microPublication BIOLOGY

6/23/2026 - *Open Access*

# Bioinformatic pipeline to identify candidate mRNA transcripts targeted by the putative Regulated Ire1-Dependent Decay pathway of *Dictyostelium discoideum*.

Connor Bingham[1], Elias Taylor-Cornejo[1§]

[1]Department of Biology, Randolph–Macon College, Ashland, VA, United States

[§]To whom correspondence should be addressed: EliasTaylor-Cornejo@rmc.edu

## Abstract

Inositol-requiring protein 1 (Ire1) is a eukaryotic stress sensor that counteracts the buildup of unfolded proteins in the endoplasmic reticulum (ER) by activating the Unfolded Protein Response (UPR) via a specific ribonuclease (RNase) activity. The amoeba *Dictyostelium discoideum* relies on an *ire1* ortholog, *ireA*, to survive ER stress, but the mRNA transcripts targeted by the IreA ribonuclease remain unknown. In this work, we developed a bioinformatic pipeline that identified 21 mRNA transcripts of *D. discoideum* that contain a consensus Ire1 cut site found within a secondary mRNA hairpin loop structure and have the potential to be cut by IreA.

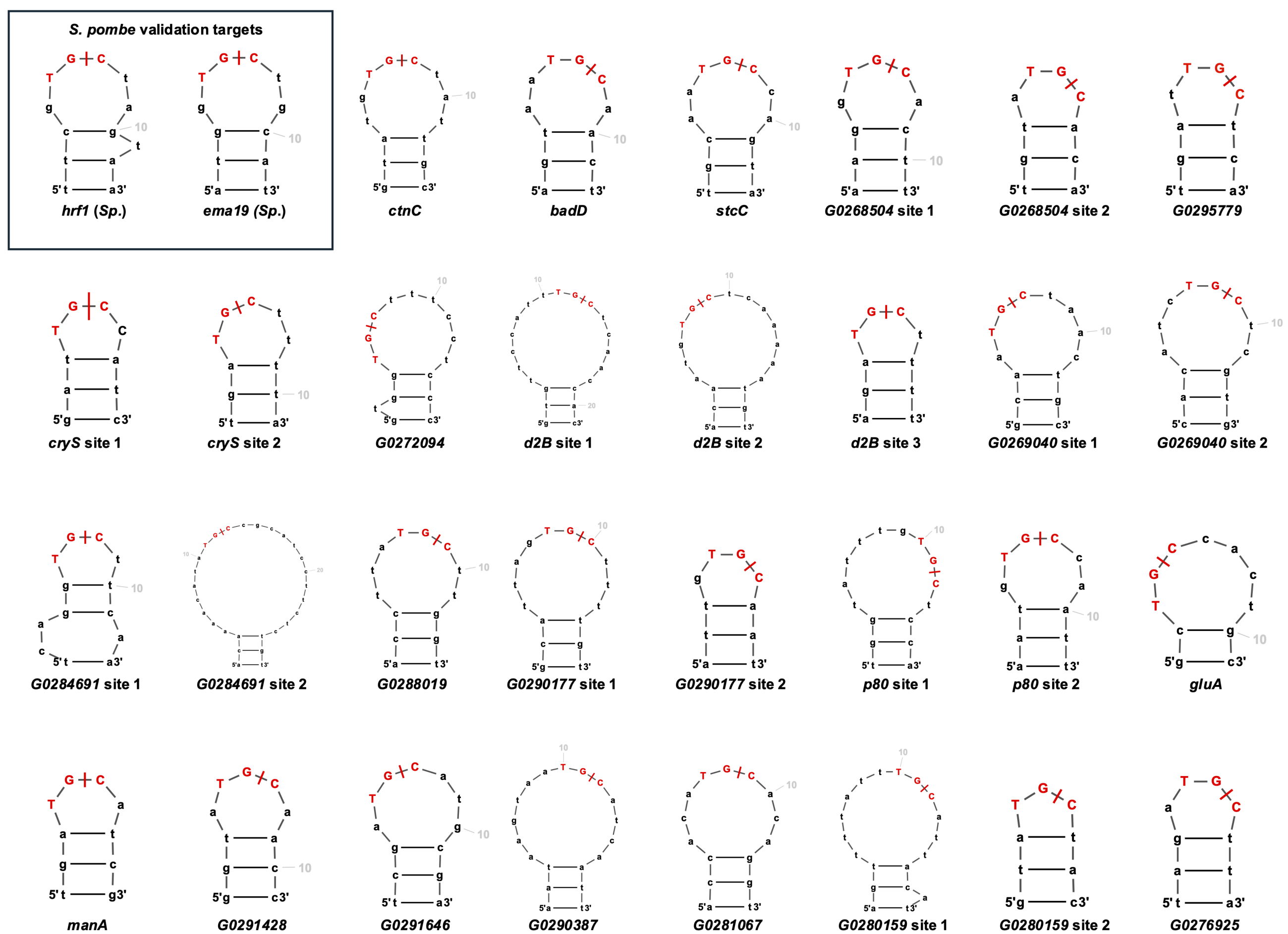


**Figure 1. Predicted stem-loop structures containing Ire1 cleavage sites in candidate RIDD transcripts:**

Predicted secondary structures of mRNA hairpin loops containing the consensus Ire1 cleavage site for the 21 candidate RIDD transcripts identified in *D. discoideum*. Two known Ire1 cleavage targets from *S. pombe* are included as representative validation examples: *ema19*, for which the pipeline identified an exact match to the known cleavage site, and *hrf1*, for which the pipeline identified the expected TGC cut site on a predicted stem-loop structure. All sequences are represented as cDNA; the mRNA cleavage sequence UGC therefore appears as TGC. Genes with multiple predicted cleavage sites are shown as separate diagrams labeled site 1, site 2, or site 3. For clarity, gene identifiers beginning with DDB_G0 have been shortened to G0 followed by the unique identifying number. Red nucleotides indicate the predicted

TGC cleavage sequence with the red line indicating the cleavage site. RNA secondary structures were predicted by RNAfold and visualized with the software RNA 2D Template (R2DT) (McCann et al. 2025).

