1 **Title**

2 A *Drosophila* natural variation screen identifies NKCC1 as a substrate of NGLY1 deglycosylation
3 and a modifier of NGLY1 deficiency

4 **Authors**

5 Dana M. Talsness[1], Katie G. Owings[1], Emily Coelho[1], Gaelle Mercenne[2], John M. Pleinis[2], Aamir
6 R. Zuberi[3], Raghavendran Partha[4], Nathan L. Clark[1], Cathleen M. Lutz[3], Aylin R. Rodan[2,5],
7 Clement Y. Chow[1]*

8 The first two authors contributed equally to this study. [1]Department of Human Genetics, University
9 of Utah School of Medicine, Salt Lake City, UT; [2]Department of Internal Medicine, Division of
10 Nephrology and Hypertension, and Molecular Medicine Program, University of Utah, Salt Lake
11 City, UT; [3]Genetic Resource Science, The Jackson Laboratory, Bar Harbor, ME.; [4]Department of
12 Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA; [5]Medical Service,
13 Veterans Affairs Salt Lake City Health Care System, Salt Lake City, UT; *For correspondence:
14 cchow@genetics.utah.edu

15 **Abstract**

16 N-Glycanase 1 (NGLY1) is a cytoplasmic deglycosylating enzyme. Loss-of-function mutations in
17 the *NGLY1* gene cause NGLY1 deficiency, which is characterized by developmental delay,
18 seizures, and a lack of sweat and tears. To model the phenotypic variability observed among
19 patients, we crossed a *Drosophila* model of NGLY1 deficiency onto a panel of genetically diverse
20 strains. The resulting progeny showed a phenotypic spectrum from 0-100% lethality. Association
21 analysis on the lethality phenotype as well as an evolutionary rate covariation analysis generated
22 lists of modifying genes, providing insight into NGLY1 function and disease. The top association
23 hit was *Ncc69* (human *NKCC1/2*), a conserved ion transporter. Analyses in NGLY1 -/- mouse
24 cells demonstrated that NKCC1 is misglycosylated and has reduced function, making it only the
25 second confirmed NGLY1 enzymatic substrate. The misregulation of this ion transporter may
26 explain the observed defects in secretory epithelium function in NGLY1 deficiency patients.

27 **Introduction**

28 NGLY1 deficiency (OMIM 615273) is an autosomal recessive, rare metabolic disorder caused by
29 loss-of-function mutations in the *NGLY1* gene. Patients with NGLY1 deficiency have a variety of
30 symptoms, including developmental delay, seizures, liver dysfunction, central and peripheral
31 nervous system abnormalities, sweat gland abnormalities, and a lack of tears (alacrima) (Enns et
32 al., 2014; Lam et al., 2017). While the first NGLY1 deficiency patient was only recently identified
33 (Need et al., 2012), there have been rapid research advances thanks to the support of two patient
34 organizations (NGLY1.org and Grace Science Foundation). Even though a great deal has been
35 learned about the genetic disorder in a short amount of time, there are currently no cures or
36 approved treatments for NGLY1 deficiency.

37 The *NGLY1* gene encodes the N-Glycanase protein (NGLY1). NGLY1 is thought to function in
38 the Endoplasmic Reticulum (ER) Associated Degradation (ERAD) pathway due to its association
39 with other ERAD members (Katiyar, Joshi, & Lennarz, 2005; McNeill, Knebel, Arthur, Cuenda, &
40 Cohen, 2004; Park, Suzuki, & Lennarz, 2001). The ERAD pathway retrotranslocates misfolded

proteins from the ER lumen to the cytoplasm where they are degraded by the proteasome (reviewed in Qi, Tsai, & Arvan, 2017). NGLY1 is localized to the cytoplasm where it is thought to remove N-linked glycans from misfolded proteins prior to their degradation (Hirsch, Blom, & Ploegh, 2003). However, it is unclear whether NGLY1 is required to deglycosylate all misfolded proteins, or just a subset, or if it is necessary for protein degradation at all. It has been shown that model substrates can be degraded regardless of glycosylation state (Hirsch et al., 2003; Kario, Tirosh, Ploegh, & Navon, 2008). However, upon loss of NGLY1 function, there is no evidence of ER stress (Owings, Lowry, Bi, Might, & Chow, 2018). ER stress is often observed when there are mutations in proteins that are necessary for ERAD due to the accumulation of misfolded proteins in the ER. It may be that NGLY1 is not necessary for ERAD, or it is involved in a non-canonical ERAD function, or it may be deglycosylating cytoplasmic proteins for an entirely different purpose. These hypothesized functions are not mutually exclusive.

NGLY1 has been shown to deglycosylate various exogenous model substrates such as TCR-α (Hirsch et al., 2003) and RNaseB (Kario et al., 2008). To identify endogenous substrates several mass spectrometry experiments have been performed (Fujihira et al., 2017; Hosomi, Fujita, Tomioka, Kaji, & Suzuki, 2016; Zolekar et al., 2018). Yet the only high-confidence substrate of NGLY1 deglycosylation is NRF1 (gene: *NFE2L1*), which was discovered in a *C. elegans* genetic screen (Lehrbach & Ruvkun, 2016). NRF1 mediates a proteasome "bounce-back" response. NRF1 is constitutively degraded by the proteasome through the ERAD pathway, until the proteasome is inhibited or overwhelmed by protein load. During this proteasome stress, NRF1 accumulates and is deglycosylated by NGLY1 (Tomlin et al., 2017). Rather than targeting the protein for degradation, the deglycosylation activates NRF1 by converting asparagine to aspartic acid residues (Lehrbach, Breen, & Ruvkun, 2019). NRF1 can then be imported into the nucleus to act as a transcription factor for proteasome subunits. The lack of NRF1 activation in NGLY1 deficient patients likely explains some of the disorder's symptoms such as motor dysfunction (Kobayashi et al., 2011) and cognitive deficits (Lee et al., 2011). However not all symptoms can be explained by this one target and therefore there is a pressing need to identify other substrates of NGLY1 deglycosylation.

