# Title: 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

Authors: Troy A. McDiarmid[1], Vinci Au[2], Aaron D. Loewen[1], Joseph Liang[1], Kota Mizumoto[2], Donald G. Moerman[2], and Catharine H. Rankin[1,3]

1. Djavad Mowafaghian Centre for Brain Health, University of British Columbia, 2211 Wesbrook Mall, Vancouver, British Columbia V6T 2B5, Canada
2. Department of Zoology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada.
3. Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, British Columbia V6T 1Z4, Canada

Correspondence concerning this article should be addressed to:
Catharine H. Rankin
Djavad Mowafaghian Center for
Brain Health and Department of Psychology
University of British Columbia
2136 West Mall,
Vancouver, British Columbia
Canada, V6T 1Z4
email: crankin@psych.ubc.ca
phone: 604-822-5449
fax: 604-822-7299
ORCID iD: 0000-0002-1781-0654

Keywords:

CRISPR-Cas9, genome editing, *Caenorhabditis elegans*, humanization, targeted knock-in, *PTEN*, *daf-18*, gene replacement, phenomics, behavioral phenotyping, variants of uncertain significance

**Abstract**

Our ability to sequence genomes has vastly surpassed our ability to interpret the genetic variation we discover. This presents a major challenge in the clinical setting, where the recent application of whole exome and whole genome sequencing has uncovered thousands of genetic variants of uncertain significance. Here, we present a strategy for targeted human gene replacement and phenomic characterization based on CRISPR-Cas9 genome engineering in the genetic model organism *Caenorhabditis elegans* that will facilitate assessment of the functional conservation of human genes and structure-function analysis of disease-associated variants with unprecedented precision. We validate our strategy by demonstrating that direct single-copy replacement of the *C. elegans* ortholog (*daf-18*) with the critical human disease-associated gene Phosphatase and Tensin Homolog (*PTEN*) is sufficient to rescue multiple phenotypic abnormalities caused by complete deletion of *daf-18*, including complex chemosensory and mechanosenory impairments. In addition, we used our strategy to generate animals harboring a single copy of the known pathogenic lipid phosphatase inactive PTEN variant (PTEN-G129E) and showed that our automated *in vivo* phenotypic assays could accurately and efficiently classify this missense variant as loss-of-function. The integrated nature of the human transgenes allows for analysis of both homozygous and heterozygous variants and greatly facilitates high-throughput precision medicine drug screens. By combining genome engineering with rapid and automated phenotypic characterization, our strategy streamlines identification of novel conserved gene functions in complex sensory and learning phenotypes that can be used as *in vivo* functional assays to decipher variants of uncertain significance.

## Introduction

The rapid development and application of whole exome and whole genome sequencing technology has dramatically increased the pace at which we associate genetic variation with a particular disease (Auton et al., 2015; Bamshad et al., 2011; Gonzaga-Jauregui et al., 2012; Lek et al., 2016; Metzker, 2010; Need et al., 2012; Ng et al., 2010). However, our ability to sequence genomes has vastly surpassed our ability to interpret the clinical implications of the genetic variants we discover. The majority of genetic variants identified in clinical populations are currently classified as "variants of uncertain significance" meaning their potential role as a causative agent in the disease in question, or their pathogenicity, is unknown (Richards et al., 2015). Many variants are exceedingly rare, making it extremely difficult to designate them as pathogenic using classical genetic methods such as segregation within a pedigree or by identifying multiple carriers of the variant. As such, it often remains challenging to predict clinical outcomes and make informed treatment decisions based on genetic data alone.

In an attempt to address this problem several computational tools have been developed that estimate the functional consequences and pathogenicity of disease-associated variants (Richards et al., 2015). These tools use a variety of predictive features such as evolutionary sequence conservation, protein structural and functional information, the prevalence of a variant in large putatively healthy control populations, or a combination of annotations (Kircher et al., 2014; Lek et al., 2016; Richards et al., 2015). Despite extensive efforts none of these tools used in isolation or combination can faithfully report on the functional effects of a large portion of disease-associated variation and their accuracy is intrinsically limited to existing experimental training data (Grimm et

al., 2015; Miosge et al., 2015; Starita et al., 2017). These limitations were clearly demonstrated in a recent study that showed *in vivo* functional assays of 21 human genes in yeast identified pathogenic variants with significantly higher precision and specificity than computational methods (Sun et al., 2016). This means that even for genes with well-characterized biological functions there are often hundreds of variants of uncertain functional significance (Landrum et al., 2014; Starita et al., 2017). This creates a challenging situation that requires direct assessment of the functional effects of disease-associated variants *in vivo* (Starita et al., 2017).

Genetically tractable model organisms are critical for discovering novel gene functions and the functional consequences of disease-associated genetic variants (Dunham and Fowler, 2013; Lehner, 2013; Manolio et al., 2017; Wangler et al., 2017). Governmental and private funding agencies are increasingly commissioning large-scale collaborative programs to use diverse genetic model organisms to decipher variants of uncertain significance (Chong et al., 2015; Gahl et al., 2016; Wangler et al., 2017). Among genetic model organisms, the nematode *Caenorhabditis elegans* has proven to be a particularly powerful animal model for the functional characterization of human genes *in vivo* (Kaletta and Hengartner, 2006). *C. elegans* is an ideal genetic model as it combines the throughput and tractability of a single-celled organism with the complex morphology and behavioral repertoire of a multicellular animal. In addition, approximately 60-80% of human genes have an ortholog in the *C. elegans* genome (Kaletta and Hengartner, 2006; Lai et al., 2000a; Shaye and Greenwald, 2011). Transgenic expression of human genes is routinely done to confirm functional conservation and to observe the effects of disease-associated mutations. Notable

examples of the utility of *C. elegans* to determine conserved human gene functions relevant to disease include the identification of presenilins as part of the gamma secretase complex, the mechanism of action of selective serotonin reuptake inhibitors, and the role of the insulin signaling pathway in normal and pathological ageing (Kaletta and Hengartner, 2006; Levitan et al., 1996; Levitan and Greenwald, 1995; Murphy et al., 2003; Ranganathan et al., 2001). However, traditional methods for expression of human genes in *C. elegans* rely on mosaic and variable over-expression of transgenes harbored as extrachromosomal arrays or specialized genetic backgrounds that can confound phenotypic analysis. This presents several challenges that inhibit precise analysis of the often critical but subtle effects of missense variants and impede the use of these transgenic strains in large-scale drug screens.

The recent advent of CRISPR-based genome editing has revolutionized structure-function analyses across model organisms (Cong et al., 2013; DiCarlo et al., 2013; Dickinson et al., 2013; Doudna and Charpentier, 2014; Friedland et al., 2013; Gratz et al., 2013; Hwang et al., 2013; Jinek et al., 2013, 2012; Li et al., 2013). This system uses a single guide RNA (sgRNA) to precisely target a nuclease (most often Cas9) to induce a DNA double strand break at a defined location (Doudna and Charpentier, 2014). The double strand break can then be repaired via the error-prone non-homologous end-joining pathway (often resulting in damaging frameshift mutations) or a more precise homology repair pathway, e.g. homology directed repair (HDR) or microhomology-mediated end joining. Following a double strand break, exogenous DNA repair templates can be used as substrates for homologous recombination, allowing virtually any desired sequence to be inserted anywhere in the genome. Importantly, CRISPR-Cas9 genome engineering is

remarkably efficient and robust in *C. elegans* (Dickinson and Goldstein, 2016; Norris et al., 2015).

Here, we present a broadly applicable strategy that adapts CRISPR-Cas9 genome engineering for targeted replacement of *C. elegans* genes with human genes. We illustrate how the library of knockout and humanized transgenics generated with this approach can be efficiently combined with automated machine vision phenotyping to rapidly discover novel gene functions, assess the functional conservation of human genes, and how this will allow for analysis of the effects of variants of uncertain significance with unprecedented precision. It is our hope that the human gene replacement and phenomic characterization strategy delineated in this article will serve both basic and health researchers alike, by serving as an open and shareable resource that will aid any genome engineer interested in understanding the functional conservation of human genes, and the functional consequences of their variants.

## Results

### A general genome editing strategy for direct replacement of a *C. elegans* gene with a single copy of its human ortholog

To replace the Open Reading Frame (ORF) of an orthologous gene with a human gene our strategy first directs an sgRNA to induce a Cas9 mediated DNA double strand break immediately downstream of the ortholog start codon (Fig. 1A). A co-injected repair template containing ~500 bp homology arms targeted to the regions immediately upstream and downstream of the ortholog ORF serve as a substrate for homology directed repair. By fusing the coding DNA sequence (CDS) of a human gene of interest

to the upstream homology arm homology directed repair integrates the human gene in place of the ortholog at a single copy in frame (Fig. 1A).

To streamline genome editing we have based our method on a recently described Dual-Marker Selection (DMS) cassette screening protocol (Norris et al., 2015). The DMS Cassette consists of an antibiotic resistance gene (*Prps-27::neoR::unc-54 UTR*) and a fluorescent marker (*Pmyo-2::GFP::unc-54 UTR*) that greatly facilitates the identification transgenic animals (Norris et al., 2015). We chose this cassette over similar methods as it: 1) can be used in any wild-type or mutant strain amenable to transgenesis and does not require any specialized genetic backgrounds, and 2) avoids the use of morphology and/or behavior altering selection markers that necessitate cassette excision prior to phenotypic analysis (Bend et al., 2016; Dickinson et al., 2015, 2013). In our strategy, the DMS cassette is placed between the human gene of interest and the downstream homology arm (Fig. 1A). This deletes the entire ORF of the ortholog and separates the human gene from the orthologs transcriptional terminator upon initial integration. In many cases this efficiently creates a useful deletion allele of the ortholog with no human gene expression. In the unlikely event human gene expression does occur without a transcription terminator a second repair template using the same validated homology arms and sgRNA and no human gene can be integrated to create an ortholog null. Importantly, the DMS cassette is flanked by two LoxP sites housed within synthetic introns that allow subsequent excision of the selection cassette via transient expression of Cre Recombinase (Norris et al., 2015). DMS cassette excision connects the human gene to the endogenous *C. elegans* orthologs transcriptional termination sequence such that a single copy of the human gene will now be expressed under the control of all of the orthologs 5' and 3' cis-

and trans-regulatory machinery (Fig. 1B). Validation of the desired edit is performed using standard PCR amplification and Sanger sequencing of both the 5' and 3' junctions of the target region (Fig. 1A).

For structure-function analysis of a human gene variant of uncertain significance the variant of interest is incorporated into the HDR plasmid using standard *in vitro* methods such as site-directed mutagenesis and the same genome editing process is repeated using the same validated sgRNA and homology arms. This process allows for: 1) initial generation and phenotypic analysis of a complete deletion allele in the *C. elegans* orthologous gene, 2) direct integration of the human gene to determine if the human gene can compensate for loss of the orthologous gene, measuring functional conservation, and 3) structure-function analysis of the effects of variants of uncertain significance on wild-type gene function (Fig. 1B). Importantly, the vast majority of variants of uncertain significance identified in patients are heterozygous and the integrated nature of the transgenes generated with this strategy allow for straightforward assessment of heterozygous alleles using standard genetic crosses (Fig. 1D).

**_PTEN_ as a prototypic disease-associated gene for targeted human gene replacement**

As a proof of principle, we demonstrated the utility of our strategy by focusing on the critical disease-associated gene phosphatase and tensin homolog (*PTEN*). *PTEN* is a lipid phosphatase that antagonizes the phosphoinositide 3-kinase (PI3K) signaling pathway by dephosphorylating phosphatidylinositol (3,4,5)-trisphosphate (PIP3) (Li et al., 1997; Maehama and Dixon, 1998). Heterozygous germline *PTEN* mutations are associated with diverse clinical outcomes including several tumor predisposition

phenotypes (collectively called *PTEN* harmatoma tumor syndrome), intellectual disability, and Autism Spectrum Disorders (Hobert et al., 2014; Li et al., 1997; Danny Liaw et al., 1997; McBride et al., 2010; O'Roak et al., 2012; Orrico et al., 2009; Sanders et al., 2015; Varga et al., 2009). Despite extensive study, it is currently impossible to predict the clinical outcome of a *PTEN* mutation carrier using sequence data alone (Matreyek et al., 2018; Mighell et al., 2018). *PTEN* also has several technical advantages that make it an ideal test case: 1) *PTEN* functions in the highly conserved insulin signaling pathway that is well characterized in *C. elegans* (Ozes et al. 2001; Ogg & Ruvkun 1998; Mihaylova et al. 1999, and Fig. 2B). 2) *C. elegans* has a single *PTEN* ortholog called *daf-18* (Fig. 2 A-D) and transgenic overexpression of human *PTEN* using extrachromosomal arrays has been shown to rescue reduced longevity and dauer defective phenotypes induced by mutations in *daf-18* (Liu and Chin-Sang, 2015; Solari et al., 2005). 3) *C. elegans* harboring homozygous *daf-18* null alleles are viable and display superficially normal morphology and spontaneous locomoter behavior (Mihaylova et al. 1999 and Fig. S1).

**An automated chemotaxis paradigm reveals a conserved nervous system role for PTEN in controlling NaCl preference**

Structure-function analyses of *PTEN* variants necessitate an *in vivo* phenotypic assay that reports on *PTEN* function. When wild-type worms are grown in the presence of NaCl and their food source (*Escherichia coli*) they display naïve attractive navigation behavior toward NaCl. This behavior is called NaCl chemotaxis and can be quantified by measuring navigation behavior of a population of animals up a controlled NaCl

concentration gradient (Ward, 1973). Previous work has shown that *daf-18* is required for attractive navigation behavior up a concentration gradient of NaCl such that animals with deletion or reduction-of-function mutations in *daf-18* display innate *aversion* to NaCl (Tomioka et al., 2006). Interestingly, Insulin/PI3K signaling normally *actively promotes salt avoidance* under naive conditions and *daf-18* functions in a single chemosensory neuron (ASER) to antagonize the Insulin/PI3K pathway and promote salt attraction (Adachi et al., 2010; Tomioka et al., 2006). Using our machine vision system, the Multi-Worm Tracker, we developed an automated high-throughput NaCl chemotaxis paradigm (Fig 3A,B) and replicated the finding that *daf-18(e1375)* reduction-of-function mutants display strong aversion to NaCl (Swierczek et al. 2011 and Fig. 3C,D). We then generated a transgenic line using traditional extrachromomosal array technology that directed pan neuronal expression of wild-type human *PTEN* using the *aex-3* promoter (Kuroyanagi et al., 2010). Pan neuronal expression of human *PTEN* is able to rescue the *daf-18* reduction of function phenotype and restore attractive NaCl chemotaxis to wild-type levels (Fig 3C,D). This work establishes NaCl chemotaxis as an *in vivo* behavioral assay of conserved nervous system *PTEN* functions.

**Complete deletion of the *daf-18* ORF causes strong NaCl avoidance that is not rescued by direct single-copy replacement with the canonical human *PTEN* CDS**

Next we used our strategy to create a *daf-18* deletion allele and single-copy replacement *daf-18* with the 1212 bp canonical human *PTEN* CDS. Complete deletion of the 4723 bp *daf-18* open reading frame resulted in strong aversion to NaCl and chemotaxis down the salt gradient (Fig. 4A). Chemotaxis avoidance of worms harboring

the *daf-18* complete deletion allele generated using CRISPR was not significantly different from that of worms carrying the previously characterized large deletion allele *daf-18(ok480)*, confirming effective inactivation of *daf-18* (Fig. 4B). DMS cassette excision and expression of a single copy of human *PTEN* was unable to substitute for *daf-18* and did not rescue attractive NaCl chemotaxis behavior (Fig. 4C). This result was observed in two independent single-copy human *PTEN* knock in lines on several independent experimental replications (Fig. 4B). Transcription of human *PTEN* was confirmed using Reverse transcription PCR in both knock in lines (Fig. 4D). Sanger sequencing confirmed error free insertion of transgenes at base pair resolution before and after cassette excision. There are several potential reasons why expression of human *PTEN* using extrachromasomal arrays rescued *daf-18* mutant phenotypes (Solari et al. 2005 and Fig. 3C,D) while targeted single-copy replacement with *PTEN* did not (Fig. 4C). The two most prominent differences between these two technologies are the expression level of the transgenes and the use of the endogenous 3' UTR in the CRISPR knock in versus the *unc-54* myosin 3' UTR used in most *C. elegans* transgenes to ensure proper processing of transcripts in all tissues, including all constructs previously shown to rescue *daf-18* phenotypes with human *PTEN* (Merritt and Seydoux, 2010; Solari et al., 2005, and Fig. 3C,D).

**A streamlined human gene replacement strategy functionally replaces *daf-18* with human *PTEN***

In order to address these limitations and increase the speed of transgenesis we created an alternate repair template strategy that includes a transcriptional termination

sequence in the upstream homology arm so that expression of the human gene begins immediately upon integration (Fig. 5A). We included the validated *unc-54 3' UTR* sequence in the upstream homology arm fused to the 3' end of human *PTEN* (Fig. 5A). Expression of wild-type human *PTEN* using this genome editing strategy rescued NaCl chemotaxis, indicating production of functional PTEN immediately upon genomic integration prior to cassette excision (Fig. 5B). This alternate approach also offers the added benefits of increased throughput (as a second injection step to excise the cassette was not required for human gene expression as it is in the first strategy) and the option for retained visual transgenic markers (either within the selection cassette or by adding a 2A sequence to drive reporter expression from the same promoter), which simplifies the generation and phenotypic analysis of heterozygotes and double mutants (Fig. 5A, Ahier and Jarriault, 2014; Calarco and Norris, 2018; Norris et al., 2017). Given the demonstrated versatility of CRISPR genome editing in *C. elegans*, these results suggest our strategy should be broadly applicable for *in vivo* analysis of diverse human disease-associated genes.

**_daf-18_ deletion causes mechanosensory hyporesponsivity that is rescued by targeted replacement with human _PTEN_**

A long-standing goal has been to understand how human disease-associated variants alter normal gene function to produce sensory and learning abnormalities characteristic of diverse neurogenetic disorders. The massive number variants of uncertain significance recently implicated in the etiology neurogenetic disorders necessitates a dramatic increase in throughput of both transgenic construction and

231 behavioral phenoptying if this goal is to be achieved (Ben-Shalom et al., 2017;
232 Geschwind and Flint, 2015; Lim et al., 2017; Sanders et al., 2015; Starita et al., 2017;
233 Wangler et al., 2017). By combing streamlined human gene integration with rapid
234 machine vision phenotypic analysis of *C. elegans* our strategy greatly simplifies the
235 identification of novel conserved gene functions in complex sensory and learning
236 phenotypes.

237 We explored whether *daf-18* mutants displayed behavioral deficits in
238 mechanosensory responding and/or habituation, a conserved form of non-associative
239 learning that is altered in several neurodevelopmental and neuropsychiatric disorders
240 (McDiarmid et al., 2018, 2017; Rankin et al., 2009; Stessman et al., 2017; Timbers et al.,
241 2017; van der Voet et al., 2014). When a non-localized mechanosensory stimulus is
242 delivered to the side of the petri plate *C. elegans* are cultured on they respond by eliciting
243 a brief reversal before resuming forward locomotion. Wild-type *C. elegans* habituate to
244 repeated stimuli by learning to decrease the probability of eliciting a reversal (Rankin et
245 al., 1990). To determine if *daf-18* is important for mechanoresponding and/or non-
246 associative learning we examined habituation of the *daf-18(e1375)* and *daf-18* complete
247 deletion mutants. Compared to wild-type animals both *daf-18(e1375)* reduction of
248 function and *daf-18* complete deletion mutants exhibited significantly reduced probability
249 of eliciting a reversal response throughout the habituation training session, indicating
250 mechanosensory hyporesponsivity (Fig. 6A,B). Despite this hyporesponsivity, the
251 plasticity of responses, or the pattern of the gradual decrement in the probability of
252 emitting of a reversal response throughout the training session was not significantly
253 altered in *daf-18* mutants (Fig. 6A,B,E). Importantly, targeted single-copy replacement of

*daf-18* with human *PTEN* was sufficient to rescue the mechanosensory hyporesponsivity phenotype across the training session towards wild-type levels (Fig. 6D). These results identify a novel conserved role for *PTEN* in mechanosensory responding, a fundamental biological process disrupted in diverse genetic disorders (Badr et al., 1987; McDiarmid et al., 2017; Orefice et al., 2016). More broadly, they illustrate how the library of transgenic animals generated with our strategy can be used to rapidly characterize the role of diverse human genes in complex sensory and learning behaviors. These novel phenotypes can then be used to investigate the functional consequences of disease-associated variants in intact and freely behaving animals and to screen for therapeutics that reverse the effects of a particular patients missense variant.

