

Modeling human monogenic diseases using the tunicate *Ciona intestinalis*

We applied a decision-making framework for identifying tractable genes from our organism selection dataset for pilot experiments in the tunicate *Ciona intestinalis*.

Contributors (A-Z)

Prachee Avasthi, Audrey Bell, Keith Cheveralls, Seemay Chou, Megan L. Hochstrasser, Alberto Stolfi, Dennis A. Sun, Ryan York

Version 2 · Jul 29, 2025

We developed a framework for selecting disease–gene pairs from our organism selection dataset for experimental testing [1]. Here, we apply that framework to genes in the tunicate *Ciona intestinalis* (recently disambiguated into two species, *Ciona intestinalis* and *Ciona robusta* [2]) to identify three genes of interest. We briefly introduce the life history and available genetic tools in *Ciona*. We also provide examples that explain our rationale for rejecting or pursuing individual genes.

Researchers developing genetic models of human disease using animals, particularly tunicates, may find it informative to read about how we considered experimental design trade-offs. In addition, we develop visualization tools for single-cell and bulk RNA-seq datasets for *Ciona* that others may find useful when designing experiments.

- All code related to this pub is available in this [GitHub repo](#). Code specific to *C. intestinalis* is in [these analysis notebooks](#) and [this set of visualization scripts](#).

Ciona intestinalis is a tunicate (sea squirt) and a well-established model organism in developmental biology [3]. Tunicates are marine invertebrate chordates that represent the closest living relatives to vertebrates [4], making them invaluable for understanding chordate evolution and development [5]. This solitary ascidian is studied extensively to understand fundamental mechanisms of embryonic development, cell fate specification, and the evolutionary origins of vertebrate characteristics.

Adult *Ciona* are sessile filter-feeders with a distinctive sac-like body plan, but their free-swimming tadpole larvae possess classic chordate features, including a notochord, dorsal neural tube, and muscular tail [6] (Figure 1, A). The larvae undergo rapid development, completing embryogenesis in just 18 hours at 18 °C [7], making them ideal for developmental studies. Following a brief planktonic phase, tadpole larvae settle and undergo dramatic metamorphosis, resorbing their tail and notochord while developing into the sessile adult form. This life cycle transformation provides unique insights into chordate body plan evolution and the developmental programs underlying major morphological transitions.

The *Ciona* embryo is particularly tractable for cell lineage studies due to its invariant cleavage pattern and relatively small cell number [8][9]. Each blastomere's fate can be precisely traced, and the embryo's optical transparency allows for detailed observation of cellular behaviors during development.

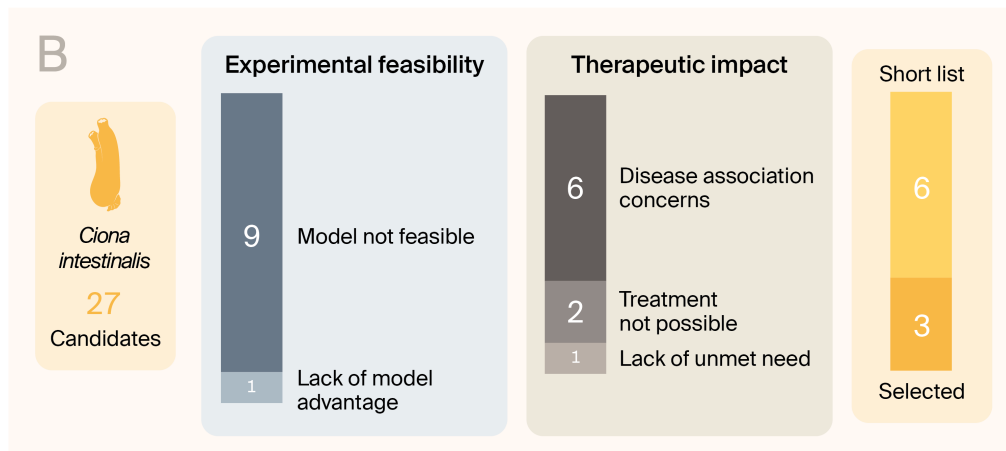
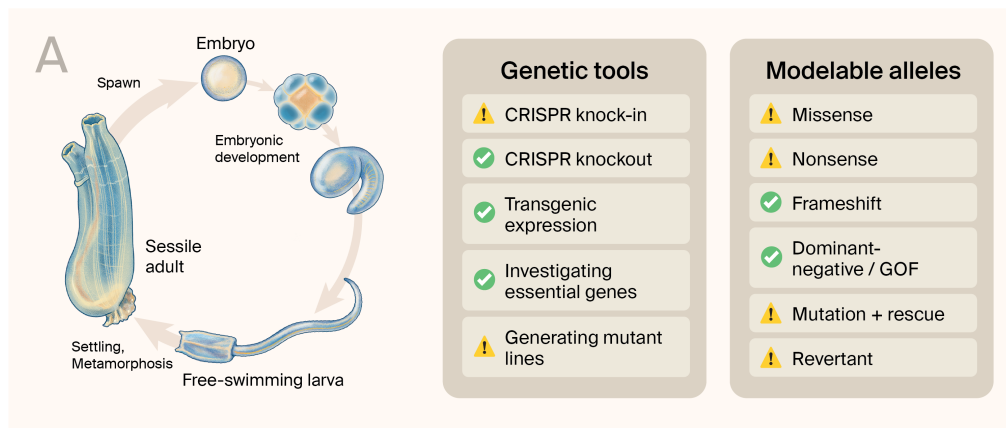


Figure 1

Overview of diligence on *Ciona intestinalis*.

(A) Illustration of life cycle of *Ciona* and schematic of available genetic tools and types of human alleles that we expect to be able to model, given available tools. Sessile adults spawn embryos, which develop into free-swimming tadpole larvae with chordate-like features. These tadpoles settle and metamorphose into the adult form.

(B) Illustration of manual diligence outcomes listing the number of genes eliminated by each failure mode. The initial shortlist was nine genes, three of which were selected for experimental testing.

State of technology

Ciona has robust genetic tools and genomic resources.

- **Genetic tools.** *Ciona* has well-developed genetic manipulation capabilities. Electroporation-mediated transgenesis is routine, allowing for tissue-specific gene expression studies using established promoter elements [10][11]. CRISPR-Cas9 genome editing is highly efficient in this system, enabling targeted mutagenesis [12] [13]. However, some barriers to reproducible CRISPR knockouts in this organism include the high degree of genome heterozygosity (1.1–1.2%) [14] and the fact that many laboratories obtain *Ciona* embryos through environmentally collected adult samples, rather than working on stable lines with a closed life cycle.

Previous reports have indicated that it's possible to generate CRISPR knock-ins within *Ciona* embryos [15], but the experts we consulted with haven't been able to reproduce this work in their lab. As such, we focused on CRISPR knockouts as our main tool for generating genetic models.