## Description

Within eukaryotic cells, transmembrane and secreted proteins are made by ribosomes located on the endoplasmic reticulum (ER) (Alberts et al. 2002). For these critical proteins to be functional, they must be properly folded by chaperones located within the ER (Alberts et al. 2002). However, when an ER attempts to make too many proteins at once, it causes ER stress by overwhelming the protein synthesis machinery, which can be lethal to cells if not mitigated (Walter and Ron 2011). The Unfolded Protein Response (UPR) is the system responsible for resolving ER stress by sensing and counteracting the buildup of unfolded proteins in the ER (Walter and Ron 2011).

Inositol-requiring enzyme-1 (Ire1) is a highly conserved UPR sensor that is vital for surviving ER stress (Cox and Walter 1996). Ire1 is a ribonuclease (RNase) that is capable of detecting the buildup of unfolded proteins in the ER and responds by cutting specific ER-targeted mRNAs to trigger mechanisms that mitigate ER stress (Cox and Walter 1996; Oikawa et al. 2010; Maurel et al. 2014). The mRNA transcripts that are cut by Ire1 all contain a specific Ire1 cut sequence (UGC) present on an mRNA hairpin loop structure (Kimmig et al. 2012). The mRNA transcripts cut by Ire1 also typically possess a signal peptide sequence that directs them to the ER via the secretory pathway and brings them in contact with Ire1, which resides in the ER membrane (Coelho and Domingos 2014). There are currently two primary mechanisms by which Ire1 cuts mRNAs in response to ER stress in yeast: mRNA splicing and regulated Ire1-dependent mRNA decay (RIDD) (Cox and Walter 1996; Maurel et al. 2014). In mRNA splicing, Ire1 removes an intron from a specific mRNA transcript, *hac1*, which enables the translation of the Hac1 transcription factor that upregulates genes associated with improved protein folding (Cox and Walter 1996). For splicing of *hac1* to occur, Ire1 cuts *hac1* mRNA transcripts in two separate hairpin loops that are then joined by RNA ligases (Sidrauski et al. 1996). On the other hand, in RIDD, Ire1 assists to cut and destroy ER-targeted mRNA transcripts to reduce unnecessary protein production at the ER until proper protein folding can be restored (Maurel et al. 2014). In this method, Ire1 cuts these mRNA targets at a single hairpin loop, where they are subsequently degraded by exonucleases (Hollien and Weissman 2006; Maurel et al. 2014).