**Human gene replacement and *in vivo* phenotypic assessment accurately identifies functional consequences of the pathogenic PTEN-G129E variant**

To demonstrate the feasibility of our human gene replacement strategy for assessing variants of uncertain significance we set out to determine whether our *in vivo* functional assays could discern known pathogenic variants. Early studies characterizing the role of PTEN as a tumor suppressor suggested impaired protein phosphatase activity was key to the etiology of PTEN disorders (Tamura et al., 1998). However, this notion was challenged by the identification of cancer patients harboring a missense mutation that changes a glycine residue in the catalytic signature motif to a glutamate, which was predicted to abolish the lipid phosphatase activity of PTEN while leaving the protein phosphatase activity intact (Liaw et al., 1997; Myers et al., 1997). Subsequent biochemical analyses supported the now widely accepted view that the lipid phosphatase

activity of PTEN is critical to its tumor suppressor activity (Myers et al., 1998).

We used site-directed mutagenesis to incorporate the PTEN-G129E (*PTEN*, c386G>A) missense variant into our repair template and used our human gene replacement strategy to replace *daf-18* with a single copy of human PTEN-G129E (Fig. 7A). Animals harboring the PTEN-G129E variant displayed strong NaCl avoidance equivalent to animals carrying the complete *daf-18* deletion allele, indicating loss-of-function (Fig. 7B). Similarly, PTEN-G129E mutants also displayed mechanosensory hyporesponsivity that was not significantly different from *daf-18* deletion carriers (Fig. 7C). These *in vivo* phenotypic results accurately classify the pathogenic PTEN-G129E as a strong loss-of-function variant. In addition, by taking advantage of a pathogenic variant with well-characterized biochemical effects these results identify a necessary role for PTEN lipid phosphatase activity in both chemotaxis and mechanosensory responding, providing novel insight into the molecular mechanisms underlying these forms of sensory processing. Taken together, these results demonstrate the potential of human gene replacement and phenomic characterization to rapidly identify the functional consequences variants of uncertain significance.

## Discussion

We have developed and validated a broadly applicable strategy for targeted human gene replacement and phenomic characterization in *C. elegans* that will facilitate assessment of the functional conservation of human genes and structure-function analysis of variants of uncertain significance with unprecedented precision. We established an automated NaCl chemotaxis paradigm and demonstrate that pan neuronal overexpression

or direct replacement of *daf-18* with its human ortholog *PTEN* using CRISPR is sufficient to rescue reversed NaCl chemotaxis preference induced by complete *daf-18* deletion. We further identified a novel mechanosensory hyporesponsive phenotype for *daf-18* mutants that could also be rescued by targeted replacement with human *PTEN*. *In vivo* characterization of mutants harboring a single copy of the known lipid phosphatase inactive G129E variant accurately classified this variant as pathogenic and revealed a critical role for PTEN lipid phosphatase activity in NaCl chemotaxis and mechanosensory responding. As a resource article, we provide novel high-throughput *in vivo* functional assays for *PTEN*, as well as validated strains, reagents, sgRNA and repair template constructs to catalyze further analysis of this critical human disease-associated gene. More broadly, we provide a conceptual framework that illustrates how genome engineering and automated machine vision phenotyping can be combined to efficiently generate and characterize a library of knockout and humanized transgenic strains that will allow for straightforward and precise analysis of human genes and disease-associated variants *in vivo* (Fig. 8).

**Comparing CRISPR targeted human gene replacement with orthology-based variant assessment methods**

To date, the most widely used genome editing-based human disease variant assessment strategy in *C. elegans* uses sequence alignments to identify and engineer the corresponding amino acid change into the orthologous *C. elegans* gene (Bend et al., 2016; Bulger et al., 2017; Canning et al., 2018; Pierce et al., 2018; Prior et al., 2017; Sorkaç et al., 2016; Troulinaki et al., 2018). A major advantage of this approach is that by

using the *C. elegans* gene, intronic regulation, protein-protein interactions, subcellular localization, and biochemical activity of the protein of interest are, by design, perfectly modeled by *C. elegans*. Even when expressed at physiologically relevant levels directly from the orthologs native loci, and with evidence of phenotypic rescue, it is not guaranteed a transgenic human protein will recapitulate all functions and interactions of the orthologous *C. elegans* protein. This presents an important consideration when attempting to replace *C. elegans* proteins that must interact in extremely precise heteromeric complexes to perform their molecular functions (e.g. certain ion channel subunits) (Bend et al., 2016; Prior et al., 2017). The most obvious limitation of this approach is that it can only be used to study orthologous amino acids. Amino acids that have been conserved throughout evolution are, by definition, the least tolerant to mutation and are thus far more likely to be detrimental to protein function when mutated (Starita et al., 2017; Weile et al., 2017). Indeed, many variant effect prediction algorithms rely on sequence conservation as the main predictor that a variant will be deleterious (Richards et al., 2015; Starita et al., 2017). It is also important to note that a large portion of amino acids will not be conserved to humans (e.g. >50% of amino acids are not conserved from DAF-18 to PTEN, Fig. 2C) and alignment algorithms that identify orthologous amino acids are imperfect. Many current implementations of orthologous amino acid engineering also require constant generation of completely new sgRNAs and repair templates for subsequent edits. Human gene replacement, in contrast, allows any coding variant of uncertain significance to be studied *in vivo* using the same validated sgRNA and homology arms.

One potential limitation of human gene replacement is that it relies on a *C. elegans*

346 ortholog to replace. Depending on the orthology prediction program used estimates
347 suggest there are *C. elegans* orthologs for 60-80% of human disease-associated genes
348 meaning their will not be an ortholog available for a minority of human genes (Kaletta
349 and Hengartner, 2006; Lai et al., 2000b; Shaye and Greenwald, 2011). This is even less
350 likely to be a problem for disease variant modelling, as disease-associated genes are more
351 likely to be highly conserved (Aerts et al., 2006; López-Bigas and Ouzounis, 2004). A
352 recent study also showed that ~10% of human disease-associated genes are able to
353 functionally substitute for their yeast paralogs (in addition to orthologs), further
354 increasing the number of human genes that can be studied by replacement (Yang et al.,
355 2017). Still, in the event there is no suitable *C. elegans* ortholog or paralog available for a
356 human gene of interest the human gene can simply be integrated at a putatively neutral
357 genomic location using CRISPR or transposon mobilization and expressed from a
358 heterologous promoter (Frøkjær-Jensen et al., 2014, 2012, 2008). This approach can be
359 used to screen for phenotypes induced by expression of any human gene and to determine
360 whether a variant exacerbates or eliminates these effects (Baruah et al., 2017).

361 With the rapidly expanding set of precise genome editing techniques available to *C.*
362 *elegans,* researchers interested in variants of uncertain significance now have the freedom
363 to choose the approach that best suits their particular needs and interests. The approaches
364 described here provide a diverse collection of methods that can be sequentially tested in a
365 pragmatic hierarchy of precision, beginning with direct replacement and working down
366 until phenotypic rescue is achieved. Regardless, it will always be ideal to have
367 corroborating evidence of variant effect from multiple techniques, indeed multiple model
368 systems, to best inform clinical decisions.

**Combining human gene replacement and automated phenomic characterization to discover conserved gene functions and establish variant functional assays**

A necessary step in the establishment of human gene functional assays is the identification of phenotypes that are rescued by or induced upon expression of the reference (wild-type) human gene, as we have done here for human *PTEN* and NaCl chemotaxis. Indeed, the establishment of functional assays remains a major bottleneck for variant assessment across species (Starita et al., 2017; Weile et al., 2017). While traditional extrachromosomal array transgenes offer a quick way to establish such assays several aspects of these transgenes can severely impede this process. These include but are not limited to: 1) variable overexpression of transgenes which can lead to silencing of transgenes in certain tissues, complicating phenotypic analysis (e.g. multi-copy transgenes are expressed in the soma but silenced in the germline while low- and single-copy transgenes are expressed in both) (Kelly et al., 1997; Merritt and Seydoux, 2010), and 2) variably mosaic expression which can make it extremely difficult to assess rescue of partially penetrant and subtle complex phenotypes, as an animal must simultaneously carry the extrachromosomal array and be one of the members of the population that displays the partially penetrant phenotype that can often be difficult to score.

The human gene replacement approach described here allows for generation of ortholog deletion alleles directly in any wild-type or mutant strain amenable to transgenesis using the same reagents designed for replacement, thereby reducing the confounding effects of background mutations on phenotype discovery. The use of excisable selectable markers that do not severely alter morphology, baseline locomotion,

and several evoked sensory behaviors further simplifies phenotypic analysis and removes the need for any specialized genetic backgrounds (Norris et al. 2015 and Figs 4, 5, 6, S1). Using this approach in combination with machine vision we provide two *in vivo* functional assays for human *PTEN*, NaCl chemotaxis and mechanosensory responsivity. In particular, NaCl chemotaxis possesses several characteristics that make it an ideal functional assay: 1) a large functional range between deletion and human gene rescue to discern potentially subtle functional differences among missense variants; 2) scalability as many plates can be run simultaneously; 3) straightforward analysis using an automatically calculated preference index (alternatively many labs score chemotaxis manually by simple blinded counting). Importantly, these reagents and functional assays can now be used in precision medicine drug screens aimed at identifying compounds that counteract the affect of a particular patient's missense variant. Further broad-scale phenomic characterization of targeted knockouts and mutant libraries combined with new databases that curate ortholog functional annotation across model organisms should expedite the process of *in vivo* functional assay development (Lee et al., 2018; McDiarmid et al., 2018; Thompson et al., 2013; Wang et al., 2017; Yemini et al., 2013).

**Further applications of targeted human gene replacement**

One of the goals of this resources article is to illustrate how Cas9-triggered homologous recombination can be adapted to directly replace *C. elegans* genes with human genes. An exciting adaptation of this approach will be to combine targeted human gene insertions with bi-partite systems for precise spatial-temporal control of human transgene expression, as has recently been done to overexpress human genes as UAS-

cDNA constructs in *Drosophila* (Luo et al. 2017; Lee et al. 2018; Wangler et al. 2017). The recent and long-awaited development of the cGAL4-UAS system should allow a similar approach to be developed for *C. elegans*, currently the only organism where the complete cell lineage and neuronal wiring diagram is known (Sulston et al., 1983; Sulston and Horvitz, 1977; Wang et al., 2016; White et al., 1986).

While one of the clearest uses for targeted replacement is precise structure-function analysis of variants of uncertain significance there are several exciting applications beyond modeling disease-associated alleles. Targeted human gene replacement is also particularly well suited for investigation of the evolutionary principles that determine the replaceability of genes. By allowing for human gene expression immediately upon genomic integration this approach should be also adaptable to essential genes. A homozygous integrant will only be obtained if the human gene can substitute for the orthologous essential gene, creating a complementation test out of the transgenesis process. This will allow for systematic and precise interrogation of the sequence characteristics and functional properties required for successful human gene replacement (Kachroo et al., 2015). A library of humanized worms would also open the door to the rich resources of tools available to visualize and manipulate human genes and proteins that are often unavailable for *C. elegans* researchers (e.g. high-quality antibodies, biochemically characterized or known pathogenic control variants, and experimentally determined crystal structures, Figs 2, 7; Berman et al. 2000). Given the throughput that has been achieved for reporter gene analysis (several thousand genes) and genome editing in *C. elegans* it should be possible to generate a humanized *C. elegans* library of similar

size to those recently created in yeast (Dickinson and Goldstein, 2016; Dupuy et al., 2007; Hamza et al., 2015; Hunt-Newbury et al., 2007; Kachroo et al., 2015; Norris et al., 2017; Sun et al., 2016; Yang et al., 2017). Integrated transgenes would offer the possibility of humanizing entire cellular processes for detailed *in vivo* analysis in a relatively complex yet tractable metazoan with increasingly powerful tools for spatial-temporal control of transgene expression and protein degradation (Armenti et al., 2014; Wang et al., 2016; S. Wang et al., 2017; Zhang et al., 2015).

Deep mutational scanning and related technologies have recently made it feasible to characterize the functional effects of virtually every possible amino acid change of a protein on a particular cellular phenotype (Fowler et al., 2010; Fowler and Fields, 2014). Several such exhaustive sequence-function maps have been recently been generated in yeast and human cell culture systems (Findlay et al., 2018, 2014; Majithia et al., 2016; Matreyek et al., 2018; Mighell et al., 2018; Weile et al., 2017). These tools offer amazing resources that serve as 'lookup tables' of functional missense variation in human genes, to enable experimentally confirmed variant interpretation immediately upon first clinical presentation (Starita et al., 2017; Weile et al., 2017). An ambitious and exciting goal for the *C. elegans* community will be to further streamline genome engineering and high-throughput phenotyping to achieve the first comprehensive sequence-function map in a metazoan.

## Materials and Methods

### Strains and culture

Worms were cultured on Nematode Growth Medium (NGM) seeded with *Escherichia coli* (OP50) as described previously (Brenner, 1974). N2 Bristol, and CB1375 *daf-18(e1375)* strains were obtained from the *Caenorhabditis* Genetics Center (University of Minnesota, USA). *daf-18(e1375)* harbors a 30–base pair insertion in the fourth exon and is predicted to insert six amino acids before introducing an early stop codon that truncates the C-terminal half of the protein while leaving the phosphatase domain intact (Ogg and Ruvkun, 1998).

The following strains were created for this work:

VG674 *daf-18(e1375); yvEx674[paex-3::PTEN::unc-54; pmyo-2::mCherry::unc-54 UTR]*

VG810-813 *daf-18(e1375); yvEx810-813[paex-3::PTEN::unc-54 UTR; pmyo-2::mCherry::unc-54 UTR]*

VG712 *daf-18(yv3[daf-18p::PTEN + LoxP pmyo-2::GFP::unc-54 UTR prps-27::NeoR::unc-54 UTR LoxP + daf-18 UTR])*

VG713 *daf-18(yv4[daf-18p::PTEN + LoxP pmyo-2::GFP::unc-54 UTR prps-27::NeoR::unc-54 UTR LoxP +daf-18 UTR])*

VG714 *daf-18(yv5[daf-18p::PTEN + LoxP + daf-18 UTR ])*

VG715 *daf-18(yv5[daf-18p::PTEN + LoxP + daf-18 UTR])*

VG817 *daf-18(yv7[daf-18p::GFP::T2A::PTEN::unc-54 UTR + LoxP pmyo-2::GFP::unc-54 UTR prps-27::NeoR::unc-54 UTR LoxP + daf-18 UTR])*

VG818 *daf-18(yv8[daf-18p::GFP::T2A::PTEN::unc-54 UTR + LoxP pmyo-2::GFP::unc-54 UTR prps-27::NeoR::unc-54 UTR LoxP + daf-18 UTR])*

VG867 *daf-18(yv14[daf-18p::GFP::T2A::PTEN*-G129E*::unc-54 UTR + LoxP pmyo-2::GFP::unc-54 UTR prps-27::NeoR::unc-54 UTR LoxP + daf-18 UTR])*

**Strain and plasmid generation**

The reference *PTEN* CDS (UniProt consensus, identifier: P60484-1) was obtained from a pCMV-PTEN plasmid (Addgene plasmid #28298) and cloned into a TOPO gateway entry clone (Invitrogen) according to manufacturers instructions. The *PTEN* entry clone was recombined with an pDEST-*aex-3p* destination vector (obtained from Dr. Hidehito Kuroyanagi; Kuroyanagi et al., 2010) to generate the *aex-3p::PTEN::unc-54 UTR* rescue construct using gateway cloning (Invitrogen), according to manufacturers instructions.

The Moerman lab guide selection tool (http://genome.sfu.ca/crispr/) was used to identify the *daf-18* targeting sgRNA. The *daf-18* sgRNA sequence: GGAGGAGGAGTAACCATTGG was cloned into the *pU6::klp-12* sgRNA vector (obtained from Calarco lab) using site-directed mutagenesis and used for all editing experiments. The *daf-18p::PTEN* CDS and *daf-18p::GFP::T2A::PTEN::unc-54 UTR*

upstream homology arms were synthesized by IDT and cloned into the loxP_myo2_neoR repair construct (obtained from Calarco lab) using Gibson Assembly.

*C. elegans* wild-type N2 strain was used for all CRISPR editing experiments. Genome edits were created as previously described (Norris et al., 2015). In brief, plasmids encoding sgRNA, Cas9 co-transformation markers pCFJ90 and pCFJ104 (Jorgensen lab, Addgene) and the selection cassette flanked by homology arms (~500 bp) containing *PTEN* were injected into wild-type worms. Animals containing the desired insertions were identified by G418 resistance, loss of extrachromosomal array markers, and uniform dim fluorescence of the inserted GFP.

**Genotype confirmation**

Correct replacement of the *daf-18* ORF with human *PTEN* was confirmed by amplifying the two regions spanning the upstream and downstream insertion borders using PCR followed by Sanger sequencing (primer binding locations depicted in Fig. 1). The genotyping strategy is essentially as described for deletion allele generation via DMS cassette insertion in (Norris et al. 2015).

The forward and reverse primers used to amplify the upstream insertion region were TGCCGTTTGAATTAGCGTGC (located within the *daf-18* genomic promoter region) and CCCTCAATGTCTCTACTTGT (located within the *myo-2* promoter of the selection cassette) respectively.

The forward and reverse primers used to amplify the downstream insertion region were TTCCTCGTGCTTTACGGTATCG (located within the Neomycin resistance gene) and

529 CTCAACACGTTCGGAGGGTAAA (located downstream of the *daf-18* genomic coding
530 region) respectively.
531
532 Following cassette excision via injection of cre-recombinase the *daf-18* promoter
533 (TGCCGTTTGAATTAGCGTGC) and *daf-18* downstream
534 (CTCAACACGTTCGGAGGGTAAA) primers were used to amplify human *PTEN* and
535 confirm error free insertion at the *daf-18* locus via Sanger sequencing (Fig. 1).
536
537 **RNA extraction, library preparation, and cDNA amplification**
538 Total RNA was isolated from mixed stage VG714 and VG715 *PTEN* knock in
539 animals using a GeneJET RNA Purification Kit (ThermoFisher) according to
540 manufacturers instructions. Total RNA was treated with DNase (New England Biolabs)
541 and purified with an RNeasy MinElute spin column (Qiagen) according to manufacturers
542 instructions. cDNA libraries were prepared from crude and purified total RNA using
543 Superscript III (Invitrogen). All genes were amplified from cDNA libraries with Platinum
544 Taq DNA Polymerase (ThermoFisher) and gene-specific primer sets.
545
546 The forward and reverse primers used to amplify the *PTEN* CDS were
547 ATGACAGCCATCATCAAAGA and TCAGACTTTTGTAATTTGTG respectively.
548
549 The forward and reverse *cmk-1* intronic control primers (Ardiel et al., 2018) were
550 AGGGTAGGCTAGAGTCTGGGATAGAT and
551 ACGACTCCGTTGTCGTGCATAAAC respectively.

552 **Protein structure modeling and visualization**

553 The PTEN 1D5R reference structure (Berman et al., 2000; Lee et al., 1999) was
554 visualized using PyMOL software (DeLano, 2002). Structural models for full-length
555 human PTEN, DAF-18 53-506, and full-length DAF-18 were predicted using Phyre2
556 (Kelley et al., 2015) and visualized using PyMOL.

