habituate prior to imaging (detailed protocol:
458 https://dx.doi.org/10.17504/protocols.io.bsicncaw).
459 For drug repurposing experiments, the MRCT FDA-approved compound library
460 (Catalog No.L1300) was supplied pre-dissolved in DMSO by LifeArc (Stevenage, UK).
461 The day prior to tracking, imaging plates were dosed with the compound library to
462 achieve a final well concentration of 100 µM prior to seeding with bacteria (see below
463 for details). Plates were left to dry (~30 mins), before being stored in the dark at room
464 temperature overnight. Following the methods described above, age synchronized
465 young adult *unc-80*(*syb1531*) worms were dispensed into the imaging plate wells and
466 incubated at 20°C for 4 hours before tracking. The behaviour of mutant worms dosed
467 with the drugs was then compared to wild-type N2 and *unc-80*(*syb1531*) worms (also
468 age-synchronized young adults) dispensed into the wells of the same tracking plates
469 dosed with an identical volume (1% w/v) of DMSO only (detailed protocol:
470 https://dx.doi.org/10.17504/protocols.io.5jyl8p5yrg2w/v1).
471

472 ### *Plate preparation*

Low peptone (0.013%) nematode growth medium (detailed protocol: https://dx.doi.org/10.17504/protocols.io.2rcgd2w) was prepared as follows: 20 g agar (Difco), 0.13 g Bactopeptone, and 3 g NaCl were dissolved in 975 mL of milliQ water. After autoclaving, 1 mL of 10 mg/mL cholesterol was added along with 1 mL CaCl2 (1 M), 1 mL MgSO4 (1 M) and 25 mL KPO4 buffer (1M, pH 6.0). Molten agar was cooled to 50-60°C and 200 µL was dispensed into each well of 96-square well plates (Whatman UNIPLATE: WHAT-77011651) using an Integra VIAFILL (detailed protocol: https://dx.doi.org/10.17504/protocols.io.bmxbk7in). Poured plates were stored agar-side up at 4°C until required.

One day prior to imaging, plates were placed without lids in a drying cabinet to lose 3-5% of weight by volume. Wells were then seeded with 5 µL OP50 (OD600 1.0) using an Integra VIAFILL dispenser, and plates were stored with lids (WHAT-77041001) on at room temperature overnight.

Image acquisition

All videos were acquired and processed following methods previously described (Barlow et al., 2022). In brief, videos were acquired at 25 frames per second in a room with a nominal temperature of 20°C using a shutter time of 25 ms and a resolution of 12.4 µm/px. Three videos were taken sequentially: a 5-minute pre-stimulus video, a 6-minute blue light recording with three 10-second blue light pulses starting at 60, 160 and 260 seconds, and a 5-minute post-stimulus recording. The script for controlling recording timings and photostimulation was made using LoopBio's API for their Motif software (https://github.com/loopbio/python-motifapi).

Image processing and feature extraction

Videos were segmented and tracked using Tierpsy Tracker (Javer, Currie, et al., 2018). After segmentation of worm skeletons, our previously described convolutional neural network classifier was used to exclude non-worm objects from being classified during feature extraction (Barlow et al., 2022). Skeletons that did not meet the following criteria were removed from the analysis: 700 – 1300 µM length, 20 – 200 µM width. As an additional quality control measure, we used Tierpsy Tracker's viewer to mark wells with visible contamination, agar damage, compound precipitation, or excess liquid as "bad", and exclude these wells from downstream analysis.

Following tracking, we extracted a previously-defined set of 3076 behavioural features for each well in each of the three videos (pre-stimulus, blue light, and post-stimulus) (Javer, Ripoll-Sánchez, et al., 2018). Feature values are averaged over tracks to produce a single feature vector for each well.

Statistical analysis

Statistically significant differences in the behavioural feature sets extracted by Tierpsy were calculated using block permutation t-tests in a pairwise manner between the mutant and/or treatment of interest and wild-type N2. Python (version 3.8.5) and the 'univariate_tests' function of Tierpsytools 'statistical_tests' module (https://github.com/Tierpsy/tierpsy-tools-python/blob/master/tierpsytools/analysis/statistical_tests.py) were used to perform these analysis. The *p*-value of each feature is calculated using $n = 10000$ permutations that randomly shuffle data labels within, but not between, the independent days of image acquisition for each strain/treatment to control for day-to-day variation in the technical replicates of each experiment. The resulting *p*-values were then corrected for multiple comparisons using the Benjamimi-Yekutieli procedure to control the false

discovery rate at 5% (Benjamini & Yekutieli, 2001) and $p > 0.05$ is considered not statistically significant.