In metazoans, such as fruit flies and mammals, Ire1 can use both mRNA slicing and RIDD to regulate the UPR in response to ER stress (Moore and Hollien 2015). However, in microbes such as yeast, Ire1 appears to mainly use either mRNA splicing or RIDD to activate their UPR. For example, the yeast species *Saccharomyces cerevisiae* exclusively performs *hac1* mRNA splicing, not RIDD, in response to ER stress (Niwa et al. 2004). However, a closely related yeast species *Schizosaccharomyces pombe* lacks an ortholog for *hac1* and therefore cannot perform *hac1* mRNA splicing; instead, *S. pombe* responds to ER stress primarily through the RIDD pathway (Kimmig et al. 2012).

Previous studies have shown that the amoeba *D. discoideum* can activate the UPR via an *ire1* ortholog called *ireA*, which is required to survive ER stress, but the specific RNase function (mRNA splicing or RIDD) is unknown (Domínguez-Martín et al. 2018). Furthermore, the mRNA transcripts that are cut by IreA in response to ER stress have not been identified, and *D. discoideum* lack an ortholog of *hac1* (Domínguez-Martín et al. 2018). Due to the lack of a *hac1* ortholog, we hypothesize that IreA responds to ER stress using the RIDD mechanism. To test this hypothesis, we developed a bioinformatic pipeline to identify ER-targeted mRNA transcripts from *D. discoideum* that contain an Ire1 cut site (UGC) found on a predicted mRNA hairpin loop structure that have been shown to be downregulated in an IreA-dependent manner in response to ER stress.

We first evaluated the effectiveness of our bioinformatic pipeline by using it to identify Ire1 cut sites present in mRNA hairpin loops of transcripts from *S. pombe* that are known to be cleaved by Ire1. Like *D. discoideum*, *S. pombe* also lacks a *hac1* orthologue, and is the best microbial reference organism that responds to ER stress using RIDD (Kimmig et al. 2012). Our bioinformatic pipeline was validated against the full set of 11 known mRNA transcripts cleaved by Ire1 in *S. pombe* and correctly identified 8 of the 11 known targets (Kimmig et al. 2012; Guydosh et al. 2017; Li et al. 2018) (Extended Data Table 1). While identification of all 11 targets would be ideal, we considered this level of accuracy sufficient to generate novel biological insights when applied to *D. discoideum*, as the pipeline demonstrated a clear ability to identify the structural and sequence features characteristic of RIDD transcripts.

We then used the validated bioinformatic pipeline to search the *D. discoideum* cDNA library for transcripts that contain an Ire1 consensus cut sequence (UGC) present on a predicted mRNA hairpin loop structure. Our bioinformatic pipeline identified 4128 out of the 13,266 *D. discoideum* transcribed genes having a “UGC” sequence located on an mRNA hairpin-loop. We narrowed down the list of 4128 potential RIDD transcripts to 41 transcripts by cross-referencing a published RNA-seq dataset reporting 96 unique transcripts that are downregulated in an IreA-dependent manner during ER stress in *D. discoideum* (Domínguez-Martín et al. 2018). Of the 41 potential RIDD transcripts, 21 of them had a greater than 95% probability of possessing a standard Sec/SPI signal peptide sequence, indicating that they were extremely likely to be targeted to the ER (Extended Data Table 2). The location of the UGC cleavage site within the predicted stem-loop structure was confirmed for each of the 21 candidate RIDD transcripts using the RNA 2D Templates (R2DT) software (McCann et al. 2025) (Figure 1). There does not appear to be a specific nucleotide position of the UGC

cut sequence within the hairpin loop, or size of the hairpin loop structure, which is consistent with previously characterized RIDD targets from *S. pombe* (Teufel et al. 2022).

Currently, these 21 candidates have not been studied extensively, and their functions remain largely unknown (Extended Data Table 2). Thirteen of the candidates are categorized as proteins of completely unknown function. DDB_G0269040 is broadly associated with reproduction (Eichinger et al. 2005) and DDB_G0281067 is broadly associated with phagocytosis (Gotthardt et al. 2006). DDB_G0288019 is associated with tRNA inosine editing, which allows for the efficient translation of certain codons (Paris et al. 2012). D2B is associated with chemical sensing and breakdown (Baumgardner et al. 2018). Three candidates are associated with responses to bacteria, including p80 through copper transport (Hao et al. 2016), GluA through beta-glucosidase activation (Dimond and Loomis 1976; Munoz-Ruiz et al. 2024), and ManA through alpha-mannosidase activation (Free and Schimke 1976; Munoz-Ruiz et al. 2024). One potential mRNA target, *cryS*, is associated with a significant number of functions within *D. discoideum*, including fruiting body formation and endocytosis (Yuan and Chia 2000). Aside from being synthesized at the ER, the reason for cleaving these 21 specific RIDD transcripts is not entirely clear but could reveal novel insights into how *D. discoideum* responds to ER stress.