Generating mutant lines in *Ciona* is possible with sufficient time and resources. However, the experts we consulted with aren't well equipped to close the life cycle on knockouts, which requires metamorphosis, rearing, and maintenance of adult animals, so we've decided to focus on phenotypes in G0 knockout embryos.

- **Phenotypic assays.** *Ciona* embryos and larvae are highly amenable to experimental manipulation and observation (for details about genetic tools and the size of the organism, see [16]). Thousands of ~0.1mm embryos can be collected and electroporated at once, allowing for high-throughput experiments [11]. Standard protocols exist for *in situ* hybridization [17], immunohistochemistry [18], and generation of reporters using transposon-mediated transgenesis [19]. The transparent embryos are particularly well-suited for live imaging approaches.

A unique feature of *Ciona* development is that knockouts of normally essential genes don't usually lead to embryonic death, due to a lack of apoptotic processes in the embryo. This means that some genes that would be lethal in human cells might have a recoverable phenotype in *Ciona*.

- **Gene expression.** Comprehensive transcriptomic resources are available for *C. intestinalis*, including developmental time-course RNA-seq data covering embryonic and larval stages [20]. Single-cell RNA sequencing has provided detailed cell-type-specific expression profiles throughout development, revealing the molecular basis of cell fate specification and differentiation programs in this chordate model system.

We manually diligenced 27 different genes in *Ciona intestinalis* to arrive at nine potentially actionable genes for pilot experiments. A summary of the diligence results can be seen in [Figure 1](#), B.

You can see a copy of the **working document** we used to catalog our thoughts for each *Ciona* gene we diligenced [here](#).

Example genes we rejected

In this section, we explain why we rejected a few genes from *C. intestinalis* to show how we applied our selection criteria.

CACNB2 (calcium voltage-gated channel auxiliary subunit beta 2)

Model not feasible

This subunit of a voltage-gated calcium channel is implicated in a subtype of Brugada syndrome, a disease that causes death by sudden heart failure and is detectable by abnormalities in electrocardiogram signals [21]. All cases of this disease recorded in studies in OMIM were heterozygous missense mutations. Because we can't easily generate targeted mutations in *Ciona*, we decided not to pursue this gene.

RP9 (RP9 pre-mRNA splicing factor)

Disease association concerns

This protein is implicated in a subtype of retinitis pigmentosa (RP), a disease that causes gradual loss of eyesight due to the death of retinal cells. While a particular variant (His137Leu) has been reported in the literature as causing RP [22], follow-up work seems to indicate that sequencing results identifying this variant may actually come from a different pseudogene locus within the genome, which could have been

amplified in previous assays [23][24]. Because there's a lack of confidence in identifying variations in this gene as causative of RP, we decided not to pursue this gene.

***MLYCD* (Malonyl-CoA decarboxylase)**

Lack of model advantage

This protein is involved in fatty acid metabolism, where it breaks down malonyl-CoA into acetyl-CoA and carbon dioxide [25]. Loss of this protein leads to metabolic defects across organ systems. When considering how *Ciona* might model this disease, we also considered whether identifying molecules that help clear malonyl-CoA in human *in vitro* cell culture systems might be more efficient. It's possible that modeling in *Ciona* might uncover phenotypes that wouldn't be obvious in cell lines or less appropriate models. However, the experts we consulted also previously indicated that metabolic phenotypes or others involved in growth may not be well-modeled by the fast-developing embryos of *Ciona*. For these reasons, we decided not to pursue this gene.

Our initial shortlist

Here, we describe a few of the best-supported genes to model in *Ciona*.

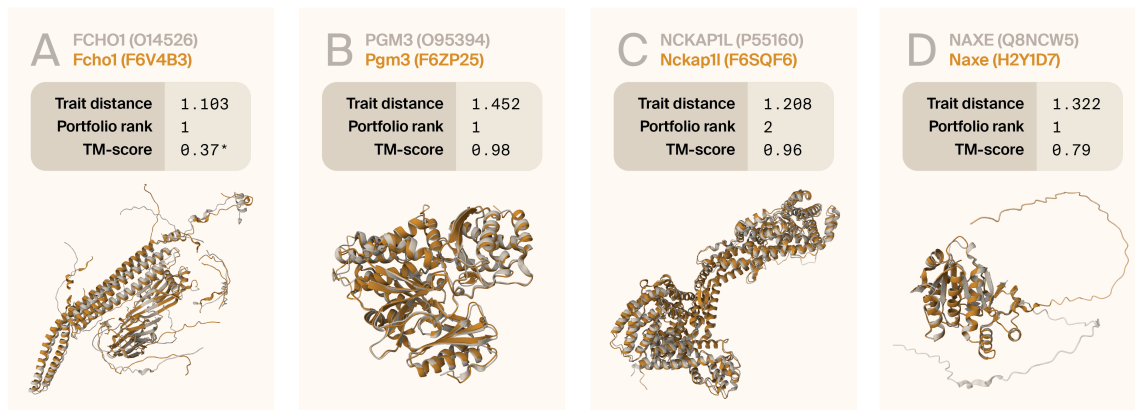


Figure 2

Structural comparisons of top-selected shortlist proteins to human proteins.

(A) FCH and mu domain-containing endocytic adaptor 1 [flexible alignment generated using jFATCAT for visualization purposes (see the “Methods” section of our previous pub for more details [26])]

(B) Phosphoglucomutase 3

(C) NCK-associated protein 1-like protein

(D) NAD(P)HX epimerase

Tan proteins are predicted human structures; gold proteins are predicted *C. intestinalis* homolog structures. “Trait distance” is a multivariate Mahalanobis distance between pairs of proteins calculated based on 10 physicochemical properties (see the “Methods” section of our previous pub for more details [26]). “Portfolio rank” is the relative rank of this protein compared to all other proteins found in the organism selection/the Zoogle portal. TM-score is a standard measure of solid-body protein structural similarity [27]. In general, TM-scores of 0.5 or above indicate the same fold.

***FCHO1* (FCH and mu domain-containing endocytic adaptor 1)**

This protein is a nucleator of clathrin-coated pits [28]. Loss of function through missense, nonsense, or splice mutations leads to an immunodeficiency characterized by recurrent infections due to T- and B-cell dysfunction and usually death in childhood [29][30].

The *Ciona* match from our organism selection pipeline shows both FCHO1 and FCHO2 as strong hits from Foldseek, but our interpretation of this similarity is confounded because this family of proteins is large and appears to contain many disordered regions, which makes it unsuitable for comparison using rigid-body structural comparison approaches such as the TM-score (Figure 2). The *Ciona* homolog of *FCHO1* appears to be expressed throughout the embryo during development (Figure 3).

Previous work in *Ciona* has investigated the function of clathrin endocytosis in the morphogenesis of the notochord [31]. During this process, a lumen is formed between adjacent cells in the notochord. Drugs that target clathrin-mediated endocytosis, including chlorpromazine, pitstop 2, and Dynasore, disrupt the formation of this lumen [31]. Given the strong possibility of a similar phenotype when mutating *Fcho1* in *Ciona*, we decided that this gene would be tractable for downstream analyses.