In addition to discovering new NGLY1 targets, there is a need to understand how background genetic variants affect the number and severity of symptoms in patients. While the majority of patients harbor two complete loss-of-function mutations in *NGLY1* (He et al., 2015), there are many symptoms such as seizures and scoliosis that are only reported in a subset of the patients (Enns et al., 2014). All patients experience developmental delay, but it ranges from slightly below average IQ to completely non-verbal (Lam et al., 2017). To begin to understand how genetic background might affect the range of symptoms and severity of disease, we utilized a collection of genetically diverse *Drosophila* strains known as the *Drosophila* Genetic Reference Panel (DGRP) (Mackay et al., 2012). By crossing a fly model of NGLY1 deficiency onto the panel, we recapitulated the variable phenotype seen in the human population. Here, we report the results of this cross and a list of candidate modifier genes derived from the genome-wide association (GWA) of the cross. To contextualize the candidate modifier list, we also performed an evolutionary rate covariation (ERC) analysis to identify genes that are co-evolving with *NGLY1*. Together these two genetic analyses have generated a list of genes that 1) may explain some of the variation seen between NGLY1 patients, 2) may encode proteins that physically interact with NGLY1 in ERAD or other cellular processes, and 3) may be direct deglycosylation targets of NGLY1. The top candidate modifier gene from the GWA is *NKCC1*, a conserved Na/K/Cl ion co-transporter. We found that *NKCC1* modifies multiple phenotypes in *Drosophila*, and in *NGLY1 -/-* mammalian cells NKCC1 displays abnormal glycosylation and has reduced activity. The misregulation of NKCC1 likely explains several prominent secretory epithelium-related phenotypes observed in NGLY1 deficiency patients.

1 **Results**

2 **Variation in lethality associated with NGLY1 deficiency**

3 We crossed a fly model of NGLY1 deficiency (*Pngl* in flies, hereon referred to as *NGLY1*) onto
4 163 strains of the DGRP in order to assess the effect of natural variation on loss of *NGLY1*
5 function. We previously reported this NGLY1 deficiency model where an *NGLY1* RNAi reduces
6 *NGLY1* transcript by >95% when driven by the ubiquitous *Tubulin*-GAL4 driver transgene
7 (*Tubulin>NGLY1^RNAi^*) (Owings et al., 2018). In order to cross a ubiquitously expressed *NGLY1^RNAi^*
8 onto the DGRP strains in a single cross, we needed to overcome the lethality associated with loss
9 of NGLY1 (Owings et al., 2018). To do this, a *Tubulin*-GAL80 transgene, which represses the
10 effect of GAL4, was crossed onto the *Tubulin>NGLY1^RNAi^* background, such that RNAi is not
11 expressed and flies from this parent strain are healthy and viable (Figure 1A). This donor strain
12 was crossed to each DGRP strain to generate F1 flies that have both ubiquitous knockdown of
13 *NGLY1* and 50% of their genome from each respective DGRP strain (Figure 1B). In this way,
14 analyzing the F1 progeny was a direct measurement of the dominant effect of the DGRP genetic
15 variants on the *NGLY1* knockdown phenotype.
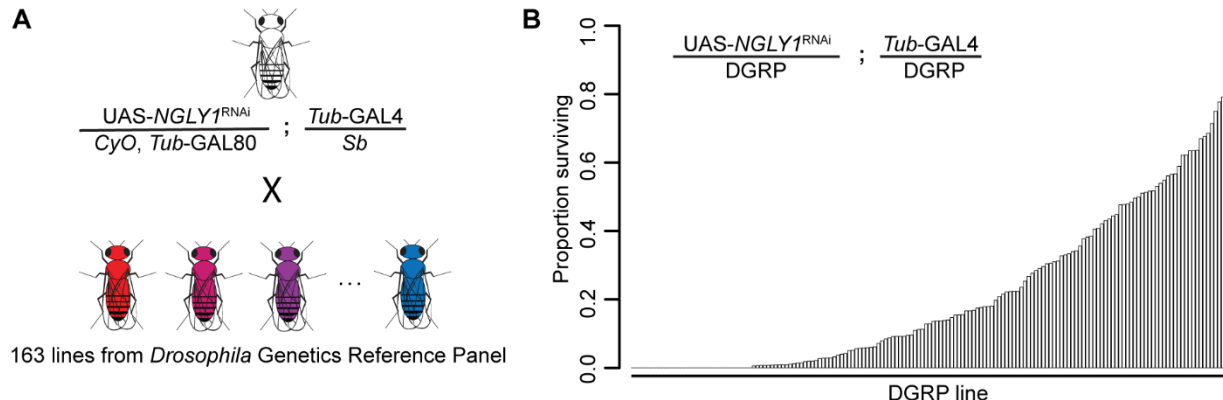
16

17 **Figure 1. Lethality phenotype of *NGLY1* knockdown is highly modifiable by strain background.** (**A**) *Drosophila* cross for *NGLY1*
18 knockdown in each *Drosophila* genetic reference panel (DGRP) strain. (**B**) Proportion of *NGLY1* knockdown flies surviving for each
19 cross was calculated based on the number eclosing compared to the expected number. Expected number was based on the largest
20 control balancer class for each cross.

21 The phenotypic outcome used for this screen was survival through eclosion. We simply scored
22 all adults emerging from each cross in the four balancer categories: *CyO*, *Sb*, double balanced,
23 or no balancers, with the no balancer flies being the *NGLY1* knockdown. If no lethality is present,
24 Mendelian segregation should produce the expected 1:1:1:1 ratio of the genotypes. Given that
25 there is a very low level of lethality associated with each balancer, the largest balancer class is
26 the closest to the expected, and was used to calculate the ratio of lethality for *NGLY1* knockdown.
27 Results of the screen reveal that survival to adulthood was strongly influenced by DGRP genetic
28 background (Figure 1B; Supplemental Table 1), with proportion of surviving flies ranging from 0.0
29 to 0.967. Survival to adulthood was not correlated with efficiency of RNAi, as there was no
30 difference in knockdown efficiency in flies from either end of the phenotypic distribution (low
31 surviving: 92.0% ± 3.7; high surviving: 94.8% ± 3.7; P = 0.4). There was no correlation between
32 proportion of surviving flies and the absolute number of flies in the balancer class (R^2 = 0.02; P =
33 0.14), indicating that the ratio is not driven by the number of the balancer control flies.

34 **Genome-Wide Association**

1 We hypothesized that the observed variable survival to adulthood in *NGLY1* knockdown flies was
2 due to the underlying genetic variation in the DGRP. Therefore, genome-wide association (GWA)
3 analysis of the fully sequenced DGRP was used to identify variants that associated with *NGLY1*
4 knockdown survival. We used a linear mixed model to test 2,007,145 SNPs (Supplemental Table
5 2). We recognize that our study suffers from a multiple testing problem, making it difficult to
6 interpret the role of any single SNP identified. Instead, the location of SNPs was used to identify
7 candidate modifier genes. This type of approach has worked well in the past for other disease
8 models (Ahlers et al., 2019; Chow, Wolfner, & Clark, 2013a, 2013b; Lavoy, Chittoor-Vinod, Chow,
9 & Martin, 2018; Palu et al., 2019) and provides an unbiased list of candidate genes that can be
10 functionally tested for interactions with *NGLY1*.