Heatmaps and principal component analysis show the *z*-normalised values of the extracted behavioural feature sets compared to the N2 reference and were calculated using in-built Seaborn (version 0.11.2) packages (Waskom, 2021). All the scripts used for statistical analysis and the generation of figures are available at: (https://github.com/Tom-OBrien/Systematic-creation-and-phenotyping-of-Mendelian-disease-models-in-C.elegans).

Data Availability

The datasets (Tierpsy features, tracking data, and metadata) and code produced in this study are available on Zenodo, DOI: 10.5281/zenodo.12547170. Any remaining information can be obtained from the corresponding author upon reasonable request.

Acknowledgements

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 714853) and was supported by the Medical Research Council through grant MC-A658-5TY30.

References

Aoyagi, K., Rossignol, E., Hamdan, F. F., Mulcahy, B., Xie, L., Nagamatsu, S., Rouleau, G. A., Zhen, M., & Michaud, J. L. (2015). A Gain-of-Function Mutation in NALCN in a Child with Intellectual Disability, Ataxia, and Arthrogryposis. *Human Mutation*, *36*(8), 753–757. https://doi.org/10.1002/humu.22797

Baek, J.-H., Cosman, P., Feng, Z., Silver, J., & Schafer, W. R. (2002). Using machine vision to analyze and classify Caenorhabditis elegans behavioral phenotypes quantitatively. *Journal of Neuroscience Methods*, *118*, 9–21. https://doi.org/10.1016/S0165-0270(02)00117-6

Barlow, I. L., Feriani, L., Minga, E., McDermott-Rouse, A., O'Brien, T. J., Liu, Z., Hofbauer, M., Stowers, J. R., Andersen, E. C., Ding, S. S., & Brown, A. E. X. (2022). Megapixel camera arrays enable high-resolution animal tracking in multiwell plates. *Communications Biology*, *5*(1), Article 1. https://doi.org/10.1038/s42003-022-03206-1

Bavisetty, S., Grody, W. W., & Yazdani, S. (2013). Emergence of pediatric rare diseases: Review of present policies and opportunities for improvement. *Rare Diseases*, *1*. https://doi.org/10.4161/rdis.23579

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, *29*(4), 1165–1188. https://doi.org/10.1214/aos/1013699998

Boycott, K. M., Rath, A., Chong, J. X., Hartley, T., Alkuraya, F. S., Baynam, G., Brookes, A. J., Brudno, M., Carracedo, A., den Dunnen, J. T., Dyke, S. O. M., Estivill, X., Goldblatt, J., Gonthier, C., Groft, S. C., Gut, I., Hamosh, A., Hieter, P., Höhn, S., … Lochmüller, H. (2017). International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *The American Journal of Human Genetics*, *100*(5), 695–705. https://doi.org/10.1016/j.ajhg.2017.04.003

570 Chih, B., Liu, P., Chinn, Y., Chalouni, C., Komuves, L. G., Hass, P. E., Sandoval, W., & Peterson,
571 A. S. (2012). A ciliopathy complex at the transition zone protects the cilia as a
572 privileged membrane domain. *Nature Cell Biology*, *14*(1), Article 1.
573 https://doi.org/10.1038/ncb2410
574 Cochet-Bissuel, M., Lory, P., & Monteil, A. (2014). The sodium leak channel, NALCN, in
575 health and disease. *Frontiers in Cellular Neuroscience*, *8*.
576 https://www.frontiersin.org/article/10.3389/fncel.2014.00132
577 Forsyth, R., & Gunay-Aygun, M. (2023). Bardet-Biedl Syndrome Overview. In M. P. Adam, G.
578 M. Mirzaa, R. A. Pagon, S. E. Wallace, L. J. Bean, K. W. Gripp, & A. Amemiya (Eds.),
579 *GeneReviews®*. University of Washington, Seattle.
580 http://www.ncbi.nlm.nih.gov/books/NBK1363/
581 Forsythe, E., & Beales, P. L. (2013). Bardet–Biedl syndrome. *European Journal of Human*
582 *Genetics*, *21*(1), Article 1. https://doi.org/10.1038/ejhg.2012.115
583 Gal, M., Magen, D., Zahran, Y., Ravid, S., Eran, A., Khayat, M., Gafni, C., Levanon, E. Y., &
584 Mandel, H. (2016). A novel homozygous splice site mutation in NALCN identified in
585 siblings with cachexia, strabismus, severe intellectual disability, epilepsy and
586 abnormal respiratory rhythm. *European Journal of Medical Genetics*, *59*(4), 204–209.
587 https://doi.org/10.1016/j.ejmg.2016.02.007
588 Geng, W., Cosman, P., Berry, C. C., Feng, Z., & Schafer, W. R. (2004). Automatic tracking,
589 feature extraction and classification of C elegans phenotypes. *IEEE Transactions on*
590 *Bio-Medical Engineering*, *51*(10), 1811–1820.
591 https://doi.org/10.1109/TBME.2004.831532
592 Goldstein, J. L., DeBose-Boyd, R. A., & Brown, M. S. (2006). Protein sensors for membrane
593 sterols. *Cell*, *124*(1), 35–46. https://doi.org/10.1016/j.cell.2005.12.022