557

558 **NaCl chemotaxis behavioral assays**

559 The chemotaxis behavioral assay was conducted on a 6 cm assay plate (2% agar),
560 where a salt gradient was formed overnight by inserting a 2% agar plug containing 50mM
561 of NaCl (approximately 5 mm in diameter) 1 cm from the edge of the plate. A control 2%
562 agar plug without NaCl was inserted 1 cm from the opposite edge of the plate. Strains
563 were grown on NGM plates seeded with *E. coli* (OP50) for 3 or 4 days. Worms on the
564 plates were collected and washed three times using M9 buffer before being pipetted onto
565 an unseeded NGM plate to remove excess buffer and select animals carrying
566 transformation markers. Adult worms were transferred and placed at the centre of the
567 assay plates and were tracked for 40 minutes on the Multi-Worm Tracker (Swierczek et
568 al., 2011). After the tracking period, the chemotaxis index was calculated as (A – B)/(A +
569 B), where A was the number of animals that were located in a 1.5 cm-wide region on the
570 side of the assay plate containing the 2% agar plug with 50mM NaCl and B was the
571 number of animals that were located in a 1.5 cm-wide region on the side of the assay
572 plate containing the 2% agar plug without NaCl (Fig. 3B). Animals not located in either
573 region (ie. the middle section of the assay plate) were not counted towards the
574 chemotaxis index. One hundred to two hundred animals were used per plate, and two or

three plate replicates were used for each line in each experiment. Any statistical comparisons were carried out on plates assayed concurrently (i.e. on the same day).

**Mechanosensory habituation behavioral assays**

Worms were synchronized for behavioral testing on Petri plates containing Nematode Growth Media (NGM) seeded with 50 µl of OP50 liquid culture 12-24 hours before use. Five gravid adults were picked to plates and allowed to lay eggs for 3-4 hours before removal. The animals were maintained in a 20°C incubator for 72 hours. Plates of worms were placed into the tapping apparatus and after a 100s acclimatization period, 30 taps were administered at a 10s ISI. Comparisons of "final response" comprised the average of the final three stimuli. Any statistical comparisons were carried out on plates assayed concurrently (i.e. on the same day).

**Multi-worm tracker behavioral analysis and statistics**

Multi-Worm Tracker software (version 1.2.0.2) was used for stimulus delivery and image acquisition (Swierczek et al., 2011). Behavioral quantification with Choreography software (version 1.3.0_r103552) used "--shadowless", "--minimum-move-body 2", and "--minimum-time 20" filters to restrict the analysis to animals that moved at least 2 body lengths and were tracked for at least 20 s. The MeasureReversal plugin was used to identify reversals occurring within 1 s ($dt = 1$) of the mechanosensory stimulus onset. Custom MatLab and R scripts organized and summarized Choreography output files. Final figures were generated using GraphPad Prism version 7.00 for Mac OS X. Each experiment was independently replicated at least twice. No blinding was

598 necessary because the Multi-Worm Tracker scores behavior objectively. Morphology
599 metrics, baseline locomotion metrics, initial and final reversal responses, habituation
600 difference scores, or chemotaxis indices from all plates were pooled and metrics were
601 compared across strains with ANOVA and Tukey honestly significant difference (HSD)
602 tests. For all statistical tests an alpha value of 0.05 was used to determine significance.
603

604 **Acknowledgements**

605 We would like to thank Dr. Evan L. Ardiel useful comments and discussion
606 regarding the manuscript. We would like to thank Dr. Kurt Haas for motivating
607 discussions to write the manuscript. We would like to thank Erica Li-Leger and the
608 Moerman lab for assistance with experiments. We would like to thank Dr. John Calarco,
609 Dr. Erik Jorgensen, and Dr. Hidehito Kuroyanagi and their labs for sharing their
610 constructs and protocols or making them publicly available. We would like to thank
611 Warren M. Meyers and Christine R. Ackerley for useful advice and discussions regarding
612 figure design and/or protein structural modeling. We would also like to thank
613 the *Caenorhabditis* Genetic Center for strains.

614 **Competing interests**

615 The authors declare no competing interests.

616 **Funding**

617 This work was supported by a Canadian Institutes of Health Research Doctoral
618 Research Award to TAM, and a CIHR operating grant (operating grant #CIHR MOP
619 130287 to CHR.
620

621 **Data availability**

622 The data sets generated during the current study are available from the
623 corresponding author on reasonable request. The code used to analyze data in the current
624 study is available from the first author on reasonable request. All strains and reagents are
625 available from the corresponding author upon reasonable request.

626

627 **Author Contribution Statement**

628 TAM, VA, KM, and DGM conceived the genome editing strategy, designed and
629 built constructs. TAM, VA, and JL generated the transgenic lines. TAM, VA, and KM
630 performed genotyping and RT-PCR experiments. TAM and CHR designed the behavioral
631 experiments. TAM, and ADL performed the behavioral experiments. TAM wrote custom
632 scripts in R to organize the data and TAM and CHR analyzed the data. TAM wrote the
633 first draft of the manuscript and made the figures. TAM, VA, KM, DGM, and CHR
634 edited and co-wrote the final manuscript.

635

636

637

638

639

640

641

642

643

644 **References**

645

646 Adachi, T., Kunitomo, H., Tomioka, M., Ohno, H., Okochi, Y., Mori, I., Iino, Y., 2010.
647 Reversal of salt preference is directed by the insulin/PI3K and Gq/PKC signaling in
648 Caenorhabditis elegans. Genetics 186, 1309–19. doi:10.1534/genetics.110.119768
649 Aerts, S., Lambrechts, D., Maity, S., Van Loo, P., Coessens, B., De Smet, F.,
650 Tranchevent, L.-C., De Moor, B., Marynen, P., Hassan, B., Carmeliet, P., Moreau,
651 Y., 2006. Gene prioritization through genomic data fusion. Nat. Biotechnol. 24,
652 537–544. doi:10.1038/nbt1203
653 Ahier, A., Jarriault, S., 2014. Simultaneous expression of multiple proteins under a single
654 promoter in Caenorhabditis elegans via a versatile 2A-based toolkit. Genetics 196,
655 605–13. doi:10.1534/genetics.113.160846
656 Ardiel, E.L., McDiarmid, T.A., Timbers, T.A., Lee, K., Safaei, J., Pelech, S.L., Rankin,
657 C.H., 2018. CaMK (CMK-1) and O-GlcNAc transferase (OGT-1) modulate
658 mechanosensory responding and habituation in an interstimulus interval-dependent
659 manner in Caenorhabditis elegans. bioRxiv 115972. doi:10.1101/115972
660 Armenti, S.T., Lohmer, L.L., Sherwood, D.R., Nance, J., 2014. Repurposing an
661 endogenous degradation system for rapid and targeted depletion of C. elegans
662 proteins. Development 141, 4640–4647. doi:10.1242/dev.115048
663 Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R.,
664 Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B.,
665 Gibbs, R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S.,
666 Lee, C., Lehrach, H., Mardis, E.R., Marth, G.T., McVean, G.A., Nickerson, D.A.,
667 Schmidt, J.P., Sherry, S.T., Wang, J., Wilson, R.K., Gibbs, R.A., Boerwinkle, E.,
668 Doddapaneni, H., Han, Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J.G.,
669 Zhu, Y., Wang, J., Chang, Y., Feng, Q., Fang, X., Guo, X., Jian, M., Jiang, H., Jin,
670 X., Lan, T., Li, G., Li, J., Li, Y., Liu, S., Liu, X., Lu, Y., Ma, X., Tang, M., Wang,
671 B., Wang, G., Wu, H., Wu, R., Xu, X., Yin, Y., Zhang, D., Zhang, W., Zhao, J.,
672 Zhao, M., Zheng, X., Lander, E.S., Altshuler, D.M., Gabriel, S.B., Gupta, N.,
673 Gharani, N., Toji, L.H., Gerry, N.P., Resch, A.M., Flicek, P., Barker, J., Clarke, L.,
674 Gil, L., Hunt, S.E., Kelman, G., Kulesha, E., Leinonen, R., McLaren, W.M.,
675 Radhakrishnan, R., Roa, A., Smirnov, D., Smith, R.E., Streeter, I., Thormann, A.,
676 Toneva, I., Vaughan, B., Zheng-Bradley, X., Bentley, D.R., Grocock, R.,
677 Humphray, S., James, T., Kingsbury, Z., Lehrach, H., Sudbrak, R., Albrecht, M.W.,
678 Amstislavskiy, V.S., Borodina, T.A., Lienhard, M., Mertes, F., Sultan, M.,
679 Timmermann, B., Yaspo, M.-L., Mardis, E.R., Wilson, R.K., Fulton, L., Fulton, R.,
680 Sherry, S.T., Ananiev, V., Belaia, Z., Beloslyudtsev, D., Bouk, N., Chen, C.,
681 Church, D., Cohen, R., Cook, C., Garner, J., Hefferon, T., Kimelman, M., Liu, C.,
682 Lopez, J., Meric, P., O'Sullivan, C., Ostapchuk, Y., Phan, L., Ponomarov, S.,
683 Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang, H., McVean, G.A.,
684 Durbin, R.M., Balasubramaniam, S., Burton, J., Danecek, P., Keane, T.M., Kolb-
685 Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Schmidt, J.P., Davies, C.J.,
686 Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Campbell, C.L., Kong, Y.,
687 Marcketta, A., Gibbs, R.A., Yu, F., Antunes, L., Bainbridge, M., Muzny, D., Sabo,

A., Huang, Z., Wang, J., Coin, L.J.M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Y., Li, Z., Lin, H., Liu, B., Luo, R., Shao, H., Xie, Y., Ye, C., Yu, C., Zhang, F., Zheng, H., Zhu, H., Alkan, C., Dal, E., Kahveci, F., Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Fung Leong, W., Stromberg, M., Ward, A.N., Wu, J., Zhang, M., Daly, M.J., DePristo, M.A., Handsaker, R.E., Altshuler, D.M., Banks, E., Bhatia, G., del Angel, G., Gabriel, S.B., Genovese, G., Gupta, N., Li, H., Kashin, S., Lander, E.S., McCarroll, S.A., Nemesh, J.C., Poplin, R.E., Yoon, S.C., Lihm, J., Makarov, V., Clark, A.G., Gottipati, S., Keinan, A., Rodriguez-Flores, J.L., Korbel, J.O., Rausch, T., Fritz, M.H., Stütz, A.M., Flicek, P., Beal, K., Clarke, L., Datta, A., Herrero, J., McLaren, W.M., Ritchie, G.R.S., Smith, R.E., Zerbino, D., Zheng-Bradley, X., Sabeti, P.C., Shlyakhter, I., Schaffner, S.F., Vitti, J., Cooper, D.N., Ball, E. V., Stenson, P.D., Bentley, D.R., Barnes, B., Bauer, M., Keira Cheetham, R., Cox, A., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J., Shaw, R., Kenny, E.E., Batzer, M.A., Konkel, M.K., Walker, J.A., MacArthur, D.G., Lek, M., Sudbrak, R., Amstislavskiy, V.S., Herwig, R., Mardis, E.R., Ding, L., Koboldt, D.C., Larson, D., Ye, K., Gravel, S., Swaroop, A., Chew, E., Lappalainen, T., Erlich, Y., Gymrek, M., Frederick Willems, T., Simpson, J.T., Shriver, M.D., Rosenfeld, J.A., Bustamante, C.D., Montgomery, S.B., De La Vega, F.M., Byrnes, J.K., Carroll, A.W., DeGorter, M.K., Lacroute, P., Maples, B.K., Martin, A.R., Moreno-Estrada, A., Shringarpure, S.S., Zakharia, F., Halperin, E., Baran, Y., Lee, C., Cerveira, E., Hwang, J., Malhotra, A., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Hyland, F.C.L., Craig, D.W., Christoforides, A., Homer, N., Izatt, T., Kurdoglu, A.A., Sinari, S.A., Squire, K., Sherry, S.T., Xiao, C., Sebat, J., Antaki, D., Gujral, M., Noor, A., Ye, K., Burchard, E.G., Hernandez, R.D., Gignoux, C.R., Haussler, D., Katzman, S.J., James Kent, W., Howie, B., Ruiz-Linares, A., Dermitzakis, E.T., Devine, S.E., Abecasis, G.R., Min Kang, H., Kidd, J.M., Blackwell, T., Caron, S., Chen, W., Emery, S., Fritsche, L., Fuchsberger, C., Jun, G., Li, B., Lyons, R., Scheller, C., Sidore, C., Song, S., Sliwerska, E., Taliun, D., Tan, A., Welch, R., Kate Wing, M., Zhan, X., Awadalla, P., Hodgkinson, A., Li, Y., Shi, X., Quitadamo, A., Lunter, G., McVean, G.A., Marchini, J.L., Myers, S., Churchhouse, C., Delaneau, O., Gupta-Hinch, A., Kretzschmar, W., Iqbal, Z., Mathieson, I., Menelaou, A., Rimmer, A., Xifara, D.K., Oleksyk, T.K., Fu, Y., Liu, X., Xiong, M., Jorde, L., Witherspoon, D., Xing, J., Eichler, E.E., Browning, B.L., Browning, S.R., Hormozdiari, F., Sudmant, P.H., Khurana, E., Durbin, R.M., Hurles, M.E., Tyler-Smith, C., Albers, C.A., Ayub, Q., Balasubramaniam, S., Chen, Y., Colonna, V., Danecek, P., Jostins, L., Keane, T.M., McCarthy, S., Walter, K., Xue, Y., Gerstein, M.B., Abyzov, A., Balasubramanian, S., Chen, J., Clarke, D., Fu, Y., Harmanci, A.O., Jin, M., Lee, D., Liu, J., Jasmine Mu, X., Zhang, J., Zhang, Y., Li, Y., Luo, R., Zhu, H., Alkan, C., Dal, E., Kahveci, F., Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Ward, A.N., Wu, J., Zhang, M., McCarroll, S.A., Handsaker, R.E., Altshuler, D.M., Banks, E., del Angel, G., Genovese, G., Hartl, C., Li, H., Kashin, S., Nemesh, J.C., Shakir, K., Yoon, S.C., Lihm, J., Makarov, V., Degenhardt, J., Korbel, J.O., Fritz, M.H., Meiers, S., Raeder, B., Rausch, T., Stütz, A.M., Flicek, P., Paolo Casale, F., Clarke, L., Smith, R.E., Stegle, O., Zheng-Bradley, X., Bentley, D.R., Barnes, B., Keira Cheetham, R., Eberle, M., Humphray, S., Kahn, S., Murray, L., Shaw, R., Lameijer, E.-W., Batzer, M.A., Konkel, M.K.,

734 Walker, J.A., Ding, L., Hall, I., Ye, K., Lacroute, P., Lee, C., Cerveira, E., Malhotra,
735 A., Hwang, J., Plewczynski, D., Radew, K., Romanovitch, M., Zhang, C., Craig,
736 D.W., Homer, N., Church, D., Xiao, C., Sebat, J., Antaki, D., Bafna, V.,
737 Michaelson, J., Ye, K., Devine, S.E., Gardner, E.J., Abecasis, G.R., Kidd, J.M.,
738 Mills, R.E., Dayama, G., Emery, S., Jun, G., Shi, X., Quitadamo, A., Lunter, G.,
739 McVean, G.A., Chen, K., Fan, X., Chong, Z., Chen, T., Witherspoon, D., Xing, J.,
740 Eichler, E.E., Chaisson, M.J., Hormozdiari, F., Huddleston, J., Malig, M., Nelson,
741 B.J., Sudmant, P.H., Parrish, N.F., Khurana, E., Hurles, M.E., Blackburne, B.,
742 Lindsay, S.J., Ning, Z., Walter, K., Zhang, Y., Gerstein, M.B., Abyzov, A., Chen, J.,
743 Clarke, D., Lam, H., Jasmine Mu, X., Sisu, C., Zhang, J., Zhang, Y., Gibbs, R.A.,
744 Yu, F., Bainbridge, M., Challis, D., Evani, U.S., Kovar, C., Lu, J., Muzny, D.,
745 Nagaswamy, U., Reid, J.G., Sabo, A., Yu, J., Guo, X., Li, W., Li, Y., Wu, R., Marth,
746 G.T., Garrison, E.P., Fung Leong, W., Ward, A.N., del Angel, G., DePristo, M.A.,
747 Gabriel, S.B., Gupta, N., Hartl, C., Poplin, R.E., Clark, A.G., Rodriguez-Flores, J.L.,
748 Flicek, P., Clarke, L., Smith, R.E., Zheng-Bradley, X., MacArthur, D.G., Mardis,
749 E.R., Fulton, R., Koboldt, D.C., Gravel, S., Bustamante, C.D., Craig, D.W.,
750 Christoforides, A., Homer, N., Izatt, T., Sherry, S.T., Xiao, C., Dermitzakis, E.T.,
751 Abecasis, G.R., Min Kang, H., McVean, G.A., Gerstein, M.B., Balasubramanian, S.,
752 Habegger, L., Yu, H., Flicek, P., Clarke, L., Cunningham, F., Dunham, I., Zerbino,
753 D., Zheng-Bradley, X., Lage, K., Berg Jespersen, J., Horn, H., Montgomery, S.B.,
754 DeGorter, M.K., Khurana, E., Tyler-Smith, C., Chen, Y., Colonna, V., Xue, Y.,
755 Gerstein, M.B., Balasubramanian, S., Fu, Y., Kim, D., Auton, A., Marcketta, A.,
756 Desalle, R., Narechania, A., Wilson Sayres, M.A., Garrison, E.P., Handsaker, R.E.,
757 Kashin, S., McCarroll, S.A., Rodriguez-Flores, J.L., Flicek, P., Clarke, L., Zheng-
758 Bradley, X., Erlich, Y., Gymrek, M., Frederick Willems, T., Bustamante, C.D.,
759 Mendez, F.L., David Poznik, G., Underhill, P.A., Lee, C., Cerveira, E., Malhotra,
760 A., Romanovitch, M., Zhang, C., Abecasis, G.R., Coin, L., Shao, H., Mittelman, D.,
761 Tyler-Smith, C., Ayub, Q., Banerjee, R., Cerezo, M., Chen, Y., Fitzgerald, T.W.,
762 Louzada, S., Massaia, A., McCarthy, S., Ritchie, G.R., Xue, Y., Yang, F., Gibbs,
763 R.A., Kovar, C., Kalra, D., Hale, W., Muzny, D., Reid, J.G., Wang, J., Dan, X.,
764 Guo, X., Li, G., Li, Y., Ye, C., Zheng, X., Altshuler, D.M., Flicek, P., Clarke, L.,
765 Zheng-Bradley, X., Bentley, D.R., Cox, A., Humphray, S., Kahn, S., Sudbrak, R.,
766 Albrecht, M.W., Lienhard, M., Larson, D., Craig, D.W., Izatt, T., Kurdoglu, A.A.,
767 Sherry, S.T., Xiao, C., Haussler, D., Abecasis, G.R., McVean, G.A., Durbin, R.M.,
768 Balasubramaniam, S., Keane, T.M., McCarthy, S., Stalker, J., Chakravarti, A.,
769 Knoppers, B.M., Abecasis, G.R., Barnes, K.C., Beiswanger, C., Burchard, E.G.,
770 Bustamante, C.D., Cai, H., Cao, H., Durbin, R.M., Gerry, N.P., Gharani, N., Gibbs,
771 R.A., Gignoux, C.R., Gravel, S., Henn, B., Jones, D., Jorde, L., Kaye, J.S., Keinan,
772 A., Kent, A., Kerasidou, A., Li, Y., Mathias, R., McVean, G.A., Moreno-Estrada,
773 A., Ossorio, P.N., Parker, M., Resch, A.M., Rotimi, C.N., Royal, C.D., Sandoval,
774 K., Su, Y., Sudbrak, R., Tian, Z., Tishkoff, S., Toji, L.H., Tyler-Smith, C., Via, M.,
775 Wang, Y., Yang, H., Yang, L., Zhu, J., Bodmer, W., Bedoya, G., Ruiz-Linares, A.,
776 Cai, Z., Gao, Y., Chu, J., Peltonen, L., Garcia-Montero, A., Orfao, A., Dutil, J.,
777 Martinez-Cruzado, J.C., Oleksyk, T.K., Barnes, K.C., Mathias, R.A., Hennis, A.,
778 Watson, H., McKenzie, C., Qadri, F., LaRocque, R., Sabeti, P.C., Zhu, J., Deng, X.,
779 Sabeti, P.C., Asogun, D., Folarin, O., Happi, C., Omoniwa, O., Stremlau, M.,