11 At a nominal P-value of $P < 10^{-5}$, 125 variants are associated with survival to adulthood. Of
12 these 125 variants, 21 fall outside of a gene region (+/- 1 kb from the 5' or 3' UTRs)
13 (Supplemental table 3). The remaining 104 variants map to 61 protein coding candidate genes
14 (Table 1). Eighty-seven of these 104 variants are in noncoding regions (UTRs, introns, or
15 upstream or downstream) and 19 are in coding regions. Of these 19, 12 are synonymous
16 changes and seven are nonsynonymous (*exp, hiw, CG30048, SP2353, CG31690, Hrd3,* and
17 *blue*). When multiple testing correction is applied to all of the variants, the top 12 remain
18 significant. Nine of these SNPs reside in an intron of the *Ncc69* gene. All 9 SNPs are in strong
19 linkage disequilibrium with each other, which is quite unusual for the DGRP. It has been shown
20 that deletions encompassing these nine *Nccc69* intronic SNPs result in a null allele (Leiserson,
21 Forbush, & Keshishian, 2011).

22 **Table 1**. **Candidate modifier genes identified from GWA.** Rank order of candidate genes was established based on
23 the most significant associated SNP in the respective gene.

rank order	gene	FBgn	human ortholog	periphery/ membrane	proteostasis
1	*exp*	FBgn0033668	---	no	no
2	*Ncc69*	FBgn0036279	*NKCC1/2*	yes	no
3	*CG5888*	FBgn0028523	---	yes	no
4	*CG16898*	FBgn0034480	---	no	no
5	*bru3*	FBgn0264001	*CELF2/3/4/5/6*	no	no
6	*CG31690*	FBgn0051690	*TMTC2*	no	yes
7	*CG7227*	FBgn0031970	*SCARB1*	no	no
8	*CR44997*	FBgn0266348	---	no	no
9	*rgn*	FBgn0261258	*Many*	no	no
10	*M6*	FBgn0037092	*GPM6A*	yes	no
11	*Rab26*	FBgn0086913	*RAB26*	yes	yes
12	*Obp56i*	FBgn0043532	---	no	no
13	*5-HT1A*	FBgn0004168	*HTR1A*	yes	no
14	*CG33012*	FBgn0053012	*ERMP1*	no	yes
15	*rst*	FBgn0003285	---	yes	no
16	*CR43926*	FBgn0264547	---	no	no
17	*CG7337*	FBgn0031374	*WDR62*	no	no
18	*hiw*	FBgn0030600	*MYCBP2*	yes	yes
19	*fid*	FBgn0259146	*TRMT9B*	no	no

| 20	*nmo*	FBgn0011817	*NLK*	no	no
21	*Sirup*	FBgn0031971	*SDHAF4*	no	no
22	*tst*	FBgn0039117	*SKIV2L*	no	no
23	*Mdr50*	FBgn0010241	*many*	yes	no
24	*Cpr49Aa*	FBgn0050045	---	no	no
25	*COX7C*	FBgn0040773	*COX7C*	no	no
26	*Eip63E*	FBgn0005640	*CDK14/15*	yes	no
27	*CG30048*	FBgn0050048	*PKD1*	no	no
28	*CG15040*	FBgn0030940	---	no	no
29	*SP2353*	FBgn0034070	*EGFLAM*	no	no
30	*Mf*	FBgn0038294	---	no	no
31	*ome*	FBgn0259175	*many*	no	no
32	*esn*	FBgn0263934	*PRICKLE1-3*	no	no
33	*haf*	FBgn0261509	*many*	no	no
34	*dally*	FBgn0263930	*GPC3/5*	yes	no
35	*robo2*	FBgn0002543	*ROBO1/2/3/4*	no	no
36	*Gyc32E*	FBgn0010197	*NPR1/2*	yes	no
37	*CG8170*	FBgn0033365	*many*	no	no
38	*CG8405*	FBgn0034071	*TMEM259*	no	yes
39	*scaf*	FBgn0033033	---	yes	no
40	*borr*	FBgn0032105	*CDCA8*	yes	no
41	*Syx7*	FBgn0267849	*STX7/12*	yes	no
42	*DIP-delta*	FBgn0085420	*many*	yes	no
43	*cv-c*	FBgn0285955	*DLC1*	yes	no
44	*Snmp2*	FBgn0035815	*CD36/SCARB1*	no	no
45	*Mer*	FBgn0086384	*NF2*	yes	no
46	*sba*	FBgn0016754	---	no	no
47	*Hsromega*	FBgn0001234	---	no	yes
48	*CCAP-R*	FBgn0039396	*NPSR1*	yes	no
49	*Hrd3*	FBgn0028475	*SEL1L*	no	yes
50	*blue*	FBgn0283709	*NEURL4*	no	yes
51	*CG6262*	FBgn0034121	*TREH*	no	no
52	*CG45186*	FBgn0266696	*SVIL*	no	no
53	*Spn*	FBgn0010905	*PPP1R9A*	yes	no
54	*dnc*	FBgn0000479	*PDE4A/B/C/D*	no	no
55	*CG4374*	FBgn0039078	*GFI1B*	no	no
56	*sff*	FBgn0036544	*many*	no	yes
57	*CG42383*	FBgn0259729	*NSFL1C*	no	yes
58	*Dyb*	FBgn0033739	*DTNB*	yes	no
59	*CG34371*	FBgn0085400	---	no	no
60	*CG4341*	FBgn0028481	*TMTC2*	no	yes
61	*CG30043*	FBgn0050043	*ERMP1*	no	yes

1   Gene ontology (GO) enrichment analysis of the 61 candidate genes did not identify enrichment in
2   any biological process or molecular function. However, GO enrichment was identified for the
3   cellular component categories "cell periphery" (GO:0071944; 19/61; q < 0.0016) and "plasma
4   membrane" (GO:0005886; 17/61; q < 0.004). At least 12/61 candidate genes are involved in
5   protein homeostasis: three are involved in ERAD (*CG8405, CG42383*, and *Hrd3*), six are ER
6   resident or membrane proteins (*CG33012, CG30043, CG31690, CG4341, Hrd3*, and *CG8405*),
7   four are involved in ubiquitination or the proteasome (*hiw, blue, CG42383*, and *Hrd3*), one
8   regulates heatshock responses (*Hsromega*), and one regulates N-linked glycosylation (*sff*).

9   Three of the identified ERAD genes already have known interactions with *NGLY1*. *CG8405* is the
10 *Drosophila* ortholog of human *TMEM259*, which physically interacts with NGLY1 in co-
11 immunoprecipitation experiments (Zhu et al., 2017). *CG42383* is the *Drosophila* ortholog of
12 human *NSFL1C* (cofactor p47). NSFL1C and NGLY1 interact with the VCP/P97 AAA-ATPase
13 complex involved in delivering misfolded proteins from the ERAD complex to the proteasome for
14 degradation (Kondo et al., 1997; McNeill et al., 2004). *Hrd3* is the *Drosophila* ortholog of *SEL1L*.
15 SEL1L is a component of the ERAD complex required for retrotranslocation of misfolded proteins
16 from the ER to the cytoplasm for degradation. Recently, the *C. elegans* orthologs of *NGLY1* and
17 *SEL1L* were both identified as modifiers of NRF1 function (Lehrbach & Ruvkun, 2016). These
18 candidate genes are a proof-of-principle that this screen has identified functionally relevant
19 modifiers.