594 Gonzaga-Jauregui, C., Lupski, J. R., & Gibbs, R. A. (2012). Human genome sequencing in
595 health and disease. *Annual Review of Medicine*, *63*, 35–61.
596 https://doi.org/10.1146/annurev-med-051010-162644
597 Harris, T. W., Arnaboldi, V., Cain, S., Chan, J., Chen, W. J., Cho, J., Davis, P., Gao, S., Grove, C.
598 A., Kishore, R., Lee, R. Y. N., Muller, H.-M., Nakamura, C., Nuin, P., Paulini, M., Raciti,
599 D., Rodgers, F. H., Russell, M., Schindelman, G., … Sternberg, P. W. (2020).
600 WormBase: A modern Model Organism Information Resource. *Nucleic Acids*
601 *Research*, *48*(D1), D762–D767. https://doi.org/10.1093/nar/gkz920
602 Iyer, S., Sam, F. S., DiPrimio, N., Preston, G., Verheijen, J., Murthy, K., Parton, Z., Tsang, H.,
603 Lao, J., Morava, E., & Perlstein, E. O. (2019). Repurposing the aldose reductase
604 inhibitor and diabetic neuropathy drug epalrestat for the congenital disorder of
605 glycosylation PMM2-CDG. *Disease Models & Mechanisms*, *12*(11), dmm040584.
606 https://doi.org/10.1242/dmm.040584
607 Javer, A., Currie, M., Lee, C. W., Hokanson, J., Li, K., Martineau, C. N., Yemini, E., Grundy, L.
608 J., Li, C., Ch'ng, Q., Schafer, W. R., Nollen, E. A., Kerr, R., & Brown, A. E. (2018). An
609 open source platform for analyzing and sharing worm behavior data. *Nature*
610 *Methods*, *15*(9), 645–646. https://doi.org/10.1038/s41592-018-0112-1
611 Javer, A., Ripoll-Sánchez, L., & Brown, A. E. X. (2018). Powerful and interpretable
612 behavioural features for quantitative phenotyping of Caenorhabditis elegans.
613 *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1758),
614 20170375. https://doi.org/10.1098/rstb.2017.0375
615 Jospin, M., Watanabe, S., Joshi, D., Young, S., Hamming, K., Thacker, C., Snutch, T. P.,
616 Jorgensen, E. M., & Schuske, K. (2007). UNC-80 and the NCA ion channels contribute