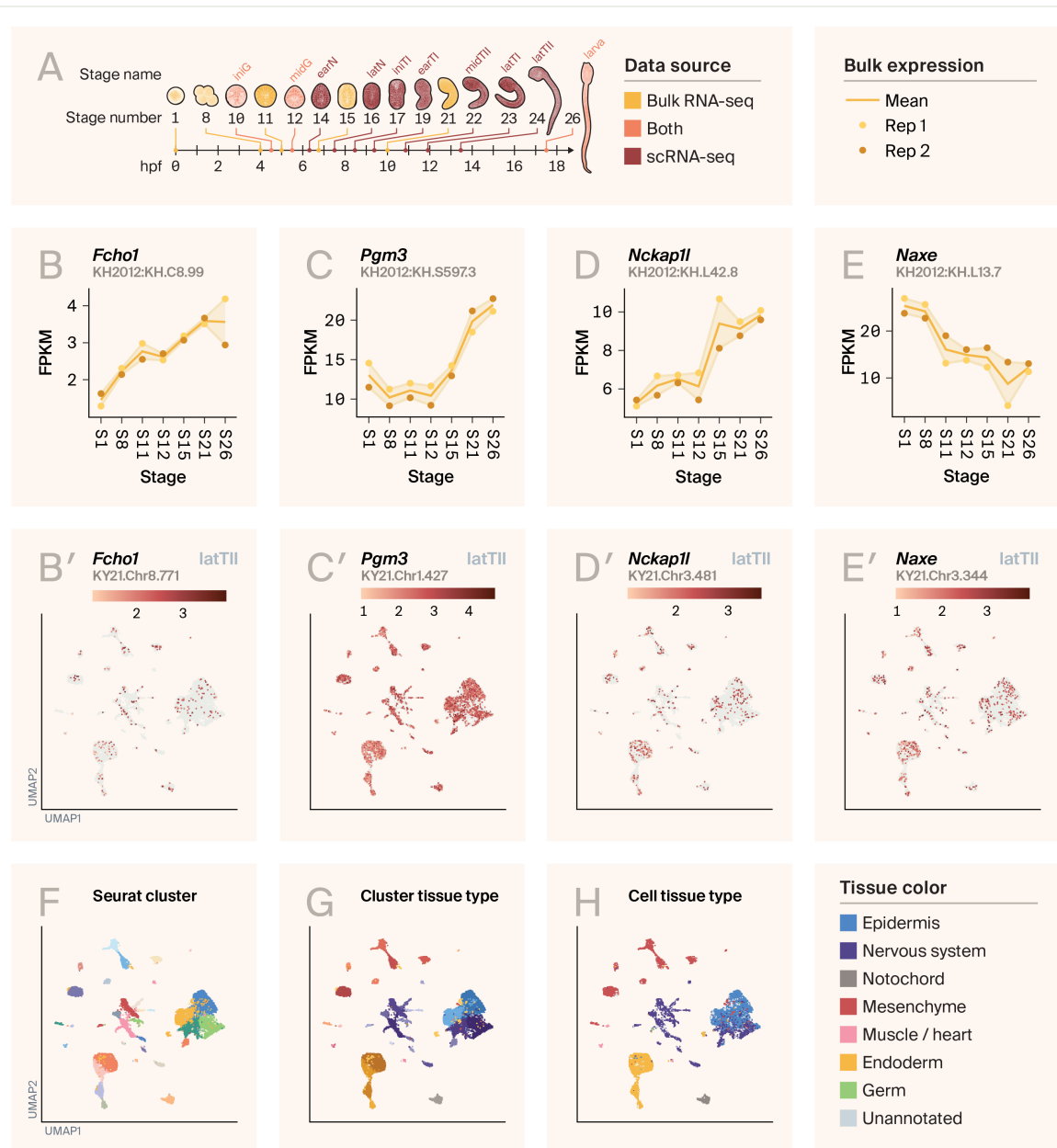


Figure 3

Bulk and single-cell RNA-seq expression in *Ciona intestinalis* embryogenesis.

(A) Illustration of *Ciona*'s developmental stages. Yellow stages are found in the bulk RNA-seq dataset from ANISEED. Red stages are found in the single-cell RNA-seq data from Cao et al. [20] and the Piekarz reanalysis. Stages colored in orange are shared between the datasets.

hpf: hours post-fertilization, FPKM: fragments per kilobase-million, latTII: late tailbud II, UMAP: uniform manifold approximation and projection

(B–E) Expression of *Fcho1* (B), *Pgm3* (C), *Nckap1l* (D), and *Naxe* (E) in the ANISEED bulk RNA-seq dataset. ANISEED data uses the KH2012 identifier, which is displayed in each panel.

(B'–E') Expression of *Fcho1* (B'), *Pgm3* (C'), *Nckap1l* (D'), and *Naxe* (E') in the Cao / Piekarz datasets, for all cells in stage LatTII. The Piekarz dataset uses the KY2021 identifier, which is displayed in each panel.

(F) Cells from panels B'–E', colored by Seurat cluster identified in the Piekarz reanalysis.

(G) Cells from panels B'–E', colored by Seurat cluster identified in the Piekarz reanalysis based on hues that represent each primary tissue type found in the dataset.

(H) Cells from panels B'–E', colored by the individual tissue type label for each individual cell, from the Cao et al dataset.

For bulk expression of additional genes, see this [supplementary figure](#).

PGM3 (phosphoglycomutase 3)

Phosphoglycomutase is a crucial enzyme involved in glycosylation processes that converts GlcNAc-6-phosphate to GlcNAc-1-phosphate [32]. Heterozygous missense, frameshift, and nonsense mutations result in immunodeficiency characterized by recurrent infections and morphological defects across organ systems [33][34][35]. Mice with partial loss of function show defects in hematopoiesis, while homozygous loss of function is embryonic-lethal [32].

PGM3 is the top hit to the *Ciona* protein in FoldSeek searches and has a very high TM-score of 0.98 (Figure 2). The *Ciona* homolog is also expressed in a diversity of cell types across development (Figure 2).

Previous work has characterized the degree of glycosylation of proteins in *Ciona* embryos across developmental stages using mass spectrometry [36]. This work observed many glycosylation patterns that contain the substrates of PGM3, suggesting that its activity would be necessary for the production of these moieties. Disruption of glycosylation through drugs like NGI-1 and tunicamycin also appears to

result in defects in notochord lumen formation [37]. Due to the strong evidence for the importance of glycosylation in *Ciona* development, particularly of the notochord, we were interested in pursuing this gene for experimental testing.

PGM3 is also being investigated as a target for treating breast cancer for its role in the hexosamine biosynthetic pathway (HBP) [38]. A commercially available PGM3 inhibitor, FR054, is available, potentially letting us cross-validate our knockouts with chemical treatments and see how similarly the *Ciona* homolog responds to this drug.