20 The four candidate genes that encode ER resident proteins are particularly interesting. *CG31690*
21 and *CG4341* are both *Drosophila* orthologs of human *TMTC2*, an ER transmembrane protein that
22 regulates calcium homeostasis. *CG33012* and *CG30043* are both *Drosophila* orthologs of human
23 *ERMP1*, an ER metalloprotease. It is striking that in both cases, both *Drosophila* orthologs of a
24 single human gene were identified as candidate modifiers, suggesting that the function of *TMTC2*
25 and *ERMP1* might be particularly important for *NGLY1* lethality. It is not clear how these genes
26 might modify *NGLY1* lethality, but their physical localization to the ER makes sense and suggests
27 a possible role in protein homeostasis as well.

28 **Gene Set Enrichment Analysis**

29 The rank-order candidate modifiers identified in our GWA ignores the majority of the association
30 data by only considering one variant at a time, rather than all of the variants associated with a
31 particular gene. Therefore we performed a gene set enrichment analysis (GSEA), which assigns
32 each variant to the closest gene and generates a per gene metric for P value enrichment (Palu et
33 al., 2019; Subramanian et al., 2005). Given a defined set of genes annotated with a certain GO
34 function, GSEA determines whether the members of that set are randomly distributed throughout
35 the ranked list or if they are found primarily at the top or bottom of that list. We identified 21 gene
36 sets positively associated with the ranked gene list (≥5 genes; > 0.25 enrichment score; P < 0.05)
37 (Figure 2; Supplemental Table 4). These data suggest these GO categories are closely linked to
38 NGLY1 activity and variation in individual genes in these categories contribute to the distribution
39 of lethality observed in our screen.
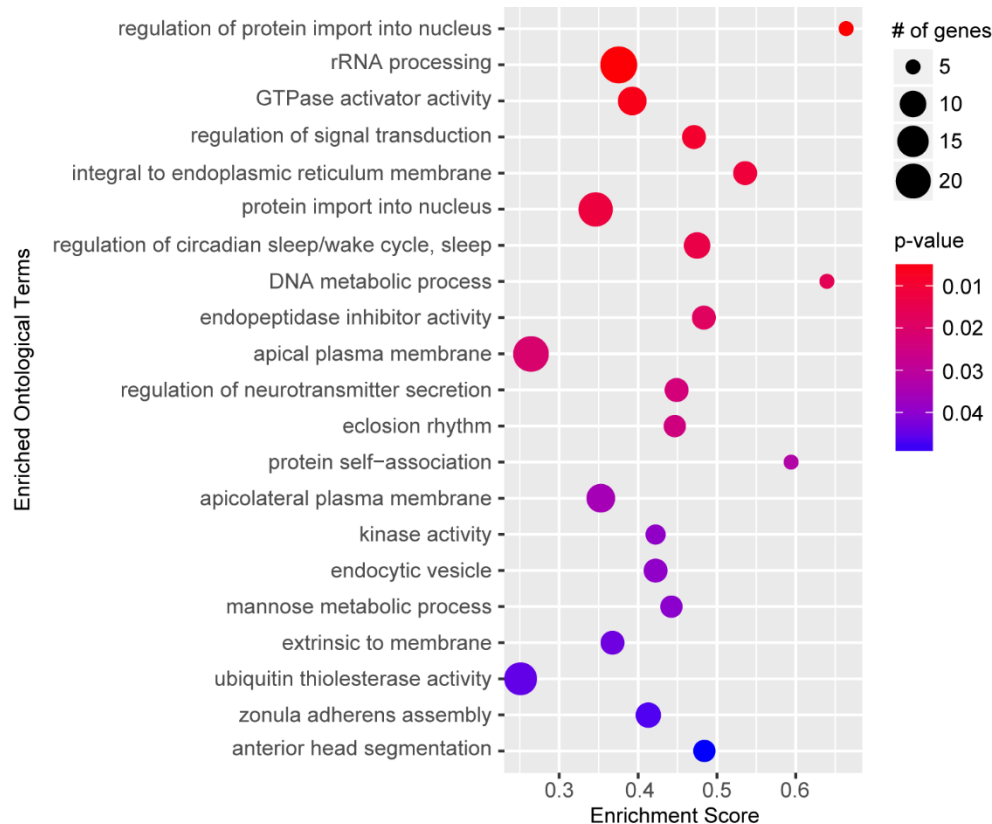
Figure 2: **Gene set enrichment analysis.** Top significant ontological categories identified by GSEA. P-values are indicated by red-to-blue gradient, with red the lowest p-values and blue the highest P-values. Gene number identified in each category is indicated by the size of the circle.

Some of the most significantly enriched categories such as nuclear transport, rRNA processing and signal transduction are broad categories that could have wide reaching implications for NGLY1 function. These processes, however, are difficult to test and require long-term investigation, beyond the scope of this study. Circadian rhythm, on the other hand, is a specific and testable category. The enriched category for circadian rhythm function contains a number of genes that directly modulate circadian rhythm, including, *clock*, *period*, *timeless*, and *cycle*. We hypothesized that if variation in circadian rhythm function modifies lethality associated with loss of NGLY1 function, then NGLY1 must affect the circadian rhythm. To test this, we knocked down *NGLY1* in the LNv pacemaker neurons in the central nervous system using the *Pdf*-GAL4 driver (Renn, Park, Rosbash, Hall, & Taghert, 1999) and assayed rhythmicity of locomotor activity in constant darkness over eight days in *Drosophila* Activity Monitors (DAM). Compared to *Pdf*-GAL4/+ and UAS-*NGLY1^RNAi* controls, flies with *NGLY1* knockdown exhibited a significantly longer period length (Supplemental Figure 1), supporting the idea that NGLY1 function can affect sleep. Indeed it has been reported that patients with NGLY1 deficiency experience disturbed sleep patterns (Enns et al., 2014; Lam et al., 2017).