617 to endocytosis defects in synaptojanin mutants. *Current Biology: CB*, *17*(18), 1595–
618 1600. https://doi.org/10.1016/j.cub.2007.08.036
619 Karmous-Benailly, H., Martinovic, J., Gubler, M.-C., Sirot, Y., Clech, L., Ozilou, C., Augé, J.,
620 Brahimi, N., Etchevers, H., Detrait, E., Esculpavit, C., Audollent, S., Goudefroye, G.,
621 Gonzales, M., Tantau, J., Loget, P., Joubert, M., Gaillard, D., Jeanne-Pasquier, C., …
622 Attié-Bitach, T. (2005). Antenatal Presentation of Bardet-Biedl Syndrome May Mimic
623 Meckel Syndrome. *American Journal of Human Genetics*, *76*(3), 493–504.
624 Kim, W., Underwood, R. S., Greenwald, I., & Shaye, D. D. (2018). OrthoList 2: A New
625 Comparative Genomic Analysis of Human and Caenorhabditis elegans Genes.
626 *Genetics*, *210*(2), 445–461. https://doi.org/10.1534/genetics.118.301307
627 Kropp, P. A., Bauer, R., Zafra, I., Graham, C., & Golden, A. (2021). Caenorhabditis elegans for
628 rare disease modeling and drug discovery: Strategies and strengths. *Disease Models*
629 *& Mechanisms*, *14*(8), dmm049010. https://doi.org/10.1242/dmm.049010
630 Lee, C. E., Singleton, K. S., Wallin, M., & Faundez, V. (2020). Rare Genetic Diseases: Nature's
631 Experiments on Human Development. *iScience*, *23*(5), 101123.
632 https://doi.org/10.1016/j.isci.2020.101123
633 Lu, B., Su, Y., Das, S., Liu, J., Xia, J., & Ren, D. (2007). The neuronal channel NALCN
634 contributes resting sodium permeability and is required for normal respiratory
635 rhythm. *Cell*, *129*(2), 371–383. https://doi.org/10.1016/j.cell.2007.02.041
636 Mak, H. Y., Nelson, L. S., Basson, M., Johnson, C. D., & Ruvkun, G. (2006). Polygenic control
637 of Caenorhabditis elegans fat storage. *Nature Genetics*, *38*(3), Article 3.
638 https://doi.org/10.1038/ng1739
639 McDermott-Rouse, A., Minga, E., Barlow, I., Feriani, L., Harlow, P. H., Flemming, A. J., &
640 Brown, A. E. X. (2021). Behavioral fingerprints predict insecticide and anthelmintic

641 mode of action. *Molecular Systems Biology*, *17*(5), e10267.
642 https://doi.org/10.15252/msb.202110267
643 McGary, K. L., Park, T. J., Woods, J. O., Cha, H. J., Wallingford, J. B., & Marcotte, E. M. (2010).
644 Systematic discovery of nonobvious human disease models through orthologous
645 phenotypes. *Proceedings of the National Academy of Sciences*, *107*(14), 6544–6549.
646 https://doi.org/10.1073/pnas.0910200107
647 McGinnis, S., & Madden, T. L. (2004). BLAST: At the core of a powerful and diverse set of
648 sequence analysis tools. *Nucleic Acids Research*, *32*(Web Server issue), W20–W25.
649 https://doi.org/10.1093/nar/gkh435
650 Mukhopadhyay, S., & Jackson, P. K. (2011). The tubby family proteins. *Genome Biology*,
651 *12*(6), 225. https://doi.org/10.1186/gb-2011-12-6-225
652 Nguengang Wakap, S., Lambert, D. M., Olry, A., Rodwell, C., Gueydan, C., Lanneau, V.,
653 Murphy, D., Le Cam, Y., & Rath, A. (2020). Estimating cumulative point prevalence of
654 rare diseases: Analysis of the Orphanet database. *European Journal of Human*
655 *Genetics*, *28*(2), Article 2. https://doi.org/10.1038/s41431-019-0508-0
656 North, M. A., Naggert, J. K., Yan, Y., Noben-Trauth, K., & Nishina, P. M. (1997). Molecular
657 characterization of TUB, TULP1, and TULP2, members of the novel tubby gene family
658 and their possible relation to ocular diseases. *Proceedings of the National Academy*
659 *of Sciences of the United States of America*, *94*(7), 3128–3133.
660 Patten, S. A., Aggad, D., Martinez, J., Tremblay, E., Petrillo, J., Armstrong, G. A. B., Fontaine,
661 A. L., Maios, C., Liao, M., Ciura, S., Wen, X.-Y., Rafuse, V., Ichida, J., Zinman, L., Julien,
662 J.-P., Kabashi, E., Robitaille, R., Korngut, L., Parker, J. A., & Drapeau, P. (2017).
663 Neuroleptics as therapeutic compounds stabilizing neuromuscular transmission in