780 Tariyal, R., Jallow, M., Sisay Joof, F., Corrah, T., Rockett, K., Kwiatkowski, D.,
781 Kooner, J., Tịnh Hiê`n, T., Dunstan, S.J., Thuy Hang, N., Fonnie, R., Garry, R.,
782 Kanneh, L., Moses, L., Sabeti, P.C., Schieffelin, J., Grant, D.S., Gallo, C., Poletti,
783 G., Saleheen, D., Rasheed, A., Brooks, L.D., Felsenfeld, A.L., McEwen, J.E.,
784 Vaydylevich, Y., Green, E.D., Duncanson, A., Dunn, M., Schloss, J.A., Wang, J.,
785 Yang, H., Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Min Kang, H.,
786 Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., Abecasis, G.R., 2015. A
787 global reference for human genetic variation. Nature 526, 68–74.
788 doi:10.1038/nature15393
789 Badr, G.G., Witt-Engerström, I., Hagberg, B., 1987. Brain stem and spinal cord
790 impairment in Rett syndrome: somatosensory and auditory evoked responses
791 investigations. Brain Dev. 9, 517–22.
792 Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A.,
793 Shendure, J., 2011. Exome sequencing as a tool for Mendelian disease gene
794 discovery. Nat. Rev. Genet. 12, 745–755. doi:10.1038/nrg3031
795 Baruah, P.S., Beauchemin, M., Parker, J.A., Bertrand, R., 2017. Expression of human
796 Bcl-xL (Ser49) and (Ser62) mutants in Caenorhabditis elegans causes germline
797 defects and aneuploidy. PLoS One 12, e0177413. doi:10.1371/journal.pone.0177413
798 Ben-Shalom, R., Keeshen, C.M., Berrios, K.N., An, J.Y., Sanders, S.J., Bender, K.J.,
799 2017. Opposing Effects on NaV1.2 Function Underlie Differences Between SCN2A
800 Variants Observed in Individuals With Autism Spectrum Disorder or Infantile
801 Seizures. Biol. Psychiatry. doi:10.1016/j.biopsych.2017.01.009
802 Bend, E.G., Si, Y., Stevenson, D.A., Bayrak-Toydemir, P., Newcomb, T.M., Jorgensen,
803 E.M., Swoboda, K.J., 2016. NALCN channelopathies. Neurology 87, 1131–1139.
804 doi:10.1212/WNL.0000000000003095
805 Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H.,
806 Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Res.
807 28, 235–42.
808 Brenner, S., 1974. The genetics of Caenorhabditis elegans. Genetics 77, 71–94.
809 Bulger, D.A., Fukushige, T., Yun, S., Semple, R.K., Hanover, J.A., Krause, M.W., 2017.
810 Caenorhabditis elegans DAF-2 as a Model for Human Insulin Receptoropathies. G3
811 (Bethesda). 7, 257–268. doi:10.1534/g3.116.037184
812 Calarco, J.A., Norris, A.D., 2018. Synthetic Genetic Interaction (CRISPR-SGI) Profiling
813 in Caenorhabditis elegans. Bio-protocol 8. doi:10.21769/BioProtoc.2756
814 Canning, P., Park, K., Gonçalves, J., Li, C., Howard, C.J., Sharpe, T.D., Holt, L.J.,
815 Pelletier, L., Bullock, A.N., Leroux, M.R., 2018. CDKL Family Kinases Have
816 Evolved Distinct Structural Features and Ciliary Function. Cell Rep. 22, 885–894.
817 doi:10.1016/j.celrep.2017.12.083
818 Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D.,
819 Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., Coban Akdemir, Z.H.,
820 Doheny, K., Scott, A.F., Avramopoulos, D., Chakravarti, A., Hoover-Fong, J.,
821 Mathews, D., Witmer, P.D., Ling, H., Hetrick, K., Watkins, L., Patterson, K.E.,
822 Reinier, F., Blue, E., Muzny, D., Kircher, M., Bilguvar, K., López-Giráldez, F.,
823 Sutton, V.R., Tabor, H.K., Leal, S.M., Gunel, M., Mane, S., Gibbs, R.A.,
824 Boerwinkle, E., Hamosh, A., Shendure, J., Lupski, J.R., Lifton, R.P., Valle, D.,
825 Nickerson, D.A., Bamshad, M.J., 2015. The Genetic Basis of Mendelian

826 Phenotypes: Discoveries, Challenges, and Opportunities. Am. J. Hum. Genet. 97,
827 199–215. doi:10.1016/j.ajhg.2015.06.009
828 Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang,
829 W., Marraffini, L.A., Zhang, F., 2013. Multiplex Genome Engineering Using
830 CRISPR/Cas Systems. Science (80-. ). 339.
831 DeLano, W.L., 2002. The PyMOL molecular graphics system. [WWW Document].
832 DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013. Genome
833 engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic
834 Acids Res. 41, 4336–4343. doi:10.1093/nar/gkt135
835 Dickinson, D.J., Goldstein, B., 2016. CRISPR-Based Methods for Caenorhabditis elegans
836 Genome Engineering. Genetics 202, 885–901. doi:10.1534/genetics.115.182162
837 Dickinson, D.J., Pani, A.M., Heppert, J.K., Higgins, C.D., Goldstein, B., 2015.
838 Streamlined Genome Engineering with a Self-Excising Drug Selection Cassette.
839 Genetics 200, 1035–1049. doi:10.1534/genetics.115.178335
840 Dickinson, D.J., Ward, J.D., Reiner, D.J., Goldstein, B., 2013. Engineering the
841 Caenorhabditis elegans genome using Cas9-triggered homologous recombination.
842 Nat. Methods 10, 1028–1034. doi:10.1038/nmeth.2641
843 Doudna, J.A., Charpentier, E., 2014. The new frontier of genome engineering with
844 CRISPR-Cas9. Science (80-. ). 346, 1258096–1258096.
845 doi:10.1126/science.1258096
846 Dunham, M.J., Fowler, D.M., 2013. Contemporary, yeast-based approaches to
847 understanding human genetic variation. Curr. Opin. Genet. Dev. 23, 658–64.
848 doi:10.1016/j.gde.2013.10.001
849 Dupuy, D., Bertin, N., Hidalgo, C.A., Venkatesan, K., Tu, D., Lee, D., Rosenberg, J.,
850 Svrzikapa, N., Blanc, A., Carnec, A., Carvunis, A.-R., Pulak, R., Shingles, J., Reece-
851 Hoyes, J., Hunt-Newbury, R., Viveiros, R., Mohler, W.A., Tasan, M., Roth, F.P., Le
852 Peuch, C., Hope, I.A., Johnsen, R., Moerman, D.G., Barabási, A.-L., Baillie, D.,
853 Vidal, M., 2007. Genome-scale analysis of in vivo spatiotemporal promoter activity
854 in Caenorhabditis elegans. Nat. Biotechnol. 25, 663–668. doi:10.1038/nbt1305
855 Findlay, G.M., Boyle, E.A., Hause, R.J., Klein, J.C., Shendure, J., 2014. Saturation
856 editing of genomic regions by multiplex homology-directed repair. Nature 513, 120–
857 3. doi:10.1038/nature13695
858 Findlay, G.M., Daza, R.M., Martin, B., Zhang, M.D., Leith, A.P., Gasperini, M., Janizek,
859 J.D., Huang, X., Starita, L.M., Shendure, J., 2018. Accurate functional classification
860 of thousands of BRCA1 variants with saturation genome editing. bioRxiv 294520.
861 doi:10.1101/294520
862 Fowler, D.M., Araya, C.L., Fleishman, S.J., Kellogg, E.H., Stephany, J.J., Baker, D.,
863 Fields, S., 2010. High-resolution mapping of protein sequence-function
864 relationships. Nat. Methods 7, 741–746. doi:10.1038/nmeth.1492
865 Fowler, D.M., Fields, S., 2014. Deep mutational scanning: a new style of protein science.
866 Nat. Methods 11, 801–807. doi:10.1038/nmeth.3027
867 Friedland, A.E., Tzur, Y.B., Esvelt, K.M., Colaiácovo, M.P., Church, G.M., Calarco,
868 J.A., 2013. Heritable genome editing in C. elegans via a CRISPR-Cas9 system. Nat.
869 Methods 10, 741–743. doi:10.1038/nmeth.2532
870 Frøkjær-Jensen, C., Davis, M.W., Ailion, M., Jorgensen, E.M., 2012. Improved Mos1-
871 mediated transgenesis in C. elegans. Nat. Methods 9, 117–8.

doi:10.1038/nmeth.1865

Frøkjær-Jensen, C., Davis, M.W., Sarov, M., Taylor, J., Flibotte, S., LaBella, M., Pozniakovsky, A., Moerman, D.G., Jorgensen, E.M., 2014. Random and targeted transgene insertion in Caenorhabditis elegans using a modified Mos1 transposon. Nat. Methods 11, 529–34. doi:10.1038/nmeth.2889

Frøkjær-Jensen, C., Wayne Davis, M., Hopkins, C.E., Newman, B.J., Thummel, J.M., Olesen, S.-P., Grunnet, M., Jorgensen, E.M., 2008. Single-copy insertion of transgenes in Caenorhabditis elegans. Nat. Genet. 40, 1375–1383. doi:10.1038/ng.248

Gahl, W.A., Mulvihill, J.J., Toro, C., Markello, T.C., Wise, A.L., Ramoni, R.B., Adams, D.R., Tifft, C.J., 2016. The NIH Undiagnosed Diseases Program and Network: Applications to modern medicine. Mol. Genet. Metab. 117, 393–400. doi:10.1016/j.ymgme.2016.01.007

Geschwind, D.H., Flint, J., 2015. Genetics and genomics of psychiatric disease. Science 349, 1489–94. doi:10.1126/science.aaa8954

Gonzaga-Jauregui, C., Lupski, J.R., Gibbs, R.A., 2012. Human Genome Sequencing in Health and Disease. Annu. Rev. Med. 63, 35–61. doi:10.1146/annurev-med-051010-162644

Gratz, S.J., Cummings, A.M., Nguyen, J.N., Hamm, D.C., Donohue, L.K., Harrison, M.M., Wildonger, J., O'Connor-Giles, K.M., 2013. Genome Engineering of *Drosophila* with the CRISPR RNA-Guided Cas9 Nuclease. Genetics 194, 1029–1035. doi:10.1534/genetics.113.152710

Grimm, D.G., Azencott, C.-A., Aicheler, F., Gieraths, U., MacArthur, D.G., Samocha, K.E., Cooper, D.N., Stenson, P.D., Daly, M.J., Smoller, J.W., Duncan, L.E., Borgwardt, K.M., 2015. The Evaluation of Tools Used to Predict the Impact of Missense Variants Is Hindered by Two Types of Circularity. Hum. Mutat. 36, 513–523. doi:10.1002/humu.22768

Hamza, A., Tammpere, E., Kofoed, M., Keong, C., Chiang, J., Giaever, G., Nislow, C., Hieter, P., 2015. Complementation of Yeast Genes with Human Genes as an Experimental Platform for Functional Testing of Human Genetic Variants. Genetics 201, 1263–74. doi:10.1534/genetics.115.181099

Hobert, J.A., Embacher, R., Mester, J.L., Frazier, T.W., Eng, C., 2014. Biochemical screening and PTEN mutation analysis in individuals with autism spectrum disorders and macrocephaly. Eur. J. Hum. Genet. 22, 273–6. doi:10.1038/ejhg.2013.114

Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., McKay, S., Okada, H.M., Pan, J., Schulz, A.K., Tu, D., Wong, K., Zhao, Z., Alexeyenko, A., Burglin, T., Sonnhammer, E., Schnabel, R., Jones, S.J., Marra, M.A., Baillie, D.L., Moerman, D.G., 2007. High-Throughput In Vivo Analysis of Gene Expression in Caenorhabditis elegans. PLoS Biol. 5, e237. doi:10.1371/journal.pbio.0050237

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.-R.J., Joung, J.K., 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat. Biotechnol. 31, 227–229. doi:10.1038/nbt.2501

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial

immunity. Science 337, 816–21. doi:10.1126/science.1225829

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., Doudna, J., 2013. RNA-programmed genome editing in human cells. Elife 2, e00471. doi:10.7554/eLife.00471

Kachroo, A.H., Laurent, J.M., Yellman, C.M., Meyer, A.G., Wilke, C.O., Marcotte, E.M., 2015. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science (80-. ). 348, 921–925. doi:10.1126/science.aaa0769

Kaletta, T., Hengartner, M.O., 2006. Finding function in novel targets: C. elegans as a model organism. Nat. Rev. Drug Discov. 5, 387–399. doi:10.1038/nrd2031

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., Sternberg, M.J.E., 2015. The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. 10, 845–58. doi:10.1038/nprot.2015.053

Kelly, W.G., Xu, S., Montgomery, M.K., Fire, A., 1997. Distinct requirements for somatic and germline expression of a generally expressed Caernorhabditis elegans gene. Genetics 146, 227–38.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., Shendure, J., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 46, 310–5. doi:10.1038/ng.2892

Kuroyanagi, H., Ohno, G., Sakane, H., Maruoka, H., Hagiwara, M., 2010. Visualization and genetic analysis of alternative splicing regulation in vivo using fluorescence reporters in transgenic Caenorhabditis elegans. Nat. Protoc. 5, 1495–1517. doi:10.1038/nprot.2010.107

Lai, C.H., Chou, C.Y., Ch'ang, L.Y., Liu, C.S., Lin, W., 2000a. Identification of novel human genes evolutionarily conserved in Caenorhabditis elegans by comparative proteomics. Genome Res. 10, 703–13.

Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., Maglott, D.R., 2014. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 42, D980-5. doi:10.1093/nar/gkt1113

Lee, J.O., Yang, H., Georgescu, M.M., Di Cristofano, A., Maehama, T., Shi, Y., Dixon, J.E., Pandolfi, P., Pavletich, N.P., 1999. Crystal structure of the PTEN tumor suppressor: implications for its phosphoinositide phosphatase activity and membrane association. Cell 99, 323–34.

Lee, P.-T., Zirin, J., Kanca, O., Lin, W.-W., Schulze, K.L., Li-Kroeger, D., Tao, R., Devereaux, C., Hu, Y., Chung, V., Fang, Y., He, Y., Pan, H., Ge, M., Zuo, Z., Housden, B.E., Mohr, S.E., Yamamoto, S., Levis, R.W., Spradling, A.C., Perrimon, N., Bellen, H.J., 2018. A gene-specificT2A-GAL4library forDrosophila. Elife 7. doi:10.7554/eLife.35574

Lee, R.Y.N., Howe, K.L., Harris, T.W., Arnaboldi, V., Cain, S., Chan, J., Chen, W.J., Davis, P., Gao, S., Grove, C., Kishore, R., Muller, H.-M., Nakamura, C., Nuin, P., Paulini, M., Raciti, D., Rodgers, F., Russell, M., Schindelman, G., Tuli, M.A., Van Auken, K., Wang, Q., Williams, G., Wright, A., Yook, K., Berriman, M., Kersey, P., Schedl, T., Stein, L., Sternberg, P.W., 2018. WormBase 2017: molting into a new stage. Nucleic Acids Res. 46, D869–D874. doi:10.1093/nar/gkx998

Lehner, B., 2013. Genotype to phenotype: lessons from model organisms for human genetics. Nat. Rev. Genet. 14, 168–178. doi:10.1038/nrg3404

Lek, M., Karczewski, K.J., Minikel, E. V, Samocha, K.E., Banks, E., Fennell, T.,

964 O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T.,
965 Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J.,
966 Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R.,
967 Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D.,
968 Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M.,
969 Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K.,
970 Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B.,
971 Won, H.-H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J.,
972 Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman,
973 C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D.,
974 McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M.,
975 Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G.,
976 Daly, M.J., MacArthur, D.G., Exome Aggregation Consortium, 2016. Analysis of
977 protein-coding genetic variation in 60,706 humans. Nature 536, 285–91.
978 doi:10.1038/nature19057
979 Levitan, D., Doyle, T.G., Brousseau, D., Lee, M.K., Thinakaran, G., Slunt, H.H., Sisodia,
980 S.S., Greenwald, I., 1996. Assessment of normal and mutant human presenilin
981 function in Caenorhabditis elegans. Proc. Natl. Acad. Sci. U. S. A. 93, 14940–4.
982 Levitan, D., Greenwald, I., 1995. Facilitation of lin-12-mediated signalling by sel-12, a
983 Caenorhabditis elegans S182 Alzheimer's disease gene. Nature 377, 351–354.
984 doi:10.1038/377351a0
985 Li, D., Qiu, Z., Shao, Y., Chen, Y., Guan, Y., Liu, M., Li, Y., Gao, N., Wang, L., Lu, X.,
986 Zhao, Y., Liu, M., 2013. Heritable gene targeting in the mouse and rat using a
987 CRISPR-Cas system. Nat. Biotechnol. 31, 681–3. doi:10.1038/nbt.2661
988 Li, J., Yen, C., Liaw, D., Podsypanina, K., Bose, S., Wang, S.I., Puc, J., Miliaresis, C.,
989 Rodgers, L., McCombie, R., Bigner, S.H., Giovanella, B.C., Ittmann, M., Tycko, B.,
990 Hibshoosh, H., Wigler, M.H., Parsons, R., 1997. PTEN, a putative protein tyrosine
991 phosphatase gene mutated in human brain, breast, and prostate cancer. Science 275,
992 1943–7.
993 Liaw, D., Marsh, D.J., Li, J., Dahia, P.L., Wang, S.I., Zheng, Z., Bose, S., Call, K.M.,
994 Tsou, H.C., Peacocke, M., Eng, C., Parsons, R., 1997. Germline mutations of the
995 PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome.
996 Nat. Genet. 16, 64–7. doi:10.1038/ng0597-64
997 Liaw, D., Marsh, D.J., Li, J., Dahia, P.L.M., Wang, S.I., Zheng, Z., Bose, S., Call, K.M.,
998 Tsou, H.C., Peacoke, M., Eng, C., Parsons, R., 1997. Germline mutations of the
999 PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome.
1000 Nat. Genet. 16, 64–67. doi:10.1038/ng0597-64
1001 Lim, C.-S., Kang, X., Mirabella, V., Zhang, H., Bu, Q., Araki, Y., Hoang, E.T., Wang, S.,
1002 Shen, Y., Choi, S., Kaang, B.-K., Chang, Q., Pang, Z.P., Huganir, R.L., Zhu, J.J.,
1003 2017. BRaf signaling principles unveiled by large-scale human mutation analysis
1004 with a rapid lentivirus-based gene replacement method. Genes Dev. 31, 537–552.
1005 doi:10.1101/gad.294413.116
1006 Liu, J., Chin-Sang, I.D., 2015. C. elegans as a model to study PTEN's regulation and
1007 function. Methods 77–78, 180–190. doi:10.1016/j.ymeth.2014.12.009
1008 López-Bigas, N., Ouzounis, C.A., 2004. Genome-wide identification of genes likely to be
1009 involved in human genetic disease. Nucleic Acids Res. 32, 3108–14.