Evolutionary Rate Covariation

Many of the GWA and GSEA results are intriguing, but appear far removed from the currently known functions of NGLY1. We hypothesized that we could contextualize some of the gene and network results by discovering which of them might be co-evolving with NGLY1. Therefore we employed evolutionary rate covariation (ERC) analysis (Wolfe & Clark, 2015). Gene pairs with high ERC values have correlated rates of substitution and are thought to function together in

1   protein complexes or related pathways. ERC analysis identified hundreds of protein-coding genes
2   with integrated ERC scores exceeding 2 with *NGLY1* (column 'sumnlogpvbest' in Supplemental
3   Table 5). Of the 38 GWA candidates that have human orthologs, two were found in this group
4   with elevated NGLY1 ERC values, *CG4374* (*GFI1B*) and *esn* (*PRICKLE1*). While this overlap is
5   not enriched above background, co-evolution suggests that these two genes might have a
6   particularly important interaction with *NGLY1.*

7   GO analysis was used to determine if there was enrichment in any biological pathways among
8   *NGLY1* co-evolving genes. Among the top enriched pathways were "rRNA/ncRNA/ribosome
9   biogenesis/metabolism-related functions" and "functions related to nuclear pore complex". This is
10 particularly exciting as both processes overlap with the top GO enrichment categories observed
11 in the GSEA analysis, suggesting that the same functional categories that contribute to variation
12 in *NGLY1*-related lethality also appear to contain genes that co-evolve with *NGLY1*. The rRNA
13 processing category (GO:0006364) contained six genes overlapping between the two analyses.
14 This overlap is higher than expected, given two equally sized random groups of genes (GSEA:
15 23 genes; ERC: 37 genes; $P < 2.6 \times 10^{-12}$). Among other ncRNA-related enriched GO categories
16 from the ERC analysis are ncRNA metabolic process (GO:0034660), ncRNA processing
17 (GO:0034470), tRNA metabolic process (GO:0006399), ribosome biogenesis (GO:0042254), and
18 tRNA modification (GO:0006400). The functions related to the nuclear pore included nuclear
19 export (GO:0051168), nuclear pore organization (GO:0006999), nuclear transport (GO:0051169),
20 and nuclear pore complex assembly (GO:0051292). While there was no overlap between ERC
21 and GSEA for exact nuclear pore function categories, GSEA results were enriched for functions
22 related to nuclear import (GO:0042306 and GO:0006606). Together, these observations suggest
23 previously unknown roles for NGLY1 in ncRNA and nuclear pore functions.

24 NGLY1 deficiency is part of a larger category of disorders known as Congenital Disorders of
25 Glycosylation (CDG), with NGLY1 being the only protein that actually deglycosylates substrates.
26 There are 151 known CDG genes. GO analysis of the ERC identified enrichment of the GPI
27 anchor biosynthetic process, which contains several of these CDG genes, leading us to believe
28 that other CDG genes may have been ERC hits. However, the 151 CDG genes do not fall into
29 one functional GO category, therefore, we manually curated the ERC list and identified 26 CDG
30 genes that co-evolve with *NGLY1*. This represents a significant overlap above what is expected
31 by chance ($P < 7.6 \times 10^{-10}$). In particular, five of the 21 genes involved in N-linked glycosylation
32 and nine of the 29 genes involved in GPI-anchor biogenesis are co-evolving with *NGLY1*. The
33 remaining ten genes are spread across the CDG functional spectrum. The identification of a
34 number of CDG genes that co-evolve with *NGLY1*, suggests that NGLY1 function might be
35 important to the broader glycosylation pathways.

36 **Genetic Interaction between *NGLY1* and *Ncc69* in *Drosophila***

37 While these genetic analyses revealed many promising modifying and co-evolving genes which
38 should be investigated, we began by investigating *Ncc69* because it was the top hit with a human
39 ortholog in our GWA analysis. Further, Ncc69 is a glycoprotein, making it a potential target of
40 NGLY1 deglycosylation. *Ncc69* has two mammalian orthologs, *NKCC1* and *NKCC2*. While *Ncc69*
41 is ubiquitously expressed in *Drosophila, NKCC1* (gene: *SLC12A2*) is highly expressed in
42 secretory epithelia and *NKCC2* (gene: *SLC12A1*) is primarily expressed in the kidney. In all cases,
43 the protein is a 12-pass transmembrane cation-chloride co-transporter that brings Na^+, K^+, and
44 Cl^- into the cell. Mutations in *NKCC2* are known to cause type I Bartter's syndrome (Simon &
45 Lifton, 1996) and a recent clinical report shows homozygous loss-of-function mutations in *NKCC1*
46 cause the novel disease Kilquist syndrome (Macnamara et al., 2019). Heterozygous mutations in

1   *NKCC1* have also been shown to cause disease (Delpire et al., 2016), likely due to a dominant
2   negative effect (Koumangoye, Omer, & Delpire, 2018).

3   To confirm the genetic interaction between *NGLY1* and *Ncc69* found in the GWA, we generated
4   ubiquitous double knockdown (DKD) *Drosophila* and scored the viable offspring (Figure 3A). The
5   fraction of knockdown (KD) flies surviving was calculated from the number of KD offspring
6   observed relative to the number of offspring with the balancer phenotype. Single knockdown of
7   either *NGLY1* or *Ncc69* alone did not cause significant lethality, however, the double knockdown
8   resulted in complete lethality. This synthetic lethality confirms *Ncc69* as a hit from the *NGLY1*
9   modifier screen.

10 Knockdown of cation-chloride cotransporters in glia has been shown previously to cause seizures
11 in *Drosophila* (Rusan, Kingsford, & Tanouye, 2014), and we wanted to test whether this
12 phenotype could be modified by *NGLY1* knockdown. We performed single and double
13 knockdowns of *NGLY1* and *Ncc69* in glial cells using the *repo*-GAL4 driver. *Drosophila* were
14 assessed for seizure phenotype using the bang sensitivity assay (Figure 3B). In *NGLY1* KD flies,
15 ~30% showed severe seizures in the form of complete immobility five seconds following vortex.
16 However, by ten seconds following vortex, *NGLY1* KD flies were completely recovered. This is
17 the first report of seizure phenotype in any NGLY1 deficiency model, mimicking what is observed
18 in patients. *Ncc69* KD flies showed severe seizures with 75% seizing at five seconds following
19 vortex, in line with previous reports (Rusan et al., 2014). In the DKD, there was a partial rescue
20 of the severe *Ncc69* phenotype. At all time-points between five and 60 seconds, the DKD flies
21 showed an intermediate phenotype relative to *NGLY1* and *Ncc69* single knockdowns. This
22 genetic effect is in the opposite direction compared to the lethality assay (rescue of phenotype
23 versus enhanced phenotype). While this makes the genetic interaction less straight forward, it is
24 completely possible for two genes to interact differently in different tissue and developmental
25 contexts. Nonetheless, both the lethality and seizure assays confirm a genetic interaction between
26 *NGLY1* and *Ncc69*.