***NCKAP1L* (NCK-associated protein 1-like protein)**

This protein, also known as HEM1, is a subunit of a larger scaffold that supports the wave regulatory complex (WRC), which itself regulates the actin cytoskeleton [39]. Missense and splice abnormalities in this protein lead to immunodeficiency characterized by recurrent viral and bacterial infections, combined with autoimmunity characterized by excessive inflammation [40][41]. Patient cells appear to have defects in lamellipodia, which are hypothesized to impede proper T-cell development.

The *Ciona* homolog appears to be a strong structural match to both NCKAP1 and NCKAP1L (TM-score = 0.96) (Figure 2). *NCKAP1* is a homolog of *NCKAP1L*; *NCKAP1L* is distinct in its high expression in the human immune system. The *Ciona* homolog appears to be expressed across developmental stages and tissue types (Figure 2).

Lamellipodia are hypothesized to play a role in *Ciona* notochord lumen development and appear to form at the leading edge of notochord cells during lumen formation [42]. Disruption of the actin cytoskeleton through drugs such as latrunculin B and blebbistatin disrupts notochord lumen morphology, suggesting that proper actin dynamics are required for this process. For these reasons, we were interested in testing whether *Nckap1l* knockout may have an assayable phenotype in *Ciona*.

***NAXE* [NAD(P)HX epimerase]**

This metabolic enzyme is necessary for detoxifying metabolic byproducts (S-NADHX, R-NADHX, and cyclic NADHX) of the essential coenzymes NADH and NADPH back into their original, useful states [43]. Missense, nonsense, frameshift, and splice mutations in *NAXE* lead to a syndrome in which patients, usually infants, experience metabolic crisis after developing a fever due to depleted NADH/NADPH and

accumulation of NADHX [44][45][46]. Most otherwise phenotypically normal patients die as a result of this syndrome due to damage to the brain from encephalopathy or leukoencephalopathy.

Curiously, work on mice knockouts of *Naxe* (also known as *Apoa1bp* or *A1bp* due to its additional function in binding apolipoprotein A1) doesn't appear to show strong effects on lifespan or NADH levels [47]. Mutations of this gene in baker's yeast (*Saccharomyces cerevisiae*) [48] and in thale cress (*Arabidopsis thaliana*) [49] indicate that *Naxe* isn't essential for these organisms. Work on this disease could benefit from additional models.

The *Ciona* homolog of *NAXE* is a strong structural match to the human homolog (TM-score = 0.79) and is expressed broadly across cell types throughout development. Given its core metabolic function and lack of effective existing models, we were interested in seeing whether we could model *NAXE* function in *Ciona*.

Recent work has suggested that treating patients with precursors of NADH, such as nicotinamide and niacin, could help relieve metabolic crisis [45]. Given these data, we were also interested in testing whether these molecules could rescue any potential phenotypic consequences of *NAXE* knockout in *Ciona*.

Other hits

Several other genes – *RBBP7*, *GAS2L2*, *DHODH*, *CWF19L1*, and *VWA8* – also made our short list ([Table 1](#)).

HGNC gene symbol (UniProt ID)	Human protein name	Associated human disease (OMIM)	Ciona protein UniProt ID	Cic ide
<i>FCHO1</i> (O14526)	FCH and mu domain containing endocytic adaptor 1	<u>Immunodeficiency 76</u>	<u>F6V4B3</u>	KH
<i>PGM3</i> (O95394)	Phosphoglucomutase 3	<u>Immunodeficiency 23</u>	<u>F6ZP25</u>	KH
<i>NCKAP1L</i> (P55160)	NCK associated protein 1-like protein	<u>Immunodeficiency 72 with autoinflammation</u>	<u>F6SQF6</u>	KH
<i>NAXE</i> (Q8NCW5)	NAD(P)HX epimerase	<u>Encephalopathy, progressive, early-onset, with brain edema and/or leukoencephalopathy</u>	<u>H2Y1D7</u>	KH
<i>RBBP7</i> (Q16576)	RB binding protein 7, chromatin remodeling factor	<u>Spermatogenic failure, X-linked, 9</u>	<u>A0A1W3JEA6</u>	KH
<i>GAS2L2</i> (Q8NHY3)	Growth arrest specific 2-like 2 protein	<u>Primary ciliary dyskinesia 41</u>	<u>H2XPE8</u>	KH
<i>DHODH</i> (Q02127)	Dihydroorotate dehydrogenase (quinone)	<u>Miller syndrome</u>	<u>F6UA87</u>	KH
<i>CWF19L1</i> (Q69YN2)	CWF19 like cell cycle control factor 1	<u>Spinocerebellar ataxia, autosomal recessive 17</u>	<u>F6T2S3</u>	KH
<i>VWA8</i> (A3KMH1)	Von Willebrand factor A domain-containing 8	<u>Retinitis pigmentosa 97</u>	<u>F6QXZ7</u>	KH

Table 1

Summary of the relevant diseases and our initial guess for phenotypes for each short-list gene.

* = not found in KH-KY-mapping Excel spreadsheet (see Methods).

Expert feedback and our final shortlist

We sought expert feedback on our short list from Alberto Stolfi, a tunicate researcher at the Georgia Institute of Technology, one of the external scientists we consulted as part of this work. After sharing our initial shortlist of genes, our expert advisor noted that many *Ciona* genes exhibit maternal expression, which can impede the ability to observe phenotypes in CRISPR knockout embryos. Maternal expression in *Ciona*, as observed in bulk RNA-seq data and according to experts, is generally characterized by high expression at the earliest developmental stages without a substantial increase at later developmental stages. For example, NAXE expression appears to be highest at the one-cell stage and shows a continual decrease in relative expression until hatching ([Figure 3](#)). For this reason, we decided not to pursue NAXE for our initial experiments. Other genes we decided not to pursue due to the potential for maternal expression were *Gas2l2*, *Dhodh*, and *Cwf19l1* (see [supplementary figure](#)). While RNAi [\[50\]](#) and morpholino [\[51\]](#) methods are also feasible in *Ciona* and able to suppress maternal gene expression, we decided to stick to a single method of genetic manipulation for our experimental plan.

Two homologs from our analyses seemed not to have strict maternal expression: *Rbbp7* and *Vwa8*. One function of RBBP7 is as a member of human polycomb repressive complex 2 (PRC2) [\[52\]](#); morpholinos against the *Ciona* homolog of *Enhancer of zeste* (E(z)), another PRC2 component, appear to have morphological consequences throughout development [\[53\]](#). VWA8 is a mitochondrial matrix-targeted protein with ATPase activity associated with a subtype of retinitis pigmentosa [\[54\]](#). Morpholino knockdown of zebrafish *Vwa8* causes gross morphological defects and degraded retinal cells [\[54\]\[55\]](#). For these two genes, we weren't sure exactly what developmental stages, tissues, or cell types would present a clear phenotype. Identifying a phenotype might require first evaluating gene expression through *in situ* hybridization or transgenic expression, and then generating knockouts and screening for defects across development, which could be time-consuming.