1010 doi:10.1093/nar/gkh605
1011 Luo, X., Rosenfeld, J.A., Yamamoto, S., Harel, T., Zuo, Z., Hall, M., Wierenga, K.J.,
1012 Pastore, M.T., Bartholomew, D., Delgado, M.R., Rotenberg, J., Lewis, R.A.,
1013 Emrick, L., Bacino, C.A., Eldomery, M.K., Coban Akdemir, Z., Xia, F., Yang, Y.,
1014 Lalani, S.R., Lotze, T., Lupski, J.R., Lee, B., Bellen, H.J., Wangler, M.F., Members
1015 of the UDN, 2017. Clinically severe CACNA1A alleles affect synaptic function and
1016 neurodegeneration differentially. PLoS Genet. 13, e1006905.
1017 doi:10.1371/journal.pgen.1006905
1018 Maehama, T., Dixon, J.E., 1998. The tumor suppressor, PTEN/MMAC1,
1019 dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-
1020 trisphosphate. J. Biol. Chem. 273, 13375–8.
1021 Majithia, A.R., Tsuda, B., Agostini, M., Gnanapradeepan, K., Rice, R., Peloso, G., Patel,
1022 K.A., Zhang, X., Broekema, M.F., Patterson, N., Duby, M., Sharpe, T., Kalkhoven,
1023 E., Rosen, E.D., Barroso, I., Ellard, S., Kathiresan, S., O'Rahilly, S., Chatterjee, K.,
1024 Florez, J.C., Mikkelsen, T., Savage, D.B., Altshuler, D., Mikkelsen, T., Savage,
1025 D.B., Altshuler, D., 2016. Prospective functional classification of all possible
1026 missense variants in PPARG. Nat. Genet. 48, 1570–1575. doi:10.1038/ng.3700
1027 Manolio, T.A., Fowler, D.M., Starita, L.M., Haendel, M.A., MacArthur, D.G., Biesecker,
1028 L.G., Worthey, E., Chisholm, R.L., Green, E.D., Jacob, H.J., McLeod, H.L., Roden,
1029 D., Rodriguez, L.L., Williams, M.S., Cooper, G.M., Cox, N.J., Herman, G.E.,
1030 Kingsmore, S., Lo, C., Lutz, C., MacRae, C.A., Nussbaum, R.L., Ordovas, J.M.,
1031 Ramos, E.M., Robinson, P.N., Rubinstein, W.S., Seidman, C., Stranger, B.E., Wang,
1032 H., Westerfield, M., Bult, C., 2017. Bedside Back to Bench: Building Bridges
1033 between Basic and Clinical Genomic Research. Cell 169, 6–12.
1034 doi:10.1016/j.cell.2017.03.005
1035 Matreyek, K.A., Starita, L.M., Stephany, J.J., Martin, B., Chiasson, M.A., Gray, V.E.,
1036 Kircher, M., Khechaduri, A., Dines, J.N., Hause, R.J., Bhatia, S., Evans, W.E.,
1037 Relling, M. V, Yang, W., Shendure, J., Fowler, D.M., 2018. Multiplex Assessment
1038 of Protein Variant Abundance by Massively Parallel Sequencing. bioRxiv 211011.
1039 doi:10.1101/211011
1040 McBride, K.L., Varga, E.A., Pastore, M.T., Prior, T.W., Manickam, K., Atkin, J.F.,
1041 Herman, G.E., 2010. Confirmation study of PTEN mutations among individuals
1042 with autism or developmental delays/mental retardation and macrocephaly. Autism
1043 Res. 3, 137–41. doi:10.1002/aur.132
1044 McDiarmid, T.A., Bernardos, A.C., Rankin, C.H., 2017. Habituation is altered in
1045 neuropsychiatric disorders - A comprehensive review with recommendations for
1046 experimental design and analysis. Neurosci. Biobehav. Rev. 80, 286–305.
1047 doi:10.1016/j.neubiorev.2017.05.028
1048 McDiarmid, T.A., Yu, A.J., Rankin, C.H., 2018. Beyond the response-High throughput
1049 behavioral analyses to link genome to phenome in *Caenorhabditis elegans*. Genes,
1050 Brain Behav. e12437. doi:10.1111/gbb.12437
1051 Merritt, C., Seydoux, G., 2010. Transgenic solutions for the germline. WormBook 1–21.
1052 doi:10.1895/wormbook.1.148.1
1053 Metzker, M.L., 2010. Sequencing technologies — the next generation. Nat. Rev. Genet.
1054 11, 31–46. doi:10.1038/nrg2626
1055 Mighell, T.L., Evans-Dutson, S., O'Roak, B.J., 2018. A saturation mutagenesis approach

to understanding PTEN lipid phosphatase activity and genotype-phenotypes relationships. bioRxiv 255265. doi:10.1101/255265

Mihaylova, V.T., Borland, C.Z., Manjarrez, L., Stern, M.J., Sun, H., 1999. The PTEN tumor suppressor homolog in Caenorhabditis elegans regulates longevity and dauer formation in an insulin receptor-like signaling pathway. Proc. Natl. Acad. Sci. U. S. A. 96, 7427–32.

Miosge, L.A., Field, M.A., Sontani, Y., Cho, V., Johnson, S., Palkova, A., Balakishnan, B., Liang, R., Zhang, Y., Lyon, S., Beutler, B., Whittle, B., Bertram, E.M., Enders, A., Goodnow, C.C., Andrews, T.D., 2015. Comparison of predicted and actual consequences of missense mutations. Proc. Natl. Acad. Sci. 112, E5189–E5198. doi:10.1073/pnas.1511585112

Murphy, C.T., McCarroll, S.A., Bargmann, C.I., Fraser, A., Kamath, R.S., Ahringer, J., Li, H., Kenyon, C., 2003. Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. Nature 424, 277–283. doi:10.1038/nature01789

Myers, M.P., Pass, I., Batty, I.H., Van der Kaay, J., Stolarov, J.P., Hemmings, B.A., Wigler, M.H., Downes, C.P., Tonks, N.K., 1998. The lipid phosphatase activity of PTEN is critical for its tumor supressor function. Proc. Natl. Acad. Sci. U. S. A. 95, 13513–8.

Myers, M.P., Stolarov, J.P., Eng, C., Li, J., Wang, S.I., Wigler, M.H., Parsons, R., Tonks, N.K., 1997. P-TEN, the tumor suppressor from human chromosome 10q23, is a dual-specificity phosphatase. Proc. Natl. Acad. Sci. U. S. A. 94, 9052–7.

Need, A.C., Shashi, V., Hitomi, Y., Schoch, K., Shianna, K. V, McDonald, M.T., Meisler, M.H., Goldstein, D.B., 2012. Clinical application of exome sequencing in undiagnosed genetic conditions. J. Med. Genet. 49, 353–361. doi:10.1136/jmedgenet-2012-100819

Ng, S.B., Buckingham, K.J., Lee, C., Bigham, A.W., Tabor, H.K., Dent, K.M., Huff, C.D., Shannon, P.T., Jabs, E.W., Nickerson, D.A., Shendure, J., Bamshad, M.J., 2010. Exome sequencing identifies the cause of a mendelian disorder. Nat. Genet. 42, 30–35. doi:10.1038/ng.499

Norris, A.D., Gracida, X., Calarco, J.A., 2017. CRISPR-mediated genetic interaction profiling identifies RNA binding proteins controlling metazoan fitness. Elife 6. doi:10.7554/eLife.28129

Norris, A.D., Kim, H.-M., Colaiácovo, M.P., Calarco, J.A., 2015. Efficient Genome Editing in Caenorhabditis elegans with a Toolkit of Dual-Marker Selection Cassettes. Genetics 201.

O'Roak, B.J., Vives, L., Fu, W., Egertson, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K., Munson, J., Hiatt, J.B., Turner, E.H., Levy, R., O'Day, D.R., Krumm, N., Coe, B.P., Martin, B.K., Borenstein, E., Nickerson, D.A., Mefford, H.C., Doherty, D., Akey, J.M., Bernier, R., Eichler, E.E., Shendure, J., 2012. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science 338, 1619–22. doi:10.1126/science.1227764

Ogg, S., Ruvkun, G., 1998. The C. elegans PTEN homolog, DAF-18, acts in the insulin receptor-like metabolic signaling pathway. Mol. Cell 2, 887–93.

Orefice, L.L., Zimmerman, A.L., Chirila, A.M., Sleboda, S.J., Head, J.P., Ginty, D.D., 2016. Peripheral Mechanosensory Neuron Dysfunction Underlies Tactile and Behavioral Deficits in Mouse Models of ASDs. Cell 166, 299–313.

1102 doi:10.1016/j.cell.2016.05.033
1103 Orrico, A., Galli, L., Buoni, S., Orsi, A., Vonella, G., Sorrentino, V., 2009. Novel PTEN
1104 mutations in neurodevelopmental disorders and macrocephaly. Clin. Genet. 75, 195–
1105 198. doi:10.1111/j.1399-0004.2008.01074.x
1106 Ozes, O.N., Akca, H., Mayo, L.D., Gustin, J.A., Maehama, T., Dixon, J.E., Donner, D.B.,
1107 2001. A phosphatidylinositol 3-kinase/Akt/mTOR pathway mediates and PTEN
1108 antagonizes tumor necrosis factor inhibition of insulin signaling through insulin
1109 receptor substrate-1. Proc. Natl. Acad. Sci. 98, 4640–4645.
1110 doi:10.1073/pnas.051042298
1111 Pierce, S.B., Stewart, M.D., Gulsuner, S., Walsh, T., Dhall, A., McClellan, J.M., Klevit,
1112 R.E., King, M.-C., 2018. De novo mutation inRING1with epigenetic effects on
1113 neurodevelopment. Proc. Natl. Acad. Sci. U. S. A. 115, 1558–1563.
1114 doi:10.1073/pnas.1721290115
1115 Prior, H., Jawad, A.K., MacConnachie, L., Beg, A.A., 2017. Highly Efficient, Rapid and
1116 Co-CRISPR Independent Genome Editing in *Caenorhabditis elegans*.
1117 G3: Genes|Genomes|Genetics g3.300216.2017.
1118 doi:10.1534/g3.117.300216
1119 Ranganathan, R., Sawin, E.R., Trent, C., Horvitz, H.R., 2001. Mutations in the
1120 Caenorhabditis elegans serotonin reuptake transporter MOD-5 reveal serotonin-
1121 dependent and -independent activities of fluoxetine. J. Neurosci. 21, 5871–84.
1122 Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F., Colombo, J.,
1123 Coppola, G., Geyer, M.A., Glanzman, D.L., Marsland, S., McSweeney, F.K.,
1124 Wilson, D.A., Wu, C.-F., Thompson, R.F., 2009. Habituation revisited: an updated
1125 and revised description of the behavioral characteristics of habituation. Neurobiol.
1126 Learn. Mem. 92, 135–8. doi:10.1016/j.nlm.2008.09.012
1127 Rankin, C.H., Beck, C.D., Chiba, C.M., 1990. Caenorhabditis elegans: a new model
1128 system for the study of learning and memory. Behav. Brain Res. 37, 89–92.
1129 Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W.,
1130 Hegde, M., Lyon, E., Spector, E., Voelkerding, K., Rehm, H.L., ACMG Laboratory
1131 Quality Assurance Committee, 2015. Standards and guidelines for the interpretation
1132 of sequence variants: a joint consensus recommendation of the American College of
1133 Medical Genetics and Genomics and the Association for Molecular Pathology.
1134 Genet. Med. 17, 405–423. doi:10.1038/gim.2015.30
1135 Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E.,
1136 Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S., Goldberg, A.P., Jinlu, C., Keaney,
1137 J.F., Klei, L., Mandell, J.D., Moreno-De-Luca, D., Poultney, C.S., Robinson, E.B.,
1138 Smith, L., Solli-Nowlan, T., Su, M.Y., Teran, N.A., Walker, M.F., Werling, D.M.,
1139 Beaudet, A.L., Cantor, R.M., Fombonne, E., Geschwind, D.H., Grice, D.E., Lord,
1140 C., Lowe, J.K., Mane, S.M., Martin, D.M., Morrow, E.M., Talkowski, M.E.,
1141 Sutcliffe, J.S., Walsh, C.A., Yu, T.W., Ledbetter, D.H., Martin, C.L., Cook, E.H.,
1142 Buxbaum, J.D., Daly, M.J., Devlin, B., Roeder, K., State, M.W., State, M.W., 2015.
1143 Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71
1144 Risk Loci. Neuron 87, 1215–1233. doi:10.1016/j.neuron.2015.09.016
1145 Shaye, D.D., Greenwald, I., 2011. OrthoList: A Compendium of C. elegans Genes with
1146 Human Orthologs. PLoS One 6, e20085. doi:10.1371/journal.pone.0020085
1147 Solari, F., Bourbon-Piffaut, A., Masse, I., Payrastre, B., Chan, A.M.-L., Billaud, M.,

2005. The human tumour suppressor PTEN regulates longevity and dauer formation in Caenorhabditis elegans. Oncogene 24, 20–27. doi:10.1038/sj.onc.1207978

Sorkaç, A., Alcantara, I.C., Hart, A.C., 2016. In Vivo Modelling of ATP1A3 G316S-Induced Ataxia in C. elegans Using CRISPR/Cas9-Mediated Homologous Recombination Reveals Dominant Loss of Function Defects. PLoS One 11, e0167963. doi:10.1371/journal.pone.0167963

Starita, L.M., Ahituv, N., Dunham, M.J., Kitzman, J.O., Roth, F.P., Seelig, G., Shendure, J., Fowler, D.M., 2017. Variant Interpretation: Functional Assays to the Rescue. Am. J. Hum. Genet. 101, 315–325. doi:10.1016/j.ajhg.2017.07.014

Stessman, H.A.F., Xiong, B., Coe, B.P., Wang, T., Hoekzema, K., Fenckova, M., Kvarnung, M., Gerdts, J., Trinh, S., Cosemans, N., Vives, L., Lin, J., Turner, T.N., Santen, G., Ruivenkamp, C., Kriek, M., van Haeringen, A., Aten, E., Friend, K., Liebelt, J., Barnett, C., Haan, E., Shaw, M., Gecz, J., Anderlid, B.-M., Nordgren, A., Lindstrand, A., Schwartz, C., Kooy, R.F., Vandeweyer, G., Helsmoortel, C., Romano, C., Alberti, A., Vinci, M., Avola, E., Giusto, S., Courchesne, E., Pramparo, T., Pierce, K., Nalabolu, S., Amaral, D.G., Scheffer, I.E., Delatycki, M.B., Lockhart, P.J., Hormozdiari, F., Harich, B., Castells-Nobau, A., Xia, K., Peeters, H., Nordenskjöld, M., Schenck, A., Bernier, R.A., Eichler, E.E., 2017. Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat. Genet. 49, 515–526. doi:10.1038/ng.3792

Sulston, J.E., Horvitz, H.R., 1977. Post-embryonic cell lineages of the nematode, Caenorhabditis elegans. Dev. Biol. 56, 110–56.

Sulston, J.E., Schierenberg, E., White, J.G., Thomson, J.N., 1983. The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev. Biol. 100, 64–119.

Sun, S., Yang, F., Tan, G., Costanzo, M., Oughtred, R., Hirschman, J., Theesfeld, C.L., Bansal, P., Sahni, N., Yi, S., Yu, A., Tyagi, T., Tie, C., Hill, D.E., Vidal, M., Andrews, B.J., Boone, C., Dolinski, K., Roth, F.P., 2016. An extended set of yeast-based functional assays accurately identifies human disease mutations. Genome Res. 26, 670–680. doi:10.1101/gr.192526.115

Swierczek, N.A., Giles, A.C., Rankin, C.H., Kerr, R.A., 2011. High-throughput behavioral analysis in C. elegans. Nat. Methods 8, 592–8. doi:10.1038/nmeth.1625

Tamura, M., Gu, J., Matsumoto, K., Aota, S., Parsons, R., Yamada, K.M., 1998. Inhibition of cell migration, spreading, and focal adhesions by tumor suppressor PTEN. Science 280, 1614–7.

Thompson, O., Edgley, M., Strasbourger, P., Flibotte, S., Ewing, B., Adair, R., Au, V., Chaudhry, I., Fernando, L., Hutter, H., Kieffer, A., Lau, J., Lee, N., Miller, A., Raymant, G., Shen, B., Shendure, J., Taylor, J., Turner, E.H., Hillier, L.W., Moerman, D.G., Waterston, R.H., 2013. The million mutation project: A new approach to genetics in Caenorhabditis elegans. Genome Res. 23, 1749–1762. doi:10.1101/gr.157651.113

Timbers, T.A., Ardiel, E.L., McDiarmid, T.A., Lee, K., Safaei, J., Pelech, S.L., Rankin, C.H., 2017. CaMK (CMK-1) and O-GlcNAc transferase (OGT-1) modulate mechanosensory responding and habituation in an interstimulus interval-dependent manner in Caenorhabditis elegans. bioRxiv.

Tomioka, M., Adachi, T., Suzuki, H., Kunitomo, H., Schafer, W.R., Iino, Y., 2006. The insulin/PI 3-kinase pathway regulates salt chemotaxis learning in Caenorhabditis

elegans. Neuron 51, 613–25. doi:10.1016/j.neuron.2006.07.024

Troulinaki, K., Büttner, S., Marsal Cots, A., Maida, S., Meyer, K., Bertan, F., Gioran, A., Piazzesi, A., Fornarelli, A., Nicotera, P., Bano, D., 2018. WAH-1/AIF regulates mitochondrial oxidative phosphorylation in the nematodeCaenorhabditis elegans. Cell death Discov. 4, 2. doi:10.1038/s41420-017-0005-6

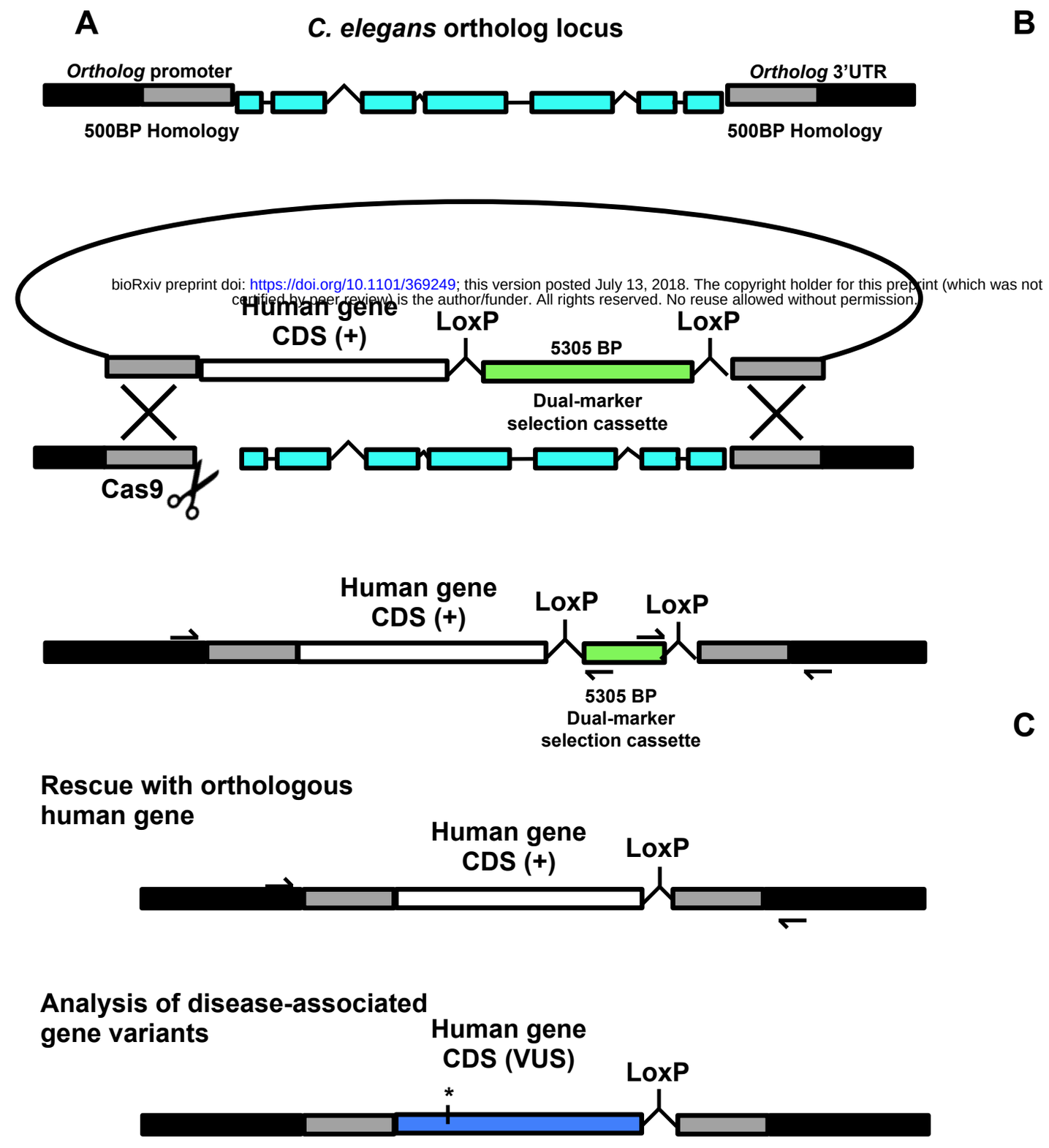
van der Voet, M., Nijhof, B., Oortveld, M.A.W., Schenck, A., 2014. Drosophila models of early onset cognitive disorders and their clinical applications. Neurosci. Biobehav. Rev. doi:10.1016/j.neubiorev.2014.01.013