Interestingly, 55 of the 96 IreA-dependently downregulated transcripts previously identified (Domínguez-Martín et al. 2018) were not identified as containing a consensus UGC cleavage site within a predicted hairpin loop. We propose two possible explanations for these unaccounted transcripts. First, their downregulation may be an indirect consequence of IreA cleaving a master positive regulator, resulting in reduced expression of its downstream gene targets. Alternatively, IreA may splice an uncharacterized mRNA in *D. discoideum*, producing a transcription factor that indirectly downregulates UPR target genes, analogous to *hac1* mRNA splicing in *S. cerevisiae*. Distinguishing between these possibilities represents an important direction for future experimental work.

A third explanation for the unaccounted transcripts is an important limitation of this bioinformatic pipeline: RNA secondary structure prediction determined purely from nucleotide sequence is not always accurate, which likely accounts for both some of the 55 unaccounted transcripts and the three *S. pombe* targets not recovered during validation. Incorporating experimental genome probing data, such as Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) reactivity data, as structural constraints into RNAfold would likely improve prediction accuracy (Lorenz et al. 2011; Sloma and Mathews 2015). Performing such probing experiments in *D. discoideum* represents an important direction for future work to validate and refine the candidate RIDD targets identified here. Nonetheless, this bioinformatic tool can be broadly applicable to search for a specific nucleotide sequence within a specific RNA structural context.

## Methods

### Bioinformatic pipeline to identify mRNA substrates of IreA

We designed a bioinformatic pipeline to identify mRNA transcripts from *D. discoideum* with the following criteria: (1) must contain at least one instance of the RIDD consensus cut sequence “UGC” (2) the cut sequence must be found within a mRNA hairpin loop structure (3) the mRNA must be downregulated in an IreA-dependent manner during ER stress (4) the mRNA was preferentially selected if it possessed a signal peptide sequence targeting it to the ER via the secretory pathway, as ER-targeted transcripts are most likely to come into contact with IreA at the ER membrane.

A cDNA library representing all known *D. discoideum* mRNA transcript sequences was obtained from GenBank (Eichinger et al. 2005). We created a Python program, named “CutSequenceFinder” that can search cDNA libraries to identify specified sequences that are predicted to be present within RNA hairpin loop structures (available at: https://github.com/Connor-Bingham/CutSequenceFinder/releases/tag/v1.0). To determine secondary structures, stem loops were predicted using RNAfold (ViennaRNA Package v2.0, Lorenz et al. 2011) with the Turner 2004 nearest-neighbor thermodynamic parameters (Mathews et al. 2004), applied at 37°C and a salt concentration of 1.021 M NaCl. Stem loop structures were defined as hairpins containing a minimum of 3 unpaired nucleotides in the loop, consistent with the Turner 2004 nearest-neighbor rules. All stem loops present in the minimum free energy (MFE)-predicted secondary structure were retained; no additional energetic cutoff was applied. The standard output of RNAfold was redirected to a text file to facilitate continued processing. After using CutSequenceFinder to identify transcripts that contain the RIDD cut sequence “UGC” on an RNA hairpin loop, we selected the *D. discoideum* transcripts previously shown to be downregulated during ER stress in an IreA-dependent manner by cross-referencing a published RNA-seq dataset (Domínguez-Martín et al. 2018). Finally, we selected the transcripts that are most likely being targeted to the ER secretory pathway using the SignalP 6.0 webservice to identify the remaining mRNA’s with a predicted N-terminal signal peptide sequence (Teufel et al. 2022). Candidate RIDD transcripts were then visualized with RNA 2D Templates (R2DT) software to confirm the predicted location of the UGC cut sequence within the mRNA hairpin loop (McCann et al. 2025).

**Acknowledgements:** The authors would like to acknowledge the Department of Biology and the Shapiro Undergraduate Research Fellowship (SURF) Program at Randolph-Macon College. In addition we would like to thank Dr. Eunice Dominguez-Martin for her invaluable insight and technical expertise in the *Dictyostelium discoideum* unfolded protein response. Authors declare no conflicts of interest.