27
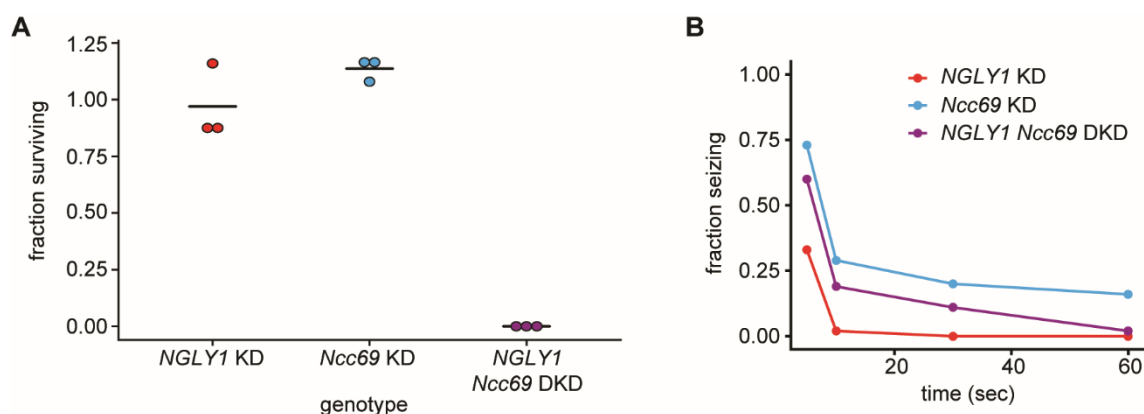
Figure 3. *NGLY1* modifies *Ncc69* seizure phenotype in *Drosophila*. (**A**) Proportion of flies surviving to eclosion in ubiquitous
knockdowns. *NGLY1* knockdown (KD) are UAS-*Pngl*[RNAi]/+; *Tubulin*-GAL4/+. *Ncc69* KD are UAS-*Ncc69*[RNAi]/+; *Tubulin*-GAL4/+. *NGLY1*
Ncc69 double knockdown (DKD) are UAS-*Pngl*[RNAi]/+ UAS-*Ncc69*[RNAi]/*Tubulin*-GAL4/+. Three separate matings were performed for
each cross with at least 100 offspring generated for the balancer control for each. Fraction surviving is calculated compared to balancer
offspring. Chi-square analysis was performed for the total number of flies compared to expected Mendelian numbers. *NGLY1* KD
χ^2=0.626, p=0.43; *Ncc69* KD χ^2=5.891, p=0.02, and *NGLY1 Ncc69* DKD χ^2=483, P < 0.0001. (**B**) Bang sensitivity assay to assess
seizures in glial knockdown flies. *NGLY1* KD are UAS-*Pngl*[RNAi]/+; *repo*-GAL4/+. *Ncc69* KD are UAS-*Ncc69*[RNAi]/+; *repo*-GAL4/+. *NGLY1*
Ncc69 DKD are UAS-*Pngl*[RNAi]/+; UAS-*Ncc69*[RNAi]/ *repo*-GAL4. For each genotype, at least 45 4-7 day old females were used to
calculate the percent seizing at a given time after vortexing. Repeated measures ANOVA p-value=0.000176.

9

Functional analysis of NKCC1 in *NGLY1* null MEFs

To understand the cell biology behind the genetic interaction that was observed in *Drosophila*, we utilized *NGLY1* knockout (-/-) mouse embryonic fibroblasts (MEFs) (jax.org/strain/027060). Fibroblasts should only express the ubiquitous ortholog, NKCC1 (Haas & Forbush, 1998). When the membrane fraction of *NGLY1* -/- MEFs was analyzed by immunoblot for NKCC1 there was a noticeable shift in molecular weight of the band compared to wildtype (+/+), control cells (Figure 4A). Using the molecular weight marker to calculate the size of the proteins, the upper limits of the bands were ~170 kDa for both +/+ and -/- cells (Figure 4B). The lower limit of the bands, however, were ~140 kDa for +/+ cells and ~150 kDa for the -/- cells (Figure 4C). This 10 kDa difference corresponds to the approximate weight of a single glycosylation event (4-20 kDa) (Landolt-Marticorena & Reithmeier, 1994).
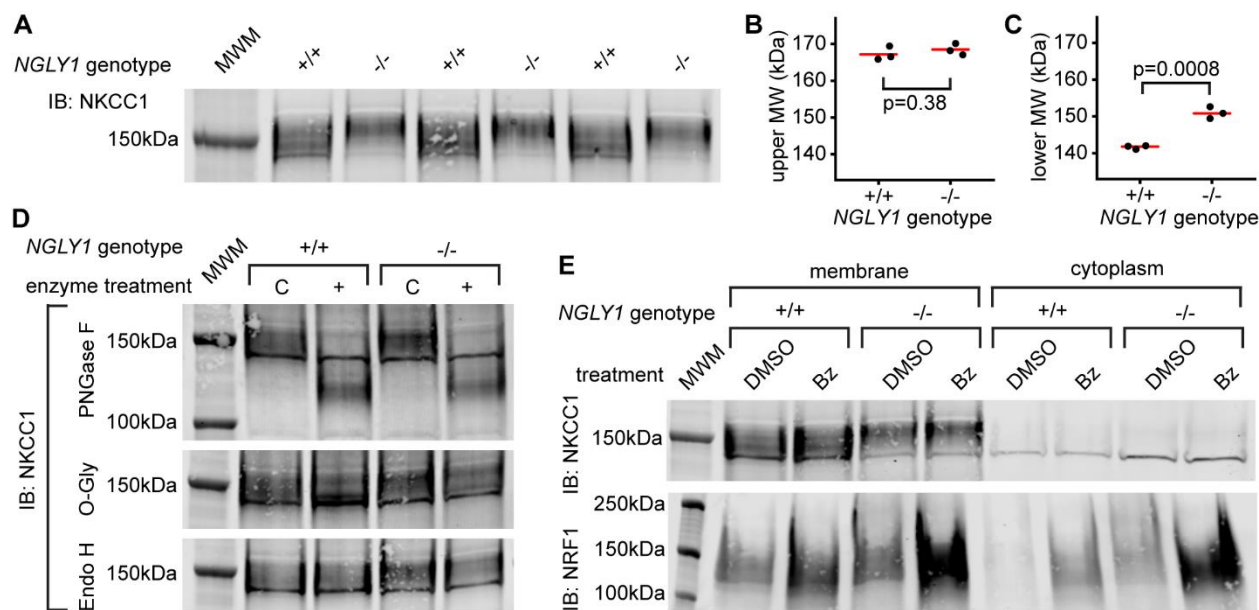

Figure 4. **Endogenous NKCC1 is misglycosylated in NGLY1 deficient MEFs.** (**A**) Control (+/+) and *NGLY1* null (-/-) MEFs were grown to confluency and then lysed to isolate the membrane and cytoplasmic fractions. Three separate membrane lysates for both genotypes were analyzed by immunoblotting for NKCC1 compared to a molecular weight marker (MWM). Blot was used for molecular weight calculations of the upper-most limit (**B**) and the lower-most limit (**C**) of the protein band. Red bar represents the mean. Two-tailed t-test was used to calculate p-values. (**D**) Membrane lysates from MEFs were treated with N-Glycosidase F (PNGase F), O-Glycosidase (O-Gly), or Endoglycosidase H (Endo H) for 1 hour then analyzed by immunoblot. Control (C) samples were treated in all the same conditions but without the added enzyme. (**E**) MEFs were treated with 500nM bortezomib (Bz) or equal volume of vehicle control (DMSO) for 4 hours then lysed to collect membrane and cytoplasmic fractions. Lysates were analyzed by immunoblotting for NKCC1. NRF1 was analyzed as a positive control of proteasome inhibition.