664 amyotrophic lateral sclerosis. *JCI Insight*, *2*(22).
665 https://doi.org/10.1172/jci.insight.97152
666 Perni, M., Challa, P. K., Kirkegaard, J. B., Limbocker, R., Koopman, M., Hardenberg, M. C.,
667 Sormanni, P., Müller, T., Saar, K. L., Roode, L. W. Y., Habchi, J., Vecchi, G., Fernando,
668 N. W., Casford, S., Nollen, E. A. A., Vendruscolo, M., Dobson, C. M., & Knowles, T. P.
669 J. (2018). Massively parallel C. elegans tracking provides multi-dimensional
670 fingerprints for phenotypic discovery. *Journal of Neuroscience Methods*.
671 https://doi.org/10.1016/j.jneumeth.2018.02.005
672 Pierce-Shimomura, J. T., Chen, B. L., Mun, J. J., Ho, R., Sarkis, R., & McIntire, S. L. (2008).
673 Genetic analysis of crawling and swimming locomotory patterns in C. elegans.
674 *Proceedings of the National Academy of Sciences of the United States of America*,
675 *105*(52), 20982–20987. https://doi.org/10.1073/pnas.0810359105
676 Ramot, D., Johnson, B. E., Berry, T. L., Carnell, L., & Goodman, M. B. (2008). The Parallel
677 Worm Tracker: A Platform for Measuring Average Speed and Drug-Induced Paralysis
678 in Nematodes. *PLoS ONE*, *3*, e2208. https://doi.org/10.1371/journal.pone.0002208
679 Rauthan, M., & Pilon, M. (2011). The mevalonate pathway in C. Elegans. *Lipids in Health and*
680 *Disease*, *10*, 243. https://doi.org/10.1186/1476-511X-10-243
681 Reiter, J. F., & Leroux, M. R. (2017). Genes and molecular pathways underpinning
682 ciliopathies. *Nature Reviews Molecular Cell Biology*, *18*(9), Article 9.
683 https://doi.org/10.1038/nrm.2017.60
684 Restif, C., Ibáñez-Ventoso, C., Vora, M. M., Guo, S., Metaxas, D., & Driscoll, M. (2014).
685 CeleST: Computer vision software for quantitative analysis of C. elegans swim
686 behavior reveals novel features of locomotion. *PLoS Computational Biology*, *10*(7),
687 e1003702. https://doi.org/10.1371/journal.pcbi.1003702

688 Roberson, E. C., Dowdle, W. E., Ozanturk, A., Garcia-Gonzalo, F. R., Li, C., Halbritter, J.,
689 Elkhartoufi, N., Porath, J. D., Cope, H., Ashley-Koch, A., Gregory, S., Thomas, S.,
690 Sayer, J. A., Saunier, S., Otto, E. A., Katsanis, N., Davis, E. E., Attié-Bitach, T.,
691 Hildebrandt, F., … Reiter, J. F. (2015). TMEM231, mutated in orofaciodigital and
692 Meckel syndromes, organizes the ciliary transition zone. *The Journal of Cell Biology*,
693 *209*(1), 129–142. https://doi.org/10.1083/jcb.201411087
694 Roessler, H. I., Knoers, N. V. A. M., van Haelst, M. M., & van Haaften, G. (2021). Drug
695 Repurposing for Rare Diseases. *Trends in Pharmacological Sciences*, *42*(4), 255–267.
696 https://doi.org/10.1016/j.tips.2021.01.003
697 Schriml, L. M., Munro, J. B., Schor, M., Olley, D., McCracken, C., Felix, V., Baron, J. A.,
698 Jackson, R., Bello, S. M., Bearer, C., Lichenstein, R., Bisordi, K., Dialo, N. C., Giglio, M.,
699 & Greene, C. (2022). The Human Disease Ontology 2022 update. *Nucleic Acids*
700 *Research*, *50*(D1), D1255–D1261. https://doi.org/10.1093/nar/gkab1063
701 Senatore, A., Monteil, A., Minnen, J. van, Smit, A. B., & Spafford, J. D. (2013). NALCN Ion
702 Channels Have Alternative Selectivity Filters Resembling Calcium Channels or Sodium
703 Channels. *PLOS ONE*, *8*(1), e55088. https://doi.org/10.1371/journal.pone.0055088
704 Shaheen, R., Szymanska, K., Basu, B., Patel, N., Ewida, N., Faqeih, E., Al Hashem, A., Derar,
705 N., Alsharif, H., Aldahmesh, M. A., Alazami, A. M., Hashem, M., Ibrahim, N.,
706 Abdulwahab, F. M., Sonbul, R., Alkuraya, H., Alnemer, M., Al Tala, S., Al-Husain, M., …
707 Alkuraya, F. S. (2016). Characterizing the morbid genome of ciliopathies. *Genome*
708 *Biology*, *17*, 242. https://doi.org/10.1186/s13059-016-1099-5
709 Stiernagle, T. (2006). Maintenance of C. elegans. *WormBook: The Online Review of C.*
710 *Elegans Biology*, 1–11. https://doi.org/10.1895/wormbook.1.101.1