## Extended Data

Description: Extended Data Table 1. Ire1 cleavage sites of known S. pombe transcripts.. Resource Type: Dataset. File: Extended Data Table 1.pdf. DOI: 10.22002/61ery-d4b13

Description: Extended Data Table 2. Potential cleavage sites of D. discoideum transcripts.. Resource Type: Dataset. File: Extended Data Table 2.pdf. DOI: 10.22002/w0vme-dy280

## References

Baumgardner K, Lin C, Firtel RA, Lacal J. 2018. Phosphodiesterase PdeD, dynacortin, and a Kelch repeat-containing protein are direct GSK3 substrates in *Dictyostelium* that contribute to chemotaxis towards cAMP. Environmental Microbiology 20: 1888-1903. DOI: 10.1111/1462-2920.14126

Coelho DS, Domingos PM. 2014. Physiological roles of regulated Ire1 dependent decay. Frontiers in Genetics 5: 10.3389/fgene.2014.00076. DOI: 10.3389/fgene.2014.00076

Cox JS, Walter P. 1996. A Novel Mechanism for Regulating Activity of a Transcription Factor That Controls the Unfolded Protein Response. Cell 87: 391-404. DOI: 10.1016/s0092-8674(00)81360-4

Dimond RL, Loomis WF. 1976. Structure and function of beta-blucosidases in Dictyostelium discoideum. J Biol Chem 251(9): 2680-7. PubMed ID: 57116

Domínguez-Martín E, Ongay-Larios L, Kawasaki L, Vincent O, Coello G, Coria R, Escalante R. 2018. IreA Controls Endoplasmic Reticulum Stress-Induced Autophagy and Survival through Homeostasis Recovery. Molecular and Cellular Biology 38: 10.1128/mcb.00054-18. DOI: 10.1128/MCB.00054-18

Eichinger L, Pachebat JA, Glöckner G, Rajandream MA, Sucgang R, Berriman M, et al., Kuspa. 2005. The genome of the social amoeba Dictyostelium discoideum. Nature 435: 43-57. DOI: 10.1038/nature03481

Free SJ, Schimke RT, Loomis WF. 1976. THE STRUCTURAL GENE FOR α-MANNOSIDASE-1 IN *DICTYOSTELIUM DISCOIDEUM*. Genetics 84: 159-174. DOI: 10.1093/genetics/84.2.159

Gotthardt D, Blancheteau V, Bosserhoff A, Ruppert T, Delorenzi M, Soldati T. 2006. Proteomics Fingerprinting of Phagosome Maturation and Evidence for the Role of a Gα during Uptake. Molecular & Cellular Proteomics 5: 2228-2243. DOI: 10.1074/mcp.M600113-MCP200

Guydosh NR, Kimmig P, Walter P, Green R. 2017. Regulated Ire1-dependent mRNA decay requires no-go mRNA degradation to maintain endoplasmic reticulum homeostasis in S. pombe. eLife 6: 10.7554/elife.29216. DOI: 10.7554/eLife.29216

Hao X, Lüthje F, Rønn R, German NA, Li X, Huang F, et al., Rensing. 2016. A role for copper in protozoan grazing – two billion years selecting for bacterial copper resistance. Molecular Microbiology 102: 628-641. DOI: doi.org/10.1111/mmi.13483

Hollien J, Weissman JS. 2006. Decay of Endoplasmic Reticulum-Localized mRNAs During the Unfolded Protein Response. Science 313: 104-107. DOI: 10.1126/science.1129631

Kimmig P, Diaz M, Zheng J, Williams CC, Lang A, Aragón T, Li H, Walter P. 2012. The unfolded protein response in fission yeast modulates stability of select mRNAs to maintain protein homeostasis. eLife 1: 10.7554/elife.00048. DOI: 10.7554/eLife.00048

Li W, Okreglak V, Peschek J, Kimmig P, Zubradt M, Weissman JS, Walter P. 2018. Engineering ER-stress dependent non-conventional mRNA splicing. eLife 7: 10.7554/elife.35388. DOI: 10.7554/eLife.35388

Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL. 2011. ViennaRNA Package 2.0. Algorithms for Molecular Biology 6: 10.1186/1748-7188-6-26. DOI: 10.1186/1748-7188-6-26

Mathews DH, Disney MD, Childs JL, Schroeder SJ, Zuker M, Turner DH. 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. Proceedings of the National Academy of Sciences 101: 7287-7292. DOI: 10.1073/pnas.0401799101