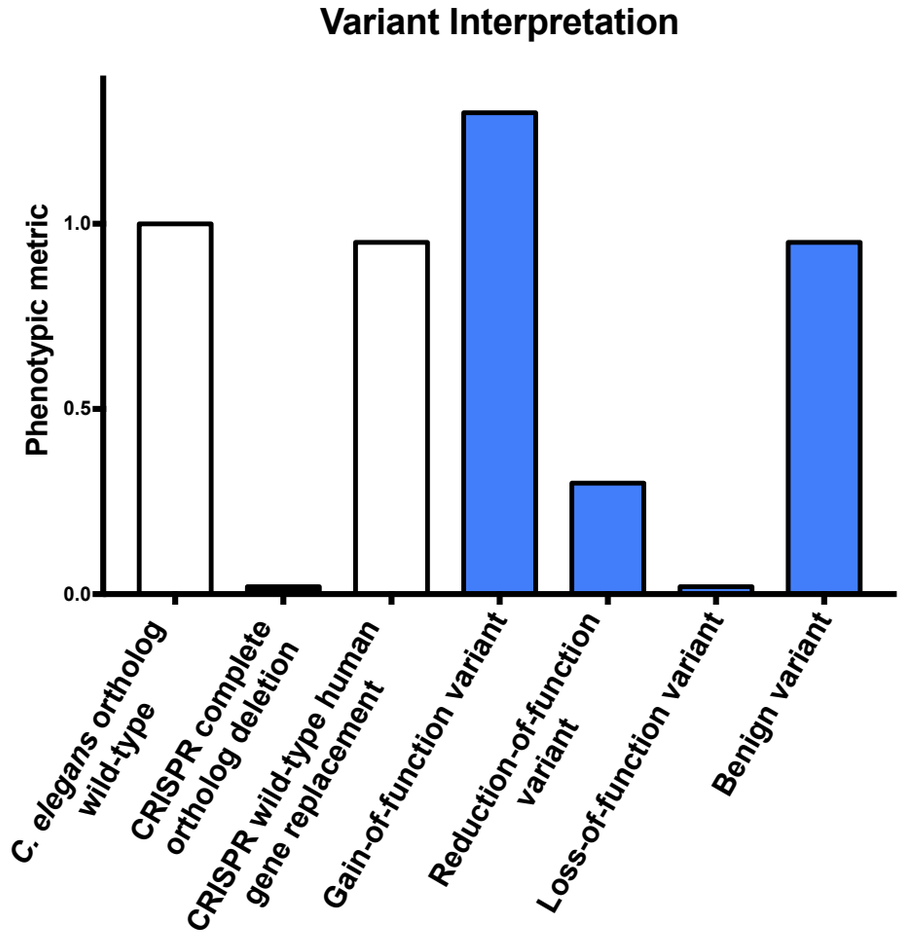
Varga, E.A., Pastore, M., Prior, T., Herman, G.E., McBride, K.L., 2009. The prevalence of PTEN mutations in a clinical pediatric cohort with autism spectrum disorders, developmental delay, and macrocephaly. Genet. Med. 11, 111–7. doi:10.1097/GIM.0b013e31818fd762

Wang, H., Liu, J., Gharib, S., Chai, C.M., Schwarz, E.M., Pokala, N., Sternberg, P.W., 2016. cGAL, a temperature-robust GAL4–UAS system for Caenorhabditis elegans. Nat. Methods 14, 145–148. doi:10.1038/nmeth.4109

Wang, J., Al-Ouran, R., Hu, Y., Kim, S.-Y., Wan, Y.-W., Wangler, M.F., Yamamoto, S., Chao, H.-T., Comjean, A., Mohr, S.E., UDN, N., Perrimon, N., Liu, Z., Bellen, H.J., Adams, D.R., Alejandro, M.E., Allard, P., Ashley, E.A., Azamian, M.S., Bacino, C.A., Balasubramanyam, A., Barseghyan, H., Beggs, A.H., Bellen, H.J., Bernstein, J.A., Bican, A., Bick, D.P., Birch, C.L., Boone, B.E., Briere, L.C., Brown, D.M., Brush, M., Burke, E.A., Burrage, L.C., Chao, K.R., Clark, G.D., Cogan, J.D., Cooper, C.M., Craigen, W.J., Davids, M., Dayal, J.G., Dell'Angelica, E.C., Dhar, S.U., Dipple, K.M., Donnell-Fink, L.A., Dorrani, N., Dorset, D.C., Draper, D.D., Dries, A.M., Eckstein, D.J., Emrick, L.T., Eng, C.M., Esteves, C., Estwick, T., Fisher, P.G., Frisby, T.S., Frost, K., Gahl, W.A., Gartner, V., Godfrey, R.A., Goheen, M., Golas, G.A., Goldstein, D.B., Gordon, M.G., Gould, S.E., Gourdine, J.-P.F., Graham, B.H., Groden, C.A., Gropman, A.L., Hackbarth, M.E., Haendel, M., Hamid, R., Hanchard, N.A., Handley, L.H., Hardee, I., Herzog, M.R., Holm, I.A., Howerton, E.M., Jacob, H.J., Jain, M., Jiang, Y., Johnston, J.M., Jones, A.L., Koehler, A.E., Koeller, D.M., Kohane, I.S., Kohler, J.N., Krasnewich, D.M., Krieg, E.L., Krier, J.B., Kyle, J.E., Lalani, S.R., Latham, L., Latour, Y.L., Lau, C.C., Lazar, J., Lee, B.H., Lee, H., Lee, P.R., Levy, S.E., Levy, D.J., Lewis, R.A., Liebendorfer, A.P., Lincoln, S.A., Loomis, C.R., Loscalzo, J., Maas, R.L., Macnamara, E.F., MacRae, C.A., Maduro, V. V., Malicdan, M.C. V., Mamounas, L.A., Manolio, T.A., Markello, T.C., Mazur, P., McCarty, A.J., McConkie-Rosell, A., McCray, A.T., Metz, T.O., Might, M., Moretti, P.M., Mulvihill, J.J., Murphy, J.L., Muzny, D.M., Nehrebecky, M.E., Nelson, S.F., Newberry, J.S., Newman, J.H., Nicholas, S.K., Novacic, D., Orange, J.S., Pallais, J.C., Palmer, C.G.S., Papp, J.C., Pena, L.D.M., Phillips, J.A., Posey, J.E., Postlethwait, J.H., Potocki, L., Pusey, B.N., Ramoni, R.B., Robertson, A.K., Rodan, L.H., Rosenfeld, J.A., Sadozai, S., Schaffer, K.E., Schoch, K., Schroeder, M.C., Scott, D.A., Sharma, P., Shashi, V., Silverman, E.K., Sinsheimer, J.S., Soldatos, A.G., Spillmann, R.C., Splinter, K., Stoler, J.M., Stong, N., Strong, K.A., Sullivan, J.A., Sweetser, D.A., Thomas, S.P., Tifft, C.J., Tolman, N.J., Toro, C., Tran, A.A., Valivullah, Z.M., Vilain, E., Waggott, D.M., Wahl, C.E., Walley, N.M., Walsh, C.A., Wangler, M.F., Warburton, M., Ward, P.A., Waters, K.M., Webb-Robertson, B.-J.M., Weech, A.A., Westerfield, M., Wheeler, M.T.,

1240 Wise, A.L., Wolfe, L.A., Worthey, E.A., Yamamoto, S., Yang, Y., Yu, G., Zornio,
1241 P.A., 2017. MARRVEL: Integration of Human and Model Organism Genetic
1242 Resources to Facilitate Functional Annotation of the Human Genome. Am. J. Hum.
1243 Genet. 100, 843–853. doi:10.1016/j.ajhg.2017.04.010
1244 Wang, S., Tang, N.H., Lara-Gonzalez, P., Zhao, Z., Cheerambathur, D.K., Prevo, B.,
1245 Chisholm, A.D., Desai, A., Oegema, K., 2017. A toolkit for GFP-mediated tissue-
1246 specific protein degradation in C. elegans. Development.
1247 Wangler, M.F., Yamamoto, S., Chao, H.-T., Posey, J.E., Westerfield, M., Postlethwait, J.,
1248 Hieter, P., Boycott, K.M., Campeau, P.M., Bellen, H.J., Bellen, H.J., 2017. Model
1249 Organisms Facilitate Rare Disease Diagnosis and Therapeutic Research. Genetics
1250 207, 9–27. doi:10.1534/genetics.117.203067
1251 Ward, S., 1973. Chemotaxis by the nematode Caenorhabditis elegans: identification of
1252 attractants and analysis of the response by use of mutants. Proc. Natl. Acad. Sci. U.
1253 S. A. 70, 817–21.
1254 Weile, J., Sun, S., Cote, A.G., Knapp, J., Verby, M., Mellor, J.C., Wu, Y., Pons, C.,
1255 Wong, C., van Lieshout, N., Yang, F., Tasan, M., Tan, G., Yang, S., Fowler, D.M.,
1256 Nussbaum, R., Bloom, J.D., Vidal, M., Hill, D.E., Aloy, P., Roth, F.P., 2017. A
1257 framework for exhaustively mapping functional missense variants. Mol. Syst. Biol.
1258 13, 957. doi:10.15252/MSB.20177908
1259 White, J.G., Southgate, E., Thomson, J.N., Brenner, S., 1986. The structure of the
1260 nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc.
1261 Lond. B. Biol. Sci. 314, 1–340.
1262 Yang, F., Sun, S., Tan, G., Costanzo, M., Hill, D.E., Vidal, M., Andrews, B.J., Boone, C.,
1263 Roth, F.P., 2017. Identifying pathogenicity of human variants via paralog-based
1264 yeast complementation. PLoS Genet. 13, e1006779.
1265 doi:10.1371/journal.pgen.1006779
1266 Yemini, E., Jucikas, T., Grundy, L.J., Brown, A.E.X., Schafer, W.R., 2013. A database of
1267 Caenorhabditis elegans behavioral phenotypes. Nat. Methods 10, 877–879.
1268 doi:10.1038/nmeth.2560
1269 Zhang, L., Ward, J.D., Cheng, Z., Dernburg, A.F., 2015. The auxin-inducible degradation
1270 (AID) system enables versatile conditional protein depletion in C. elegans.
1271 Development 142, 4374–4384. doi:10.1242/dev.129635
1272



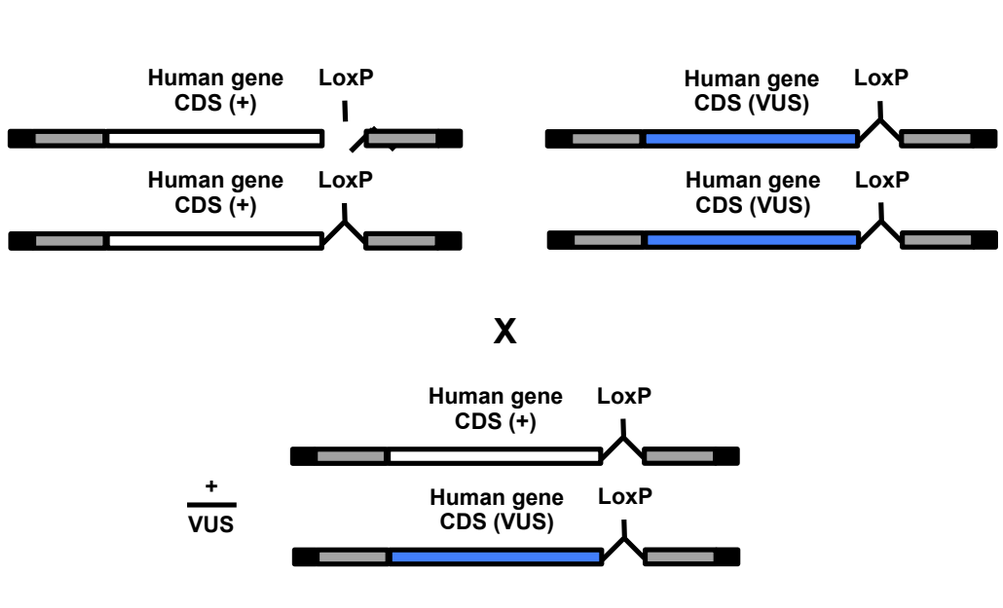
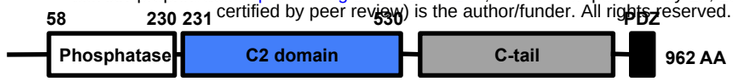



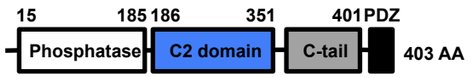

**Figure 1 | A general strategy for direct single copy replacement of C. elegans genes with human genes at the orthologs native genomic loci. A)** A schematic of the genome editing strategy. (top left) A sgRNA targets Cas9 to induce a DNA double strand break immediately downstream of the orthologs start codon. A co-injected repair template containing ~500BP homology arms targeted to the regions immediately upstream and downstream of the ortholog ORF serve as a substrate for homology directed repair. By fusing the CDS of a human gene of interest to the upstream homology arm homology directed repair integrates the human gene in place of the ortholog at a single copy in frame. A co-integrated Dual-Marker selection cassette consisting of an antibiotic resistance gene (*Prps-27::neoR::unc-54 UTR*) and a fluorescent marker (*Pmyo-2::GFP::unc-54 UTR*) greatly facilitates the identification transgenic animals without inducing morphological or phenotypic abnormalities. (middle left) initial integration deletes the entire open reading frame of the C. elegans ortholog while separating the human gene from the orthologs transcriptional terminator to inhibit expression, creating an ortholog deletion allele for phenotypic analysis (Note: cassette is not shown to scale for most human gene CDS). (bottom left) Subsequent injection of Cre Recombinase excises the selection cassette and connects the human gene to the orthologous transcriptional termination sequence such that a single copy of the human gene will now be expressed under the control of all of the orthologs 5' and 3' cis- and trans-regulatory machinery. Validation of the desired edit is preformed using standard amplification and Sanger sequencing of the target region (primer binding locations represented by half arrows in the schematic). For analysis of a human gene variant of uncertain significance (VUS) the variant of interest is incorporated into the HDR plasmid using standard *in vitro* methods such as site-directed mutagenesis and the same genome editing process is repeated using the same validated sgRNA and homology arms. **B)** Human gene replacement allows for straight-forward interpretation of variant functional effect. This process allows for: 1) initial generation and phenotypic analysis of a complete null allele in the *C. elegans* orthologous gene 2) Direct integration of the human gene to determine if the human gene can compensate for loss of the orthologous gene, measuring functional conservation, 3) structure-function analysis of the effects of variants of uncertain significance on WT gene function. **C)** This strategy allows for straightforward assessment of heterozygous alleles using standard genetic crosses.

A

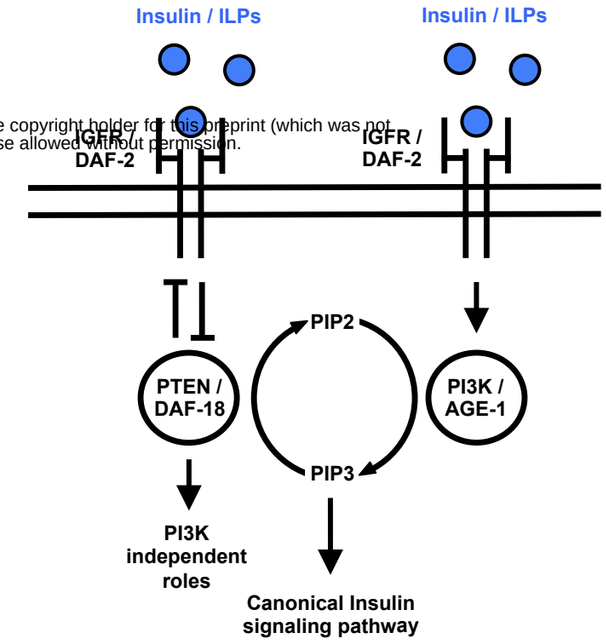
DAF-18




PTEN




B




C

## PTEN vs. DAF-18 Alignment

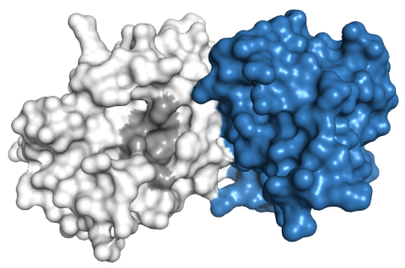
**Identity = 13.6%**
**Similarity = 20%**
**Gaps = 62.8%**