After discussion with experts, we decided that it would be more efficient to focus on a single, well-established assay in *Ciona*, rather than to survey for potentially unknown phenotypes across development. Broader phenotypic profiling in *Ciona* could be time-consuming or difficult to develop, given that phenotypes could emerge in various tissues across developmental time. Focusing our work on phenotypes in the notochord also allows us to control for the possibility of non-cell-autonomous effects by generating tissue-specific knockouts. Based on this assay choice, we decided to

focus on three genes: *Fcho1*, *Pgm3*, and *Nckap1l*. All three of these homologs have the potential for phenotypes in the notochord. Below is a rough overview of our experimental plan.

Experimental plan

1. **Generate and validate mutants in G0 embryos.**
2. **Phenotype notochord effects.**
3. **Generate tissue-specific mutants.** To evaluate whether the notochord effects are non-cell-autonomous, generate mutants using a notochord-specific promoter to express Cas9.
4. **Perform genetic rescue experiments using complementation with wild-type protein.**

For details about the technical analyses described in this pub, see [the associated pub](#) describing our evaluation framework [1].

All **code** related to this pub is available in [this GitHub repo](#) (DOI: [10.5281/zenodo.15707938](https://doi.org/10.5281/zenodo.15707938)). Code specific to *C. intestinalis* is in [these analysis notebooks](#) and [this set of visualization scripts](#).

Ciona identifier mapping

After consulting with experts, we learned that the *Ciona* community doesn't use UniProt identifiers regularly in their work. For our organism selection work and in the Zoogler interface, we rely on these identifiers to distinguish proteins in the *Ciona* proteome. Thus, to bridge our data and relevant gene-expression resources, we needed maps between UniProt identifiers and other identifiers. The community uses multiple gene and protein identifiers for previous versions of the *Ciona* genome, including "KH" identifiers for the KH2012 genome ([GCF_000224145.1](#)) (used in

resources like ANISEED [56]) and “KY” identifiers for the KY2021 genome [57] (used for a recently re-analyzed Ciona single-cell transcriptome, described at the top of the next section). To generate a map between the KY identifiers and UniProt identifiers, we ran all-versus-all BLASTp [58] using default parameters, setting the KY proteome as the database and the UniProt proteome as the query. For each KY protein, we selected the top UniProt identifier as a crude best match. We also obtained an Excel spreadsheet from Alberto Stolfi containing previously generated maps between the KY and KH identifiers, which we provide in our GitHub repo. Identifier mapping functionality can be found in our zoogletools Python package.

You can find a copy of the **KH2012 to KY2021 identifier map** [here](#) in our GitHub repository. The **zoogletools** Python package is available [here](#).

***Ciona* single-cell developmental gene expression**

We visualized single-cell RNA-seq data analyzed by Katarzyna Piekarz, as described in [this GitHub repository](#) using scripts in our [zoogletools](#) Python package. This [dataset](#) is a re-analysis of data from a 2019 paper by Cao et al. [59]. The major difference between the two datasets is that Cao et al. mapped their reads to the KH2012 genome, whereas Piekarz mapped their data to the KY2021 genome. In reusing the Piekarz data, we identified some discrepancies in file names, replicate labels, and developmental stage labels, which are corrected by helper scripts in the zoogletools package. You can find a description of those issues in [this GitHub pull request](#). We’ve discussed some of these discrepancies with the authors of the Piekarz repo in [this GitHub issue](#). There may be problems with stage labeling in the data released with the Cao et al. manuscript, with potential inconsistencies between the [Gene Expression Omnibus deposition](#) and the [Broad Institute Single Cell Portal](#). We treated the cell labels in the portal dataset as a ground truth for the purposes of our exploration.

In addition to visualizing expression counts from the Piekarz re-analysis of the Cao et al. manuscript, we downloaded tissue type annotations from Cao et al. from the [Broad Institute Single Cell Portal](#) [60] and merged them into the Piekarz re-analysis. The [zoogletools](#) Python package also provides functionality for performing this merge operation.

***Ciona* bulk developmental gene expression**

One major barrier to observing CRISPR knockout phenotypes can be the presence of high levels of maternally deposited RNA in *Ciona* embryos [61]. To understand the relative expression of genes across developmental stages, we visualized bulk RNA-seq data downloaded from the ANISEED database of *Ciona* gene expression [56]. This database uses the KH2012 genome for its identifiers. The [zoogletools](#) Python package provides scripts to generate gene expression visualizations from these data.

AI tool usage

We used Claude to help write code, clean up code, comment our code, draft text that we edited, suggest wording ideas for small phrases and sentence structures, and clarify and streamline text that we wrote. We also used Cursor to access Claude.

Visualization

We used plotly (v5.17.0) [62] arcadia-pycolor (v0.6.3) [63] to generate figures before manual adjustment.

Advantages of *Ciona* as a disease model

During our diligence process, we came to realize a variety of advantages that *Ciona* has as a potential model for human diseases, including:

- **Rapid development.** Organoids derived from human iPSCs and mouse embryos have long differentiation or developmental timelines. *Ciona* embryos develop rapidly, reaching the larval stage overnight, allowing for rapid examination of genetic effects on diverse tissue systems during development.

- **Morphogenetic processes.** *Ciona* embryos undergo a variety of relevant morphogenetic processes, such as cell intercalation, division, lumen formation, and others. These allow scientists to probe the effects of genes and drugs on such processes.
- **Ease of genetic experimentation.** *Ciona* embryos can be subjected to transgenic overexpression, tissue-specific expression, RNAi, and gene knockout using electroporation or microinjection approaches. Moreover, *Ciona* embryos can be collected in the thousands, allowing for experimentation at scale.

Challenges of working with community resources

A major hurdle to our diligence work was learning how to access and use different resources maintained by the *Ciona* community. Resources such as the Ghost database [64] and ANISEED [56] contain a wealth of information, but use community-derived identifiers (see [this explanation](#)) that aren't easy to understand without speaking with experts who can describe their history.

In addition to resolving gene identifiers, we also faced challenges in understanding and using publicly available data from the *Ciona* community. For example, two analyses of the same dataset generated by Cao et al., one by the original authors and another by a community member, had conflicting labels for developmental stages and replicates. We weren't able to determine the source of this discrepancy, which could have arisen within the original dataset.

We winnowed our list down to three genes that are promising to investigate in *Ciona*. Others who evaluate the same genes may arrive at different conclusions about what's practical to pursue, based on their expertise. Among the genes we decided not to study, we'd encourage others to try to generate mutants or knockdowns of NAXE to see whether there's any temperature-dependent effect on development, and whether supplementation with niacin derivatives could rescue mutant phenotypes. We also

encourage others to check out our [working diligence document](#) and comment on whether you agree with our feasibility determinations.

We're funding downstream work in the [Alberto Stolfi lab](#) at the Georgia Institute of Technology, who'll also share their results openly during the course of their experiments. Stay tuned for more information.