Maurel M, Chevet E, Tavernier J, Gerlo S. 2014. Getting RIDD of RNA: IRE1 in cell fate regulation. Trends in Biochemical Sciences 39: 245-254. DOI: 10.1016/j.tibs.2014.02.008

McCann H, Meade CD, Williams LD, Petrov AS, Johnson PZ, Simon AE, et al., Petrov. 2025. R2DT: a comprehensive platform for visualizing RNA secondary structure. Nucleic Acids Research 53: 10.1093/nar/gkaf032. DOI: 10.1093/nar/gkaf032

Moore K, Hollien J. 2015. Ire1-mediated decay in mammalian cells relies on mRNA sequence, structure, and translational status. Molecular Biology of the Cell 26: 2873-2884. DOI: 10.1091/mbc.E15-02-0074

Munoz-Ruiz R, Lamrabet O, Jauslin T, Guilhen C, Bourbon A, Cosson P. 2024. Antibacterial effectors in *Dictyostelium discoideum* : specific activity against different bacterial species. mSphere 9: 10.1128/msphere.00471-24. DOI: 10.1128/msphere.00471-24

Niwa M, Patil CK, DeRisi J, Walter P. 2004. Genome-scale approaches for discovering novel nonconventional splicing substrates of the Ire1 nuclease. Genome Biology 6: 10.1186/gb-2004-6-1-r3. DOI: 10.1186/gb-2004-6-1-r3

Oikawa D, Tokuda M, Hosoda A, Iwawaki T. 2010. Identification of a consensus element recognized and cleaved by IRE1α. Nucleic Acids Research 38: 6265-6273. DOI: 10.1093/nar/gkq452

Paris Zk, Fleming IMC, Alfonzo JD. 2012. Determinants of tRNA editing and modification: Avoiding conundrums, affecting function. Seminars in Cell & Developmental Biology 23: 269-274. DOI: 10.1016/j.semcdb.2011.10.009

Schwarz DS, Blower MD. 2015. The endoplasmic reticulum: structure, function and response to cellular signaling. Cellular and Molecular Life Sciences 73: 79-94. DOI: 10.1007/s00018-015-2052-6

Sidrauski C, Cox JS, Walter P. 1996. tRNA Ligase Is Required for Regulated mRNA Splicing in the Unfolded Protein Response. Cell 87: 405-413. DOI: 10.1016/S0092-8674(00)81361-6

Sloma MF, Mathews DH. 2015. Improving RNA Secondary Structure Prediction with Structure Mapping Data. Methods in Enzymology,Computational Methods for Understanding Riboswitches : 91-114. DOI: 10.1016/bs.mie.2014.10.053

Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MsHr, Pihl SI, Tsirigos KD, et al., Nielsen. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nature Biotechnology 40: 1023-1025. DOI: 10.1038/s41587-021-01156-3

Walter P, Ron D. 2011. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334: 1081-1086. DOI: 10.1126/science.1209038

Yuan A, Chia CP. 2000. Role of Esterase gp70 and Its Influence on Growth and Development of Dictyostelium discoideum. Experimental Cell Research 261: 336-347. DOI: 10.1006/excr.2000.5055

**Funding:** Connor Bingham was supported by the Shapiro Undergraduate Research Fellowship (SURF) Program at Randolph-Macon College, Beta Beta Beta (TriBeta) Research Grant, and National Science Foundation (NSF Grant: 2217908). Elias Taylor-Cornejo was supported by the National Science Foundation (NSF Grant: 2217908).

**Conflicts of Interest:** The authors declare that there are no conflicts of interest present.

**Author Contributions:** Connor Bingham: conceptualization, investigation, methodology, validation, data curation, writing - original draft, writing - review editing. Elias Taylor-Cornejo: investigation, conceptualization, formal analysis, data curation, methodology, validation, project administration, supervision, resources, writing - review editing.

**Reviewed By:** Julie Hollien, Anonymous

**History: Received** March 12, 2026 **Revision Received** May 22, 2026 **Accepted** June 17, 2026 **Published Online** June 23, 2026 **Indexed** July 7, 2026

**Copyright:** © 2026 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Citation:** Bingham C, Taylor-Cornejo E. 2026. Bioinformatic pipeline to identify candidate mRNA transcripts targeted by the putative Regulated Ire1-Dependent Decay pathway of *Dictyostelium discoideum*.. microPublication Biology. 10.17912/micropub.biology.002098