**Phosphatase domain identity = 46%**

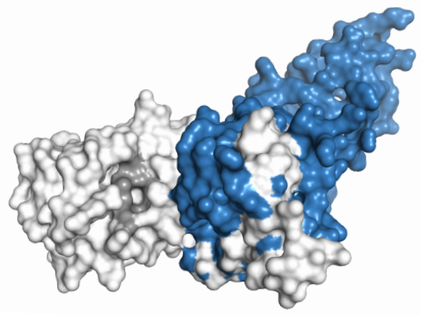
```
PTEN        ---------------------------------------------MTAIIKEIVSRNKRRY    16
DAF-18      MVTPPPDVPSTSTRSMARDLQENPNRQPGEPRVSEPYHNSIVERIRHIFRTAVSSNRCRT    60
                                                          :  *::  ** *: *

PTEN        QEDGFDLDLTYIYPNIIAMGFPAERLEGVYRNNIDDVVRFLDSKH-KNHYKIYNLCAERH    75
DAF-18      EYQNIDLDCAYITDRIIAIGYPATGIEANFRNSKVQTQQFLTRRHGKGNVKVFNLRGGYY    120
            : :.:*** :**  .***:*:**  :*. :**.  :. :**  :* *.: *::** .  :

PTEN        YDTAKFNCRVAQYPFEDHNPPQLELIKPFCEDLDQWLSEDDNHVAAIHCKAGKGRTGVMI    135
DAF-18      YDADNFDGNVICFDMTDHHPPSLELMAPFCREAKEWLEADDKHVIAVHCKAGKGRTGVMI    180
            **: :*: .*  : : **:**.***: ***.: .:**. **:** *:************

PTEN        CAYLLHRGKFLKAQEALDFYGEVRTRDKKGVTIPSQRRYVYYYSYLLKNHLDYRPVALLF    195
DAF-18      CALLIYINFYPSPRQILDYYSIIRTKNNKGVTIPSQRRYIYYYHKLRERELNYLPLRMQL    240
            ** *:: . : . :: **:*. :**:::***********:***  * :..*:* *: : :

PTEN        HKMMFETIPMFSGGTCNPQFVVCQLKVKIYS------SNSGPTRR---------------    234
DAF-18      IGVYVERPPKTWGGGSKIKVEVGNGSTILFKPDPLIISKSNHQRERATWLNNCDTPNEFD    300
              : .*  *   ** .: :. * : .. ::.      *:*.  *.

PTEN        --ED--------KFMYFEFPQ--PLPVCGDIKVEFFHK-QNKMLKKDKMFHFWVNTFFIP    281
DAF-18      TGEQKYHGFVSKRAYCFMVPEDAPVFVEGDVRIDIREIGFLKKFSDGKIGHVWFNTMFAC    360
              *:        :   * .*:  *: * **::::: .    * :...*: *.*.**:*

PTEN        GPEETSEKVENGSLCDQEIDSICSIERADNDKEYLVLTLTKNDLDKANKDKANRYFSPNF    341
daf-18      DGGLNGGHFEYVDKTQPYIGDDTSIGRKNGMR------RNETPMRKIDPETGNEFESPWQ    414
            .   .. :.*  .  :  *.. ** * :. :       .:. : * : :..*.: **

PTEN        KVKL-YFTKTVEEPSNPEASSSTSVTPDVSDNEPDHYRYSDTTD----------------    384
DAF-18      IVNPPGLEKHITEEQAMENYTNYG-------MIPPRYTISKILHEKHEKGIVKDDYNDRK    467
             *:   : * : * .  *  :. .         * :*  *.  .

PTEN        --------------------SDPENEPFDEDQHTQITKV--------------------    403
DAF-18      LPMGDKSYTESGKSGDIRGVGGPFEIPYKAEEHVLTFPVYEMDRALKSKDLNNGMKLHVV    527
                              ..* : *:. ::*.    *

PTEN        ------------------------------------------------------------    403
DAF-18      LRCVDTRDSKMMEKSEVFGNLAFHNESTRRLQALTQMNPKWRPEPCAFGSKGAEMHYPPS    587

PTEN        ------------------------------------------------------------    403
DAF-18      VRYSSNDGKYNGACSENLVSDFFEHRNIAVLNRYCRYFYKQRSTSRSRYPRKFRYCPLIK    647

PTEN        ------------------------------------------------------------    403
DAF-18      KHFYIPADTDDVDENGQPFFHSPEHYIKEQEKIDAEKAAKGIENTGPSTSGSSAPGTIKK    707

PTEN        ------------------------------------------------------------    403
DAF-18      TEASQSDKVKPATEDELPPARLPDNVRRFPVVGVDFENPEEESCEHKTVESIAGFEPLEH    767

PTEN        ------------------------------------------------------------    403
DAF-18      LFHESYHPNTAGNMLRQDYHTDSEVKIAEQEAKAFVDQLLNGQGVLQEFMKQFKVPSDNS    827

PTEN        ------------------------------------------------------------    403
DAF-18      FADYVTGQAEVFKAQIALLEQSEDFQRVQANAEEVDLEHTLGEAFERFGHVVEESNGSSK    887

PTEN        ------------------------------------------------------------    403
DAF-18      NPKALKTREQMVKETGKDTQKTRNHVLLHLEANHRVQIERRETCPELHPEDKIPRIAHFS    947

PTEN        ---------------      403
DAF-18      ENSFSDSNFDQAIYL      962
```

D

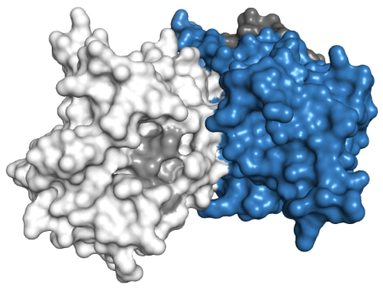
PTEN 1D5R



DAF-18 53-506



PTEN



DAF-18

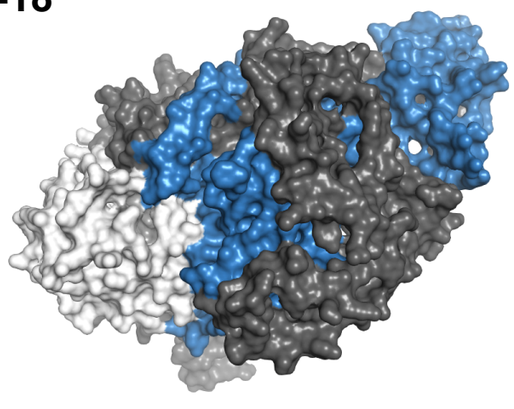


**Figure 2 | Functional and structural similarity of *C. elegans* DAF-18 and human PTEN. A)** Protein domain annotations for DAF-18 and PTEN. The canonical DAF-18 amino acid (AA) sequence is more than twice as long as the canonical PTEN AA sequence primarily due to elongation of the C2 membrane targeting and C-tail domains. **B)** Both DAF-18 and PTEN function as lipid and protein phosphatases to antagonize the highly conserved canonical insulin signaling pathway. **C)** Clustal alignment of DAF-18 and PTEN. DAF-18 and PTEN share a highly conserved phosphatase domain (46% identity) and fully conserved catalytic site (residues highlighted in grey). DAF-18 has a markedly longer and less conserved C-terminal region than human PTEN resulting in low overall amino acid similarity (20%) and identity (13.6%). Note that although the C-terminal region is much longer there are small conserved motifs spread throughout that are not illustrated by this alignment (see Liu et al., 2014 oncogene). D) Comparison of DAF-18 and PTEN structural models. **D)** (top left) Solved crystal structure of human PTEN (1D5R reference structure; Lee et al. 1999). (top right) Predicted structure of DAF-18 AA 53-506 indicating similar 3D structure to human PTEN. (bottom left) Predicted structure of full-length PTEN and (bottom right) DAF-18 illustrating the increased size of DAF-18. Note the full length DAF-18 model is likely to be inaccurate due to poor homology-based modeling of the non-conserved C-terminal region. Domain colour mapping matches panel 2A except that the catalytic site within the phosphatase domain is shaded light gray.

**Figure 3 | A conserved neuronal role for *PTEN* in NaCl preference revealed by a novel automated chemotaxis paradigm. A)** (top) All phenotypic analysis is conducted using our machine vision system the Multi-Worm Tracker. The Multi-Worm Tracker delivers stimuli and preforms image acquisition, object selection, and parameter selection in real time while choreography software extracts detailed phenotypic information offline (A bottom panels) I) petri plate of *C. elegans* II) A petri plate of *C. elegans* selected for analysis by the Multi-Worm Tracker III) A Multi-worm tracker digital representation showing the degree of phenotypic detail. An example behaviour scored by the Multi-Worm Tracker: the *C. elegans* response to a mechanosensory tap to the side of the Petri plate is brief backwards locomotion (from III to IV). B) Behavioural track tracing of a plate of worms from a novel Automated Multi-Worm Tracker NaCl chemotaxis paradigm illustrating attractive navigation behaviour of wild-type animals toward a point source of NaCl. **B)** (bottom left to right) circles and arrows and **C)** (left to right) worm tracks represent navigation trajectories of wild-type attraction to a point source of NaCl, a *daf-18(e1375)* reduction-of-function decrease in NaCl chemotaxis, and a transgenic rescue of NaCl preference via pan neuronal overexpression of WT human *PTEN* in *daf-18(e1375)* reduction of function mutants. **D)** Quantitative chemotaxis index scores across genotypes. Pan neuronal expression of human *PTEN* rescues the reversed NaCl preference of *daf-18(e1375)* mutants to wild-type levels. Circles represent plate replicates run on the same day and inset number represent the number of individual animals registered by the tracker and located outside the center starting region (i.e. included in preference score) across the three plate replicates for each genotype. Error bars represent standard deviation using the number of plates as n (n = 3). (****) P < 0.0001, n.s. not significant, one-way ANOVA and Tukey's post-hoc test.

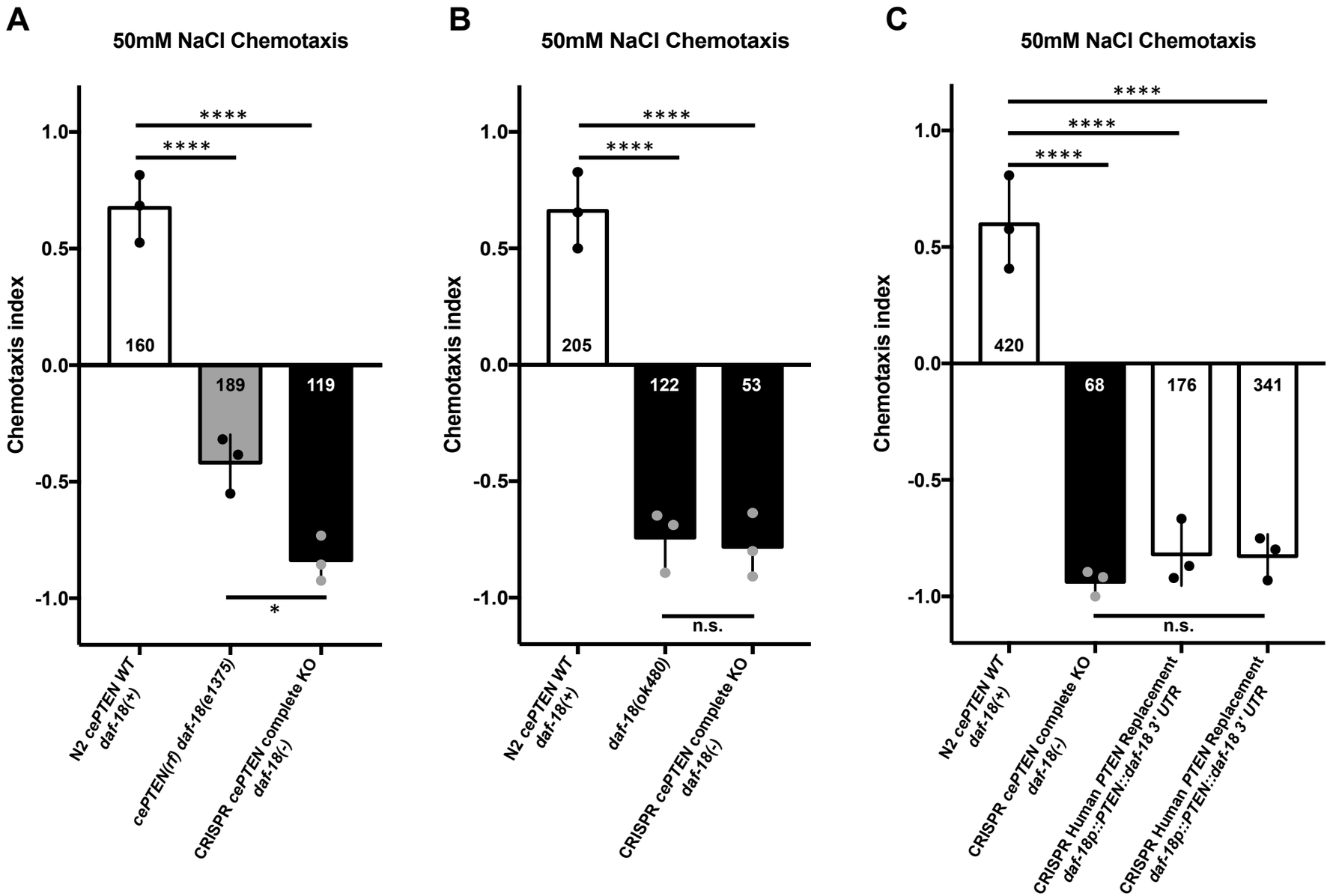
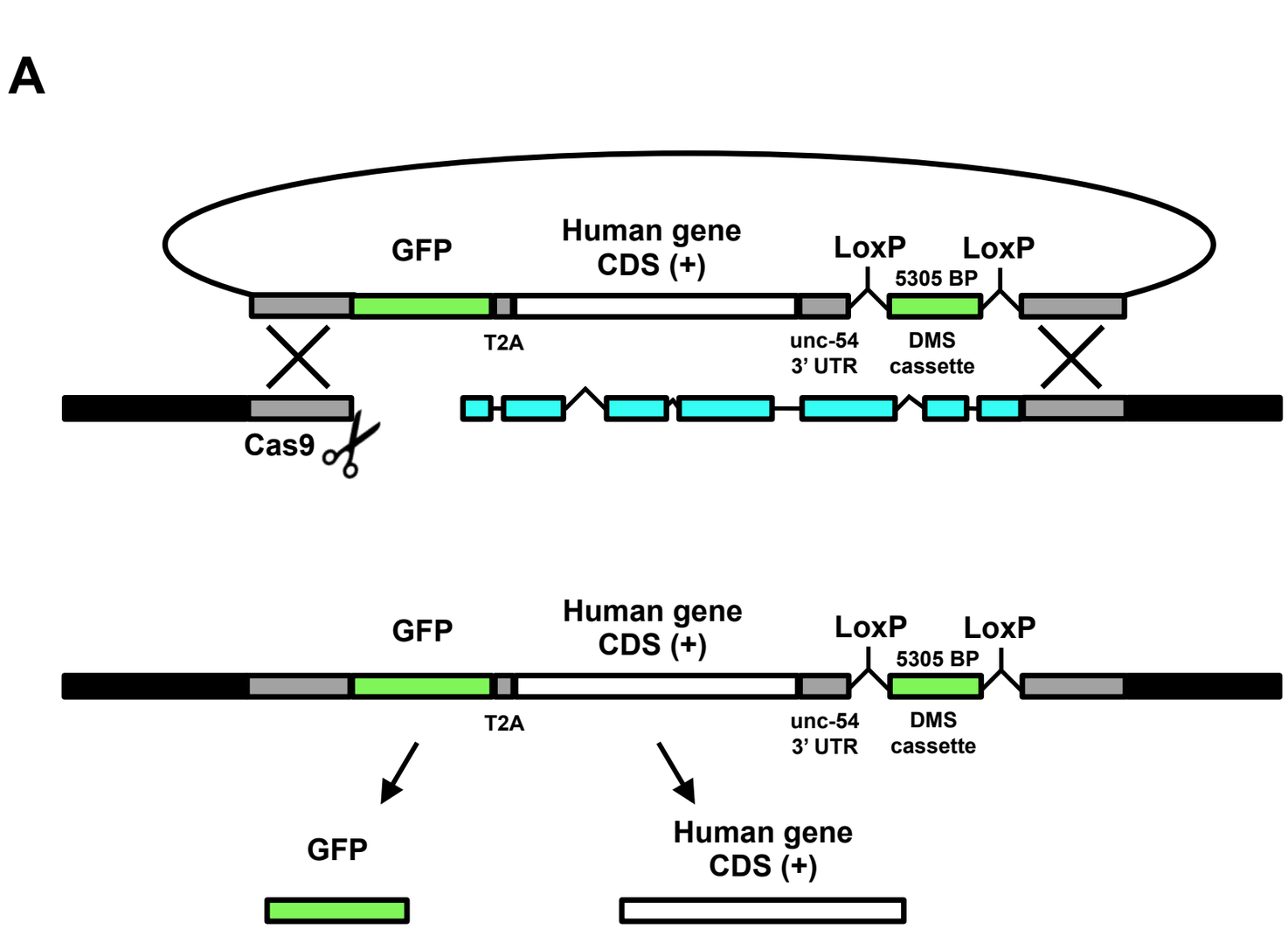



**Figure 4 | Complete *daf-18* ORF deletion causes strong NaCl avoidance that is not rescued by direct replacement with human *PTEN*. A)** NaCl chemotaxis preference scores of wild-type, *daf-18(e1375)* reduction of function, and CRISPR *daf-18* complete deletion mutants. *daf-18* complete deletion mutants show significantly stronger NaCl avoidance than *daf-18(e1375)* reduction of function mutants. **B)** CRSPR *daf-18* complete deletion mutants are not significantly different from animals harboring the putative null *daf-18(ok480)* deletion allele. **C)** Expression of a single copy *PTEN* transgene from the native *daf-18* locus is insufficient to significantly rescue NaCl chemotaxis towards wild-type levels. **D)** RT-PCR confirming expression of full length PTEN mRNA in the two independent knock-in lines used for behavioural analysis. Previously validated primers that target *cmk-1* intronic regions of genomic DNA do not produce products following DNase treatment, confirming purity of the cDNA. Circles represent plate replicates run on the same day and inset number represent the number of individual animals registered by the tracker and located outside the center starting region (i.e. included in preference score) across the three plate replicates for each genotype. Error bars represent standard deviation using the number of plates as n (n = 3). (****) $P < 0.0001$, (*) $P < 0.05$ n.s. not significant, one-way ANOVA and Tukey's post-hoc test.



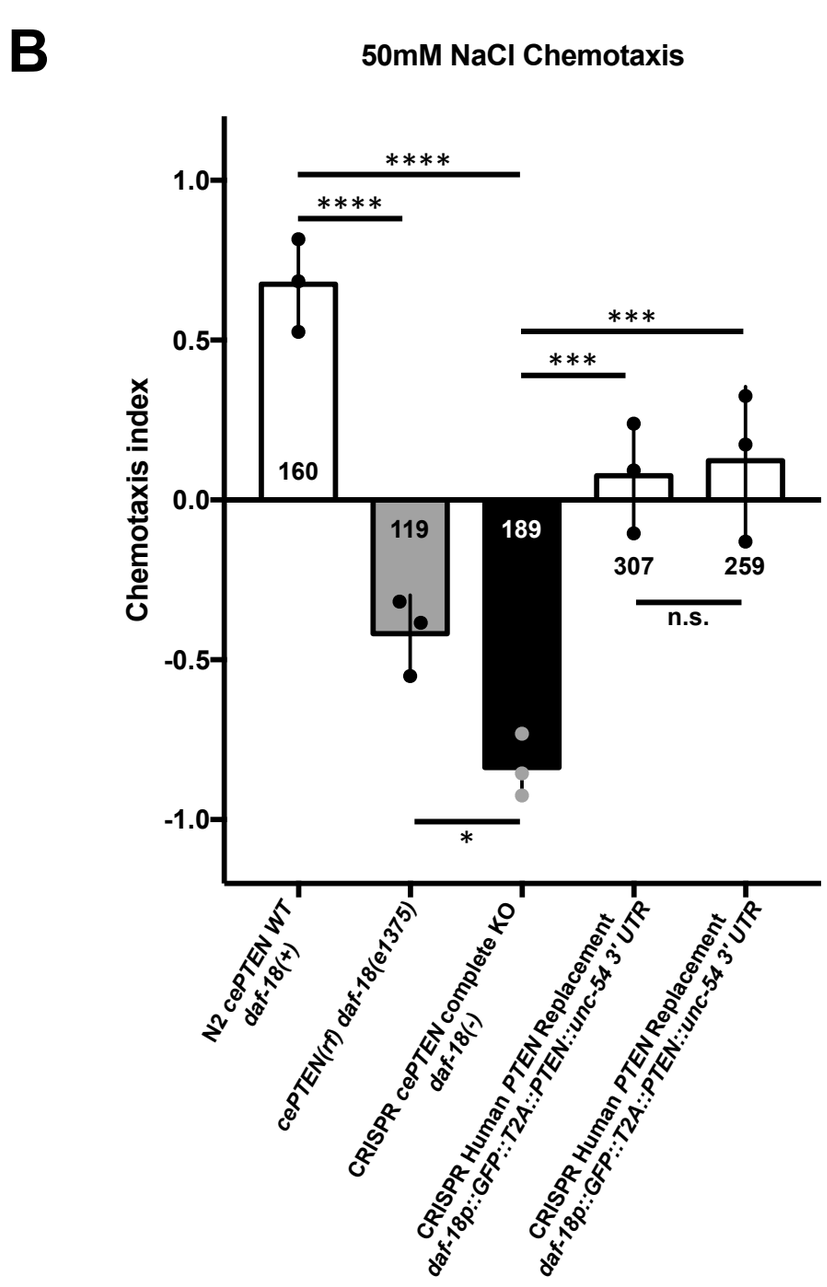




**Figure 5 | A streamlined human gene replacement strategy functionally replaces *daf-18* with human *PTEN*. A)** Streamlined CRISPR gene replacement strategy. Inclusion of the validated *unc-54* 3' UTR in the upstream homology arm increases the speed of transgenesis by removing the need for cassette excision to induce transgene expression. This alternate approach also offers the option for retained visual transgenic markers, which simplifies the generation and phenotypic analysis of heterozygotes and double mutants. The inclusion of a GFP::T2A cassette is an optional addition to allow for confirmation of transgene expression without altering human gene function (Ahier and Jarriault, 2014) . **B)** Expression of wild-type human *PTEN* using this genome editing strategy significantly rescued NaCl chemotaxis toward wild-type levels. Circles represent plate replicates run on the same day and inset number represent the number of individual animals registered by the tracker and located outside the center starting region (i.e. included in preference score) across the three plate replicates for each genotype. Error bars represent standard deviation using the number of plates as n (n = 3). (****) $P < 0.0001$, (***) $P < 0.001$, (*) $P < 0.05$, n.s. not significant, one-way ANOVA and Tukey's post-hoc test.