To determine what type of glycosylation event might be responsible for this size difference, cell lysates were treated with deglycosylating enzymes (Figure 4D). PNGase F removes all N-linked glycans and this treatment caused a large decrease in the molecular weight, to ~125 kDa in NKCC1 proteins from both *NGLY1* +/+ and -/- cells. The expected weight of mouse NKCC1 without any post-translational modifications is 130 kDa indicating that all glycosylation sites are likely N-linked. This is in accordance with the prediction of two canonical N-linked glycosylation sites (Payne et al., 1995). The fact that there is no difference in molecular weight between the +/+ and -/- after PNGase treatment indicates the difference observed in the untreated state between these two proteins is likely a single N-linked glycan. Treatment with O-Glycosidase had no effect

10

1   on the molecular weight of the band in either the +/+ or -/- lysates. Although O-Glycosidase does
2   not cleave every type of O-linked glycan, these results coupled with the PNGase results indicate
3   there are likely no O-linked glycans on NKCC1. Finally, to determine the maturation state of the
4   N-linked glycans Endoglycosidase H was used. No change in molecular weight was seen,
5   indicating each N-linked glycan is no longer in the high-mannose state.

6   Currently NRF1 is the only validated substrate of NGLY1 deglycosylation. Under normal
7   conditions NRF1 is degraded by the proteasome, and only when the proteasome is stressed or
8   inhibited does NRF1 become active. Indeed, NGLY1 is thought to act in the ERAD pathway and
9   therefore all of its substrates may be regulated in some way by proteasomal degradation. To test
10   if NKCC1 abundance is affected by the proteasome, MEFs were treated with the proteasome
11   inhibitor bortezomib (Bz). As expected, there was an increase in NRF1 abundance during
12   proteasome inhibition (Figure 4E). However, there was no noticeable increase in NKCC1 protein
13   for either the +/+ or -/- MEFs. These data demonstrate that NKCC1 is being deglycosylated by
14   NGLY1 for purposes other than regulation by the proteasome.

15   Given the altered glycosylation state of NKCC1 observed in *NGLY1* -/- MEFs, we wanted to
16   determine the functionality of NKCC1 in these cells. Previous reports have shown that inhibiting
17   N-linked glycosylation can decrease functionality for both NKCC1 (Singh, Almutairi, Pacheco-
18   Andrade, Almiahuob, & Di Fulvio, 2015) and for NKCC2 (Paredes et al., 2006), therefore we
19   hypothesized that an extra glycan might also decrease functionality. The NKCC proteins and the
20   Na^+/K^+-ATPase can both transport Rb^+ in place of K^+, so we incubated cells with radioactive ^{86}Rb
21   and measured cellular uptake. We confirmed that these two ion transporters accounted for the
22   total ^{86}Rb uptake by assaying uptake in the presence or absence of 10 μM bumetanide, an NKCC
23   inhibitor, or 100 μM ouabain, a Na^+/K^+-ATPase inhibitor, as compared to vehicle control
24   (Supplemental Figure 2). Furthermore, we found that bumetanide-sensitive ^{86}Rb flux, but not
25   ouabain-sensitive flux, was impaired by ~50% in the -/- MEFs, indicating a specific defect in
26   NKCC1 activity without impairment in the Na^+/K^+-ATPase (Figure 5). We assayed ion transport
27   activity under three conditions of isotonic, hypertonic, or hypotonic baths. Although both
28   hypertonic and hypotonic low chloride baths can stimulate NKCC1 activity in other cell types
29   (Darman & Forbush, 2002), we did not see any effect of bathing medium in the MEFs, consistent
30   with a recent report of lack of hypertonic stimulation of NKCC1 in human fibroblasts (Delpire et
31   al., 2016). Together, these data demonstrate that loss of NGLY1 results in a change in the
32   glycosylation state of NKCC1 and a significant reduction in NKCC1 function.

33
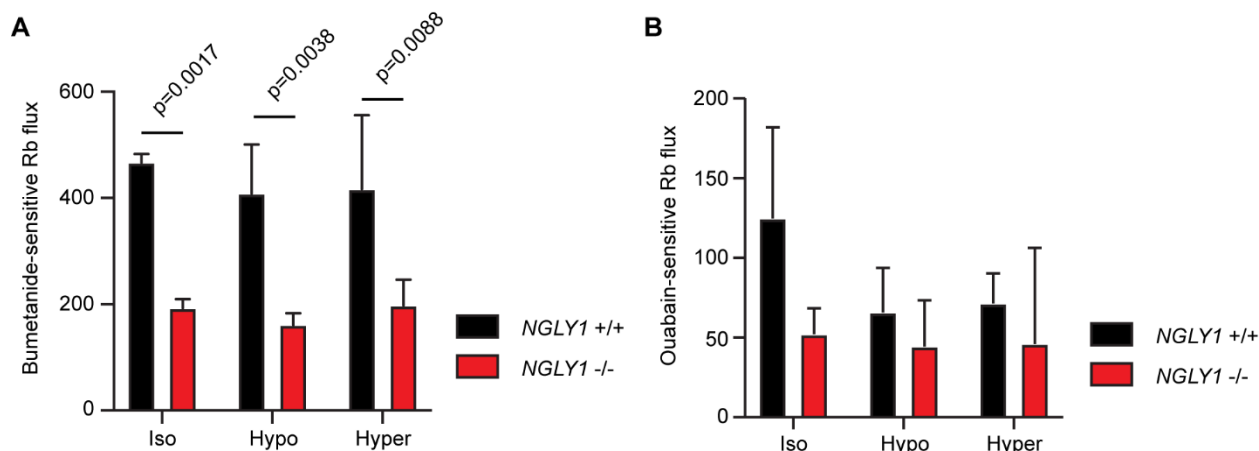
1 **Figure 5**. *NGLY1 -/- MEFs show decreased NKCC1-specific ion flux*. (**A**) Bumetanide-sensitive ^{86}Rb flux was measured in *NGLY1*
2 *+/+* and *NGLY1 -/-* MEFs to measure NKCC1 activity. Flux was examined in three bath conditions, isotonic (iso), hypotonic (hypo),
3 and hypertonic (hyper). There was a significant effect of genotype (P < 0.0001) in two-way ANOVA, with no significant effect of
4 condition (p=0.5756) or interaction (p=0.8075). Adjusted p-values for Sidak's multiple comparisons test between *NGLY1 +/+* and
5 *NGLY1 -/-* are shown in the figure. (**B**) Ouabain-sensitive ^{86}Rb flux was measured in *NGLY1 +/+* and *NGLY1 -/-* MEFs to measure
6 Na$^+$/K$^+$-ATPase activity in the same three conditions as in A. There were no significant effects of genotype (p=0.0516), condition
7 (p=0.3047) or interaction (p=0.4711) by two-way ANOVA, indicating the *NGLY1* knockout has a specific effect on NKCC1 activity
8 without affecting Na$^+$/K$^+$-ATPase activity.