711 Sun, X., Haley, J., Bulgakov, O. V., Cai, X., McGinnis, J., & Li, T. (2012). Tubby is required for
712 trafficking G protein-coupled receptors to neuronal cilia. *Cilia*, *1*(1), 21.
713 https://doi.org/10.1186/2046-2530-1-21
714 Swierczek, N. A., Giles, A. C., Rankin, C. H., & Kerr, R. A. (2011). High-throughput behavioral
715 analysis in C. elegans. *Nature Methods*, *8*(7), 592–598.
716 https://doi.org/10.1038/nmeth.1625
717 Taylor, S. R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M. B., Xu, C., Varol, E.,
718 Oikonomou, P., Glenwinkel, L., McWhirter, R., Poff, A., Basavaraju, M., Rafi, I.,
719 Yemini, E., Cook, S. J., Abrams, A., Vidal, B., Cros, C., Tavazoie, S., … Miller, D. M.
720 (2021). Molecular topography of an entire nervous system. *Cell*, *184*(16), 4329-
721 4347.e23. https://doi.org/10.1016/j.cell.2021.06.023
722 Topalidou, I., Cooper, K., Pereira, L., & Ailion, M. (2017). Dopamine negatively modulates
723 the NCA ion channels in C. elegans. *PLoS Genetics*, *13*(10), e1007032.
724 https://doi.org/10.1371/journal.pgen.1007032
725 Tsibidis, G. D., & Tavernarakis, N. (2007). Nemo: A computational tool for analyzing
726 nematode locomotion. *BMC Neuroscience*, *8*, 86. https://doi.org/10.1186/1471-
727 2202-8-86
728 Wang, S. J., & Wang, Z.-W. (2013). Track-A-Worm, An Open-Source System for Quantitative
729 Assessment of C. elegans Locomotory and Bending Behavior. *PLoS ONE*, *8*(7),
730 e69653. https://doi.org/10.1371/journal.pone.0069653
731 Ward, A., Liu, J., Feng, Z., & Xu, X. Z. S. (2008). Light-sensitive neurons and channels mediate
732 phototaxis in C. elegans. *Nature Neuroscience*, *11*(8), Article 8.
733 https://doi.org/10.1038/nn.2155

734 Waskom, M. L. (2021). seaborn: Statistical data visualization. *Journal of Open Source*
735 *Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021
736 Yeh, E., Ng, S., Zhang, M., Bouhours, M., Wang, Y., Wang, M., Hung, W., Aoyagi, K., Melnik-
737 Martinez, K., Li, M., Liu, F., Schafer, W. R., & Zhen, M. (2008). A putative cation
738 channel, NCA-1, and a novel protein, UNC-80, transmit neuronal activity in C.
739 elegans. *PLoS Biology*, *6*(3), e55. https://doi.org/10.1371/journal.pbio.0060055
740 Yemini, E., Jucikas, T., Grundy, L. J., Brown, A. E. X., & Schafer, W. R. (2013). A database of
741 Caenorhabditis elegans behavioral phenotypes. *Nature Methods*, *10*(9), 877–879.
742 https://doi.org/10.1038/nmeth.2560
743 Zhang, X., Liu, J., Pan, T., Ward, A., Liu, J., & Xu, X. Z. S. (2022). A cilia-independent function
744 of BBSome mediated by DLK-MAPK signaling in C. elegans photosensation.
745 *Developmental Cell*, *57*(12), 1545-1557.e4.
746 https://doi.org/10.1016/j.devcel.2022.05.005
747 Zhou, C., Luo, J., He, X., Zhou, Q., He, Y., Wang, X., & Ma, L. (2019). The NALCN Channel
748 Regulator UNC-80 Functions in a Subset of Interneurons To Regulate Caenorhabditis
749 elegans Reversal Behavior. *G3: Genes|Genomes|Genetics*, *10*(1), 199–210.
750 https://doi.org/10.1534/g3.119.400692
751