- 1 10.57844/arcadia-ugg5-emyd
- 2 Brunetti R, Gissi C, Pennati R, Caicci F, Gasparini F, Manni L. (2015). Morphological evidence that the molecularly determined *Ciona intestinalis* type A and type B are different species: *Ciona robusta* and *Ciona intestinalis*. <https://doi.org/10.1111/jzs.12101>
- 3 Passamanek YJ, Di Gregorio A. (2005). *Ciona intestinalis*: Chordate development made simple. <https://doi.org/10.1002/dvdy.20300>
- 4 Delsuc F, Philippe H, Tsagkogeorga G, Simion P, Tilak M-K, Turon X, López-Legentil S, Piette J, Lemaire P, Douzery EJP. (2018). A phylogenomic framework and timescale for comparative studies of tunicates. <https://doi.org/10.1186/s12915-018-0499-2>
- 5 Nanglu K, Lerosey-Aubril R, Weaver JC, Ortega-Hernández J. (2023). A mid-Cambrian tunicate and the deep origin of the ascidiacean body plan. <https://doi.org/10.1038/s41467-023-39012-4>
- 6 Kourakis MJ, Smith WC. (2015). An organismal perspective on *C. intestinalis* development, origins and diversification. <https://doi.org/10.7554/elife.06024>
- 7 Hotta K, Mitsuhashi K, Takahashi H, Inaba K, Oka K, Gojobori T, Ikeo K. (2007). A web-based interactive developmental table for the ascidian *Ciona intestinalis*, including 3D real-image embryo reconstructions: I. From fertilized egg to hatching larva. <https://doi.org/10.1002/dvdy.21188>
- 8 Conklin E. (1905). The organization and cell-lineage of the ascidian egg/by Edwin G

Conklin, E. G. "The organization and cell-lineage of the ascidian egg/by Edwin G." Conklin.[Academy of Natural Sciences] (1905).

- 9 Satoh N. (1994). Developmental biology of ascidians

Satoh, Noriyuki. "Developmental biology of ascidians." (*No Title*) (1994).

- 10 Kari W, Zeng F, Zitzelsberger L, Will J, Rothbacher U. (2016). Embryo Microinjection and Electroporation in the Chordate *Ciona intestinalis* <https://doi.org/10.3791/54313>
- 11 Jindal GA, Lim F, Tellez K, Song BP, Bantle AT, Farley EK. (2024). Protocol to electroporate DNA plasmids into *Ciona robusta* embryos at the 1-cell stage. <https://doi.org/10.1016/j.xpro.2024.103107>
- 12 Gandhi S, Razy-Krajka F, Christiaen L, Stolfi A. (2018). CRISPR Knockouts in *Ciona* Embryos. https://doi.org/10.1007/978-981-10-7545-2_13
- 13 Popsuj S, Cohen L, Ward S, Lewis A, Yoshida S, Herrera RA, Cota CD, Stolfi A. (2024). CRISPR/Cas9 Protocols for Disrupting Gene Function in the Non-vertebrate Chordate *Ciona*. <https://doi.org/10.1093/icb/icae108>
- 14 Satou Y, Nakamura R, Yu D, Yoshida R, Hamada M, Fujie M, Hisata K, Takeda H, Satoh N. (2019). A Nearly Complete Genome of *Ciona intestinalis* Type A (*C. robusta*) Reveals the Contribution of Inversion to Chromosomal Evolution in the Genus *Ciona*. <https://doi.org/10.1093/gbe/evz228>
- 15 Pickett CJ, Zeller RW. (2018). Efficient genome editing using CRISPR-Cas-mediated homology directed repair in the ascidian *Ciona robusta*. <https://doi.org/10.1002/dvg.23260>
- 16 Stolfi A, Christiaen L. (2012). Genetic and Genomic Toolbox of the Chordate *Ciona intestinalis*. <https://doi.org/10.1534/genetics.112.140590>
- 17 Christiaen L, Wagner E, Shi W, Levine M. (2009). Whole-Mount In Situ Hybridization on Sea Squirt (*Ciona intestinalis*) Embryos. <https://doi.org/10.1101/pdb.prot5348>
- 18 Sakurai H, Shiba K, Takamura K, Inaba K. (2024). Immunohistochemical Characterization of Spermatogenesis in the Ascidian *Ciona robusta*. <https://doi.org/10.3390/cells13221863>
- 19 Sasakura Y, Oogai Y, Matsuoka T, Satoh N, Awazu S. (2007). Transposon mediated transgenesis in a marine invertebrate chordate: *Ciona intestinalis*.

<https://doi.org/10.1186/gb-2007-8-s1-s3>

- 20 Cao C, Lemaire LA, Wang W, Yoon PH, Choi YA, Parsons LR, Matese JC, Wang W, Levine M, Chen K. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. <https://doi.org/10.1038/s41586-019-1385-y>
- 21 Antzelevitch C, Brugada P, Borggrefe M, Brugada J, Brugada R, Corrado D, Gussak I, LeMarec H, Nademanee K, Perez Riera AR, Shimizu W, Schulze-Bahr E, Tan H, Wilde A. (2005). Brugada Syndrome: Report of the Second Consensus Conference. <https://doi.org/10.1161/01.cir.0000152479.54298.51>
- 22 Keen TJ, Hims MM, McKie AB, Moore AT, Doran RM, Mackey DA, Mansfield DC, Mueller RF, Bhattacharya SS, Bird AC, Markham AF, Inglehearn CF. (2002). Mutations in a protein target of the Pim-1 kinase associated with the RP9 form of autosomal dominant retinitis pigmentosa. <https://doi.org/10.1038/sj.ejhg.5200797>
- 23 Sullivan LS, Bowne SJ, Birch DG, Hughbanks-Wheaton D, Heckenlively JR, Lewis RA, Garcia CA, Ruiz RS, Blanton SH, Northrup H, Gire AI, Seaman R, Duzkale H, Spellicy CJ, Zhu J, Shankar SP, Daiger SP. (2006). Prevalence of Disease-Causing Mutations in Families with Autosomal Dominant Retinitis Pigmentosa: A Screen of Known Genes in 200 Families. <https://doi.org/10.1167/iovs.05-1443>
- 24 Bischof JM, Chiang AP, Scheetz TE, Stone EM, Casavant TL, Sheffield VC, Braun TA. (2006). Genome-wide identification of pseudogenes capable of disease-causing gene conversion. <https://doi.org/10.1002/humu.20335>
- 25 Brown GK, Scholem RD, Bankier A, Danks DM. (1983). Malonyl coenzyme a decarboxylase deficiency. <https://doi.org/10.1007/bf01805615>
- 26 Avasthi P, McGeever E, Patton AH, York R. (2024). Leveraging evolution to identify novel organismal models of human biology. <https://doi.org/10.57844/ARCADIA-33B4-4DC5>
- 27 Zhang Y. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. <https://doi.org/10.1093/nar/gki524>
- 28 Henne WM, Boucrot E, Meinecke M, Evergren E, Vallis Y, Mittal R, McMahon HT. (2010). FCHo Proteins Are Nucleators of Clathrin-Mediated Endocytosis. <https://doi.org/10.1126/science.1188462>
- 29 Łyszkiewicz M, Ziętara N, Frey L, Pannicke U, Stern M, Liu Y, Fan Y, Puchałka J, Hollizeck S, Somekh I, Rohlf M, Yilmaz T, Ünal E, Karakukcu M, Patiroğlu T, Kellerer C, Karasu E, Sykora K-W, Lev A, Simon A, Somech R, Roesler J, Hoenig M, Keppler OT, Schwarz K, Klein C. (2020). Human FCHO1 deficiency reveals role

for clathrin-mediated endocytosis in development and function of T cells.