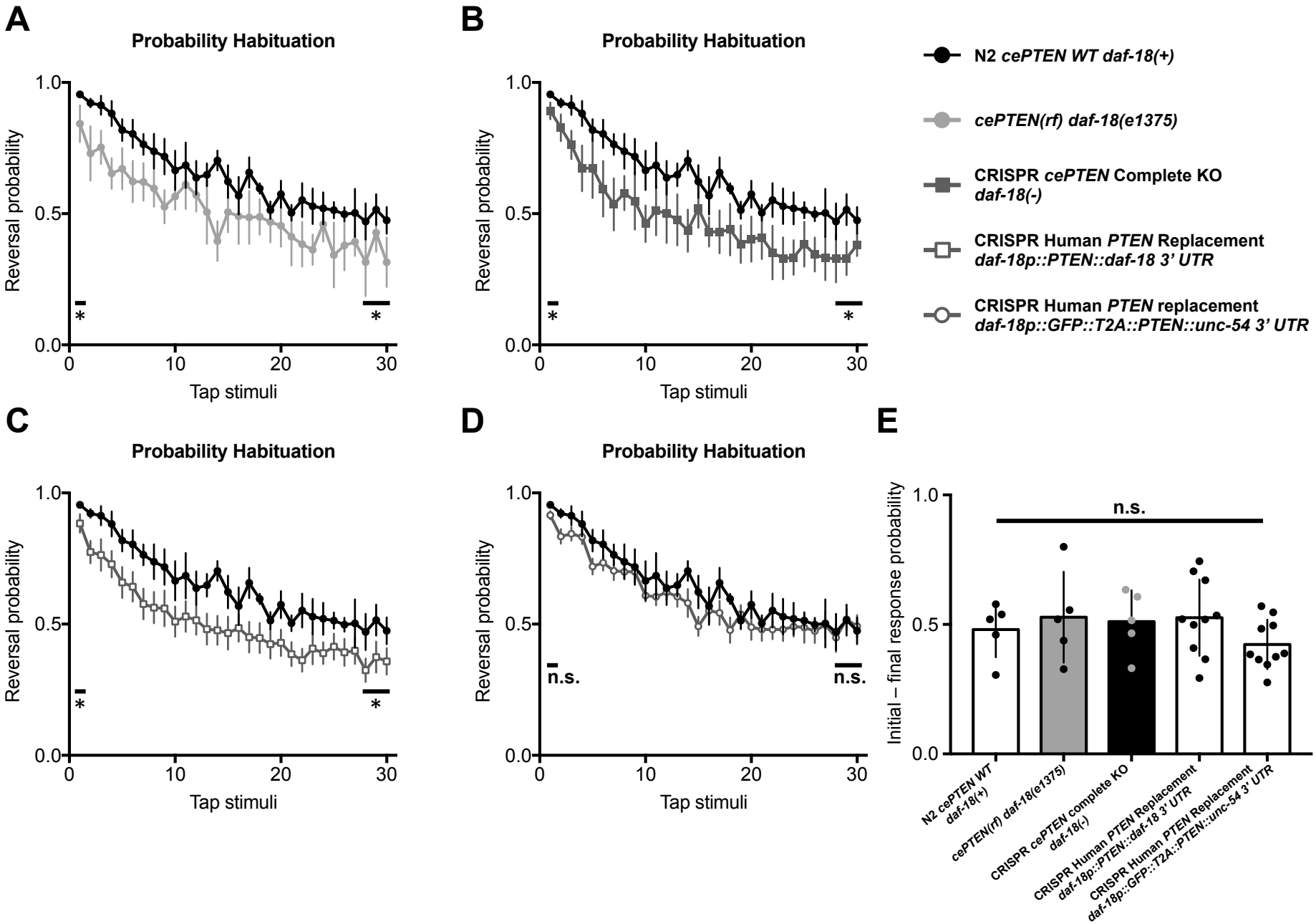



**Figure 6 | *daf-18* deletion causes mechanosensory hyporesponsivity that is rescued by targeted replacement with human *PTEN*. A-D)** Average probability of eliciting a reversal response to 30 consecutive tap stimuli delivered at a 10s ISI. A) *daf-18(e1375)* reduction of function and **B)** *daf-18* complete deletion mutants exhibit significantly reduced probability of eliciting a reversal response throughout the habituation training session, indicating mechanosensory hyporesponsivity. **C)** Replacement of *daf-18* with human *PTEN* using the original strategy does not rescue mechanosensory hyporesponsivity. **D)** Expression of human *PTEN* using the streamlined replacement strategy rescues mechanosensory responding to wild-type levels. Error bars represent standard error of the mean. **E)** Habituation, or the ability to learn to decrease the probability of eliciting of a reversal response throughout the training session was not significantly altered in *daf-18* mutants. Circles represent plate replicates run on the same day. Error bars represent standard deviation of the mean using the number of plates as n (n = 5 or 10). (*) $P < 0.05$ n.s. not significant, one-way ANOVA and Tukey's post-hoc test.

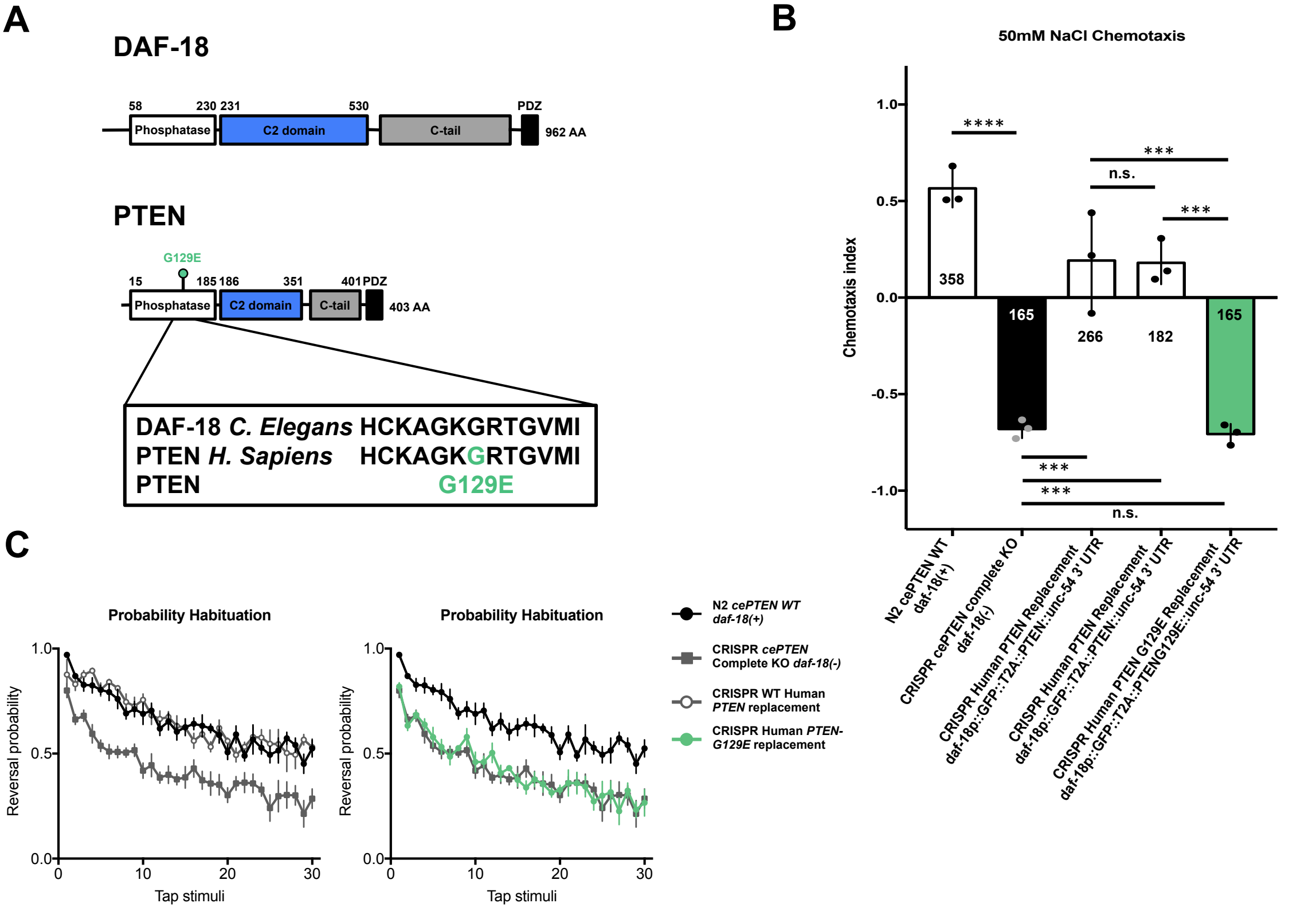
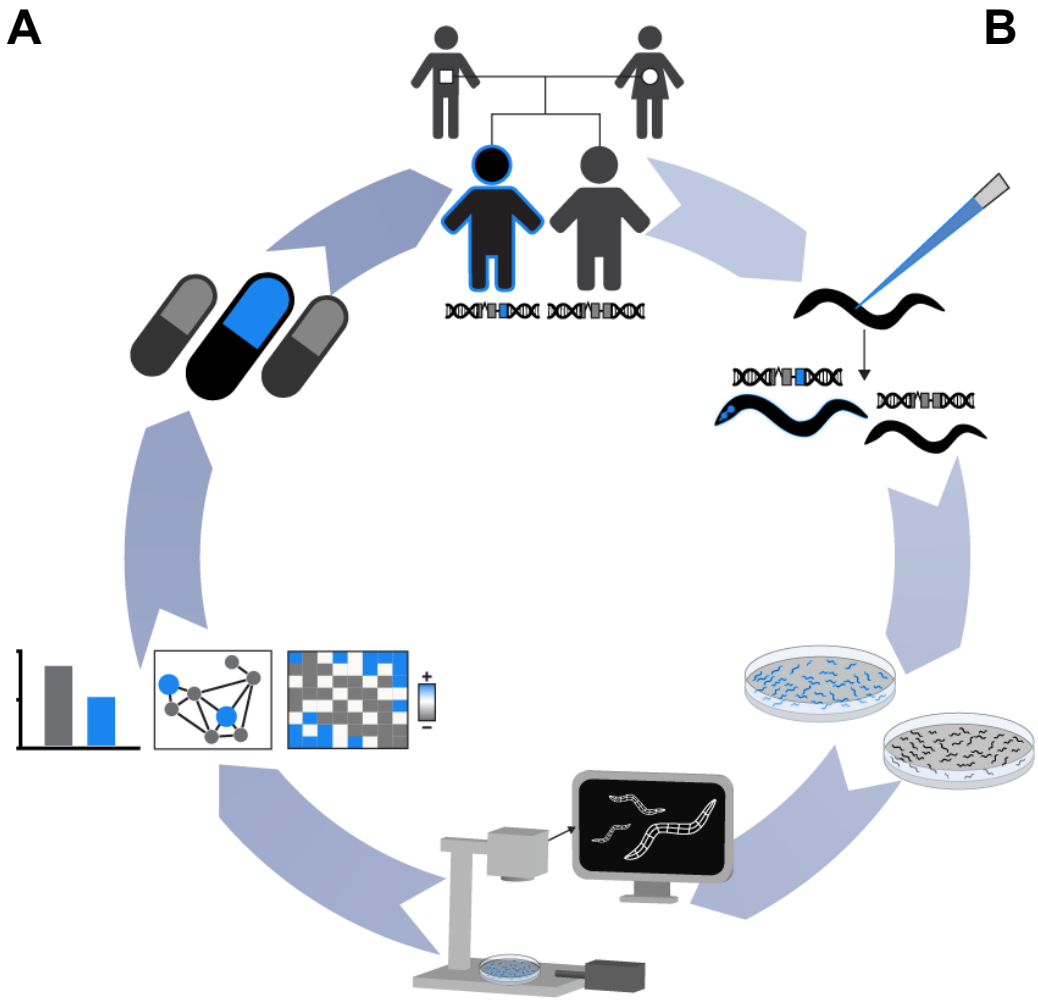



**Figure 7 | Human gene replacement and *in vivo* phenotypic assessment accurately identifies functional consequences of the pathogenic PTEN-G129E variant. A)** Location (top) and conserved amino acid sequence change (bottom) of the pathogenic lipid phosphatase inactive PTEN-G129E variant within the PTEN phosphatase domain. **B)** Animals harboring the PTEN G129E variant displayed strong NaCl avoidance equivalent to animals carrying the complete *daf-18* deletion allele, indicating loss-of-function (Fig. 7B). Circles represent plate replicates run on the same day and inset number represent the number of individual animals registered by the tracker and located outside the center starting region (i.e. included in preference score) across the three plate replicates for each genotype. Error bars represent standard deviation using the number of plates as n (n = 3). **C)** Similarly, PTEN-G129E mutants also displayed mechanosensory hyporesponsivity that was not significantly different from *daf-18* deletion mutants. Error bars represent standard error of the mean. (****) $P < 0.0001$, (***) $P < 0.001$, (*) $P < 0.05$, n.s. not significant, one-way ANOVA and Tukey's post-hoc test.




**Advantages**

- Rapid and precise gene manipulation and phenotypic assessment
- Cell specific manipulation to determine origin of pathology
- Inducible manipulation to determine critical time
- Dissection of molecular pathways

**Fundamental Knowledge**

- Gene function
- Molecular networks
- Mechanisms of neural and behavioural plasticity
- Genetic networks

**Translational Benefits**

- Insights into disease pathology
- Understanding clinical variability
- Targets for therapeutics

**Figure 8 | A conceptual framework for *in vivo* functional analysis of human genetic variation using *C. elegans*. A)** (top working clockwise) A human gene and/or variant of uncertain significance is implicated in disease etiology through clinical sequencing. Targeted CRISPR human gene replacement or analogous methods are used to generate a library of knock-out, human wild-type and variant transgenic strains. Large isogenic synchronous colonies of these transgenic worms are grown and their morphology, baseline locomotion, and sensory phenotypes are rapidly characterized using machine vision to establish novel functional assays and interpret variant effects. In vivo functional data can be used to probe epistatic network disruptions and cluster variants based on multi-parametric phenotypic profiles. The integrated humanized transgenic lines and functional assays greatly facilitate downstream applications including precision medicine drug screens designed to identify compounds that reverse the effects of a particular patients missense variant. **B)** Advantages of targeted human gene replacement using *C. elegans.*

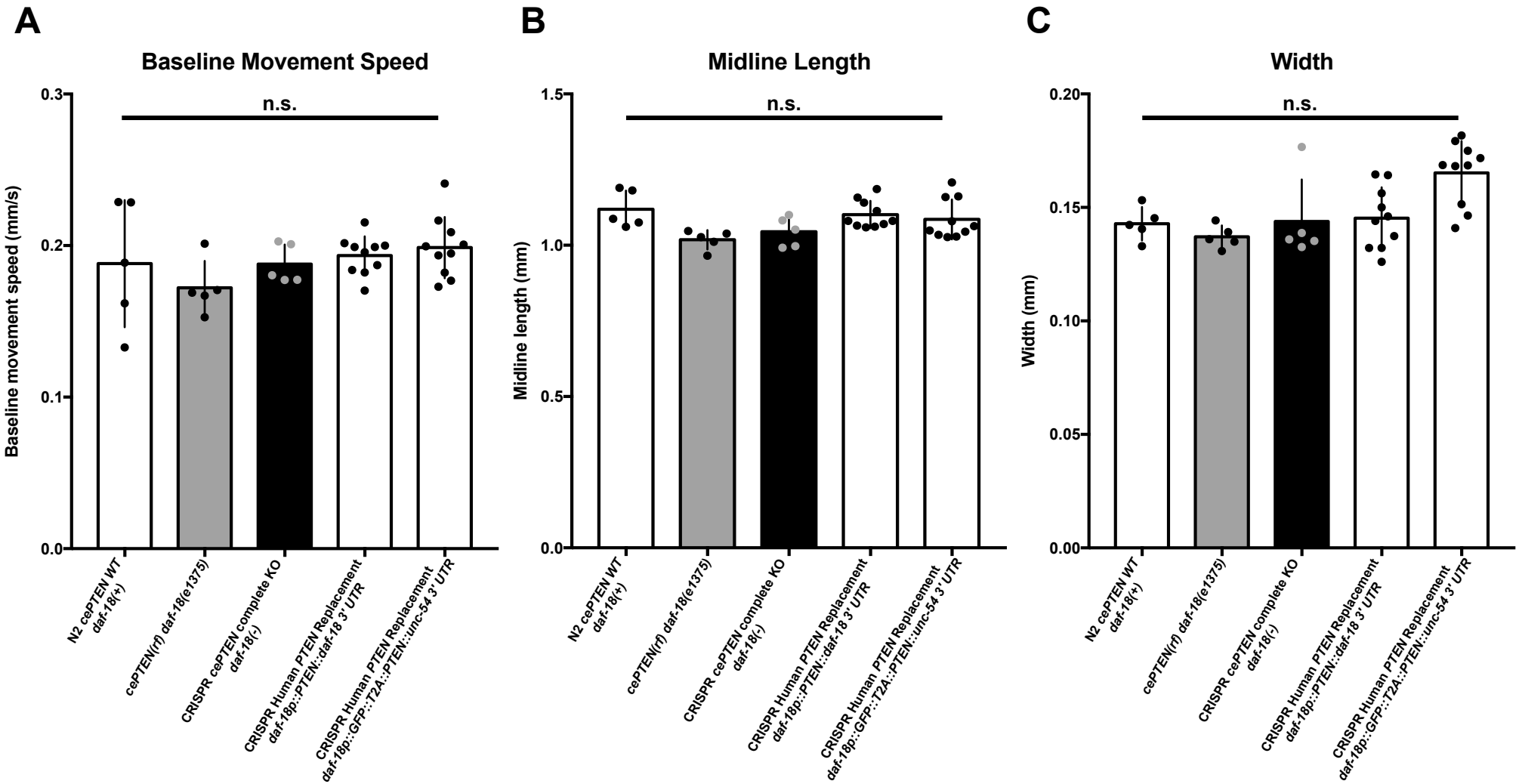



**Figure S1 | Morphology and baseline locomotion are superficially normal in *daf-18* mutants and *PTEN* transgenic animals. A)** Baseline movement speed, **B)** midline length, and **C)** width are not significantly different across genotypes. Circles represent plate replicates run on the same day. Error bars represent standard deviation using the number of plates as n (n = 5 or 10). n.s. not significant, one-way ANOVA and Tukey's post-hoc test.