<https://doi.org/10.1038/s41467-020-14809-9>

- 30 Calzoni E, Platt CD, Keles S, Kuehn HS, Beaussant-Cohen S, Zhang Y, Pazmandi J, Lanzi G, Pala F, Tahiat A, Artac H, Heredia RJ, Dmytrus J, Reisli I, Uygun V, Uygun D, Bingol A, Basaran E, Djenouhat K, Benhalla N, Bendahmane C, Emiroglu M, Kirchhausen T, Pasham M, Jones J, Wallace JG, Zheng L, Boisson B, Porta F, Rosenzweig SD, Su H, Giliani S, Lenardo M, Geha RS, Boztug K, Chou J, Notarangelo LD. (2019). F-BAR domain only protein 1 (FCHO1) deficiency is a novel cause of combined immune deficiency in human subjects.
<https://doi.org/10.1016/j.jaci.2019.02.014>
- 31 Ouyang X, Wu B, Yu H, Dong B. (2023). DYRK1-mediated phosphorylation of endocytic components is required for extracellular lumen expansion in ascidian notochord. <https://doi.org/10.1186/s40659-023-00422-9>
- 32 Greig KT, Antonchuk J, Metcalf D, Morgan PO, Krebs DL, Zhang J-G, Hacking DF, Bode L, Robb L, Kranz C, de Graaf C, Bahlo M, Nicola NA, Nutt SL, Freeze HH, Alexander WS, Hilton DJ, Kile BT. (2007). Agm1/Pgm3-Mediated Sugar Nucleotide Synthesis Is Essential for Hematopoiesis and Development.
<https://doi.org/10.1128/mcb.00802-07>
- 33 Yang L, Zerbato B, Pessina A, Brambilla L, Andreani V, Frey-Jakobs S, Fliegauf M, Barbouche M-R, Zhang Q, Chiaradonna F, Proietti M, Du X, Grimbacher B. (2024). PGM3 insufficiency: a glycosylation disorder causing a notable T cell defect.
<https://doi.org/10.3389/fimmu.2024.1500381>
- 34 Winslow A, Jalazo ER, Evans A, Winstead M, Moran T. (2022). A De Novo Cause of PGM3 Deficiency Treated with Hematopoietic Stem Cell Transplantation.
<https://doi.org/10.1007/s10875-021-01196-z>
- 35 Stray-Pedersen A, Backe PH, Sorte HS, Mørkrid L, Chokshi NY, Erichsen HC, Gambin T, Elgstøen KBP, Bjørås M, Wlodarski MW, Krüger M, Jhangiani SN, Muzny DM, Patel A, Raymond KM, Sasa GS, Krance RA, Martinez CA, Abraham SM, Speckmann C, Ehl S, Hall P, Forbes LR, Merckoll E, Westvik J, Nishimura G, Rustad CF, Abrahamsen TG, Rønnestad A, Osnes LT, Egeland T, Rødningen OK, Beck CR, Boerwinkle EA, Gibbs RA, Lupski JR, Orange JS, Lausch E, Hanson IC. (2014). PGM3 Mutations Cause a Congenital Disorder of Glycosylation with Severe Immunodeficiency and Skeletal Dysplasia.
<https://doi.org/10.1016/j.ajhg.2014.05.007>
- 36 Yagi H, Nakagawa M, Takahashi N, Kondo S, Matsubara M, Kato K. (2007). Neural complex-specific expression of xylosyl N-glycan in *Ciona intestinalis*.
<https://doi.org/10.1093/glycob/cwm128>

- 37 Wang Z, Tan Z, Bi J, Liu A, Jiang A, Dong B. (2023). Proteomic identification of intracellular vesicle trafficking and protein glycosylation requirements for lumen inflation in *Ciona* notochord. <https://doi.org/10.1002/pmic.202200460>
- 38 Ricciardiello F, Votta G, Palorini R, Raccagni I, Brunelli L, Paiotta A, Tinelli F, D'Orazio G, Valtorta S, De Gioia L, Pastorelli R, Moresco RM, La Ferla B, Chiaradonna F. (2018). Inhibition of the Hexosamine Biosynthetic Pathway by targeting PGM3 causes breast cancer growth arrest and apoptosis. <https://doi.org/10.1038/s41419-018-0405-4>
- 39 Park H, Staehling-Hampton K, Appleby MW, Brunkow ME, Habib T, Zhang Y, Ramsdell F, Liggitt HD, Freie B, Tsang M, Carlson G, Friend S, Frevert C, Iritani BM. (2008). A point mutation in the murine *Hem1* gene reveals an essential role for Hematopoietic Protein 1 in lymphopoiesis and innate immunity. <https://doi.org/10.1084/jem.20080340>
- 40 Castro CN, Rosenzweig M, Carapito R, Shahrooei M, Konantz M, Khan A, Miao Z, Groß M, Tranchant T, Radosavljevic M, Paul N, Stemmelen T, Pitoiset F, Hirschler A, Nespola B, Molitor A, Rolli V, Pichot A, Faletti LE, Rinaldi B, Friant S, Mednikov M, Karauzum H, Aman MJ, Carapito C, Lengerke C, Ziaee V, Eyaid W, Ehl S, Alroqi F, Parvaneh N, Bahram S. (2020). *NCKAP1L* defects lead to a novel syndrome combining immunodeficiency, lymphoproliferation, and hyperinflammation. <https://doi.org/10.1084/jem.20192275>
- 41 Cook SA, Comrie WA, Poli MC, Similuk M, Oler AJ, Faruqi AJ, Kuhns DB, Yang S, Vargas-Hernández A, Carisey AF, Fournier B, Anderson DE, Price S, Smelkinson M, Abou Chahla W, Forbes LR, Mace EM, Cao TN, Coban-Akdemir ZH, Jhangiani SN, Muzny DM, Gibbs RA, Lupski JR, Orange JS, Cuvelier GDE, Al Hassani M, Al Kaabi N, Al Yafei Z, Jyonouchi S, Raje N, Caldwell JW, Huang Y, Burkhardt JK, Latour S, Chen B, ElGhazali G, Rao VK, Chinn IK, Lenardo MJ. (2020). HEM1 deficiency disrupts mTORC2 and F-actin control in inherited immunodysregulatory disease. <https://doi.org/10.1126/science.aay5663>
- 42 Dong B, Deng W, Jiang D. (2011). Distinct cytoskeleton populations and extensive crosstalk control *Ciona* notochord tubulogenesis. <https://doi.org/10.1242/dev.057208>
- 43 Marbaix AY, Tyteca D, Niehaus TD, Hanson AD, Linster CL, Van Schaftingen E. (2014). Occurrence and subcellular distribution of the NAD(P)HX repair system in mammals. <https://doi.org/10.1042/bj20131482>
- 44 Kremer LS, Danhauser K, Herebian D, Petkovic Ramadža D, Piekutowska-Abramczuk D, Seibt A, Müller-Felber W, Haack TB, Płoski R, Lohmeier K, Schneider D, Klee D, Rokicki D, Mayatepek E, Strom TM, Meitinger T, Klopstock T, Pronicka E, Mayr JA, Baric I, Distelmaier F, Prokisch H. (2016). NAXE Mutations

Disrupt the Cellular NAD(P)HX Repair System and Cause a Lethal Neurometabolic Disorder of Early Childhood.
<https://doi.org/10.1016/j.ajhg.2016.07.018>

- 45 Manor J, Calame D, Gijavanekar C, Fisher K, Hunter J, Mizerik E, Bacino C, Scaglia F, Elsea SH. (2022). NAXE deficiency: A neurometabolic disorder of NAD(P)HX repair amenable for metabolic correction.
<https://doi.org/10.1016/j.ymgme.2022.04.003>
- 46 Ozaki K, Yatsuka Y, Oyazato Y, Nishiyama A, Nitta KR, Kishita Y, Fushimi T, Shimura M, Noma S, Sugiyama Y, Tagami M, Fukunaga M, Kinoshita H, Hirata T, Suda W, Murakawa Y, Carninci P, Ohtake A, Murayama K, Okazaki Y. (2024). Biallelic GGGCC repeat expansion leading to NAXE-related mitochondrial encephalopathy. <https://doi.org/10.1038/s41525-024-00429-5>
- 47 Kim J, Zhu L, Sun Q, Fang L. (2021). Systemic metabolite profiling reveals sexual dimorphism of AIBP control of metabolism in mice.
<https://doi.org/10.1371/journal.pone.0248964>
- 48 Marbaix AY, Noël G, Detroux AM, Vertommen D, Van Schaftingen E, Linster CL. (2011). Extremely Conserved ATP- or ADP-dependent Enzymatic System for Nicotinamide Nucleotide Repair. <https://doi.org/10.1074/jbc.c111.310847>
- 49 Niehaus TD, Richardson LGL, Gidda SK, ElBadawi-Sidhu M, Meissen JK, Mullen RT, Fiehn O, Hanson AD. (2014). Plants Utilize a Highly Conserved System for Repair of NADH and NADPH Hydrates . <https://doi.org/10.1104/pp.114.236539>
- 50 Nishiyama A, Fujiwara S. (2008). RNA interference by expressing short hairpin RNA in the *Ciona intestinalis* embryo. <https://doi.org/10.1111/j.1440-169x.2008.01039.x>
- 51 Satou Y, Imai KS, Satoh N. (2001). Action of morpholinos in *Ciona* embryos.
<https://doi.org/10.1002/gene.1040>
- 52 Shi Y, Wang X, Zhuang Y, Jiang Y, Melcher K, Xu HE. (2017). Structure of the PRC2 complex and application to drug discovery. <https://doi.org/10.1038/aps.2017.7>
- 53 Le Goff E, Martinand-Mari C, Martin M, Feuillard J, Boublik Y, Godefroy N, Mangeat P, Baghdiguian S, Cavalli G. (2015). Enhancer of zeste acts as a major developmental regulator of *Ciona intestinalis* embryogenesis.
<https://doi.org/10.1242/bio.010835>
- 54 Kong L, Chu G, Ma W, Liang J, Liu D, Liu Q, Wei X, Jia S, Gu H, He Y, Luo W, Cao S, Zhou X, He R, Yuan Z. (2023). Mutations in *VWA8* cause autosomal-dominant retinitis pigmentosa via aberrant mitophagy activation.
<https://doi.org/10.1136/jmg-2022-108888>

- 55 Umair M, Farooq Khan M, Aldrees M, Nashabat M, Alhamoudi KM, Bilal M, Alyafee Y, Al Tuwaijri A, Aldarwish M, Al-Rumayyan A, Alkhalaf H, Wadaan MAM, Alfadhel M. (2021). Mutated VWA8 Is Associated With Developmental Delay, Microcephaly, and Scoliosis and Plays a Novel Role in Early Development and Skeletal Morphogenesis in Zebrafish. <https://doi.org/10.3389/fcell.2021.736960>
- 56 Brozovic M, Dantec C, Dardaillon J, Dauga D, Faure E, Gineste M, Louis A, Naville M, Nitta KR, Piette J, Reeves W, Scornavacca C, Simion P, Vincentelli R, Bellec M, Aicha SB, Fagotto M, Guérout-Bellone M, Haeussler M, Jacox E, Lowe EK, Mendez M, Roberge A, Stolfi A, Yokomori R, Brown CT, Cambillau C, Christiaen L, Delsuc F, Douzery E, Dumollard R, Kusakabe T, Nakai K, Nishida H, Satou Y, Swalla B, Veeman M, Volff J-N, Lemaire P. (2017). ANISEED 2017: extending the integrated ascidian database to the exploration and evolutionary comparison of genome-scale datasets. <https://doi.org/10.1093/nar/gkx1108>
- 57 Satou Y, Tokuoka M, Oda-Ishii I, Tokuhiko S, Ishida T, Liu B, Iwamura Y. (2022). A Manually Curated Gene Model Set for an Ascidian, *Ciona robusta* (*Ciona intestinalis* Type A). <https://doi.org/10.2108/zs210102>
- 58 Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. (2009). BLAST+: architecture and applications. <https://doi.org/10.1186/1471-2105-10-421>
- 59 Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, Adey A, Waterston RH, Trapnell C, Shendure J. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. <https://doi.org/10.1126/science.aam8940>
- 60 Tarhan L, Bistline J, Chang J, Galloway B, Hanna E, Weitz E. (2023). Single Cell Portal: an interactive home for single-cell genomics data. <https://doi.org/10.1101/2023.07.13.548886>
- 61 Matsuoka T, Ikeda T, Fujimaki K, Satou Y. (2013). Transcriptome dynamics in early embryos of the ascidian, *Ciona intestinalis*. <https://doi.org/10.1016/j.ydbio.2013.10.003>
- 62 Plotly Technologies Inc. (2015). Collaborative data science. <https://plot.ly>
- 63 arcadia-pycolor. (2025). <https://github.com/Arcadia-Science/arcadia-pycolor>
- 64 Satou Y, Kawashima T, Shoguchi E, Nakayama A, Satoh N. (2005). An Integrated Database of the Ascidian, *Ciona intestinalis*: Towards Functional Genomics. <https://doi.org/10.2108/zsj.22.837>