9 ## **Discussion**

10 Like many rare diseases, research into the pathogenesis of NGLY1 deficiency has been narrowly
11 focused, based on early hypotheses. This often limits how we understand the connection between
12 a particular disease and other pathways. For NGLY1 deficiency specifically, basic research and
13 potential therapies have focused intensely on the first and only well-established substrate of
14 NGLY1 deglycosylation, NRF1. Motivated by the extensive phenotypic variation among NGLY1
15 deficiency patients, we took advantage of natural genetic variation in *Drosophila* to identify
16 modifiers of NGLY1 deficiency. This unique screen demonstrated that 1) we can model the
17 extensive phenotypic variation observed and 2) that genetic variation can cause this phenotypic
18 variability. Association analysis then identified a number of exciting candidate modifier genes.
19 Here we have validated the novel and conserved modifier *NKCC1* (*Drosophila Ncc69*), a new
20 potential therapeutic target for NGLY1 deficiency.

21 A major advantage of screens is the identification of previously unanticipated biological
22 connections. Firstly, our association analysis of the lethality screen has generated a list of ~60
23 genes that we hope the scientific and patient communities will be able to use. Secondly, GSEA
24 identified several pathways, including rRNA metabolism and nuclear transport, that are surprising
25 based on known NGLY1 functions. Thirdly, ERC analysis identified genes that are coevolving with
26 NGLY1 across the animal kingdom, including both rRNA/ncRNA pathways and nuclear transport.
27 It appears that rRNA metabolism and nuclear transport are likely important to NGLY1 function,
28 yet it remains unclear how NGLY1 is connected to these pathways. Components of the ribosome
29 and the nuclear pore are often O-glycosylated. While there is no direct connection between
30 NGLY1 and O-glycosylation, we have previously demonstrated that loss of NGLY1 impacts UDP-
31 GlcNAc levels (Owings et al., 2018). It is highly plausible that a misregulation of UDP-GlcNAc
32 levels could affect O-glycosylated proteins. More work is needed to determine exactly how NGLY1
33 is connected to these unexpected pathways. Finally, the ERC analysis also identified 26/151
34 known Congenital Disorders of Glycosylation (CDG) genes. While NGLY1 is also classified as a
35 CDG, it is unclear why there might be co-evolution with other CDG genes. Perhaps there is a
36 feedback mechanism, again, related to UDP-GlcNAc biosynthesis that connects these genes.
37 These results suggest that there is a previously unknown connection between these loosely
38 connected CDG genes.

39 When analyzing the list of modifier genes, it is apparent that many of the candidates are involved
40 in ERAD. This offers a proof-of-principle that this screen is well suited for identifying *bona fide*
41 biologically relevant modifiers. Several previous studies associated NGLY1 with the ERAD
42 process (Bebök, Mazzochi, King, Hong, & Sorscher, 1998; Katiyar et al., 2005; Park et al., 2001).
43 Yet NGLY1 does not appear to be required for proper ERAD function (Hirsch et al., 2003; Misaghi,
44 Pacold, Blom, Ploegh, & Korbel, 2004). While perturbations to ERAD often result in ER stress,
45 we have previously reported that there was no functional or transcriptome evidence for ER stress
46 in a *Drosophila* model of NGLY1 deficiency (Owings et al., 2018). Indeed, in this current screen,

1   we did not identify any genes involved in canonical ER stress responses, suggesting that ER
2   stress is not a major contributor to the pathogenesis of this disease. In fact, the only validated
3   substrate of NGLY1 deglycosylation, NRF1, requires NGLY1 enzymatic activity for regulation by
4   sequence editing when the proteasome is *not* degrading it (Lehrbach et al., 2019). Here we report
5   NKCC1 as a second NGLY1 substrate, and we found that inhibiting the proteasome had no effect
6   on the protein abundance. Thus, it may be that NGLY1 functions to regulate proteins in various
7   ways that are closely related to ERAD, but that do not result in direct proteasomal degradation,
8   and thereby, do not cause an accumulation of misfolded protein and ER stress.

9   As with most hypothesis-free approaches, our analyses produced many exciting new avenues for
10   exploration. In order to keep our work relevant and translatable to the clinic, we focus our follow
11   up experiments on genes with human orthologs. In this report we began with our top hit, *Ncc69*
12   (human *NKCC1/2*), which encodes an SLC12 Na$^+$/K$^+$/2Cl$^-$ transporter. To verify the GWA
13   observation, we analyzed both lethality and seizure phenotypes in *Drosophila*, using ubiquitous
14   or glial cell-specific double knockdown, respectively. These two tests produced signatures of
15   genetic interactions, however in opposite directions (increased lethality but partial rescue of
16   seizures). These results are less straightforward to interpret, but are still completely plausible.
17   Ubiquitous loss of activity is quite different than loss of activity in specific cell types, and possible
18   explanations include different physical interactions in different tissues and non-autonomous
19   effects.

20   This complicated genetic interaction data made it difficult to further dissect the relationship
21   between *NGLY1* and *Ncc69* in *Drosophila*, thus we turned to mammalian cell culture. Mouse
22   embryonic fibroblasts (MEFs) only express the ubiquitous ortholog *NKCC1*, and not *NKCC2*. We
23   hypothesized that, because NGLY1 has deglycosylation activity, it might regulate NKCC1 through
24   glycosylation state. Indeed, we found that NKCC1 protein migrated at a higher size in *NGLY1 -/-*
25   MEFs that approximates one extra glycosylation site, relative to NKCC1 in +/+ MEFs. Treatment
26   with PNGase eliminated this size difference, confirming that this is due to N-linked glycosylation.
27   We found this mis-glycosylation was accompanied by a ~50% reduction in NKCC1 activity. Given
28   that NGLY1 is assumed to function in the ERAD pathway and that the only other confirmed
29   substrate of NGLY1 deglycosylation (NRF1) is known to be degraded through ERAD as part of
30   its regulation, we tested whether there was any change in NKCC1 abundance after proteasome
31   inhibition. Inhibition of the proteasome in *NGLY1* null MEFs resulted in no increase in NKCC1
32   levels in the cytoplasm. Furthermore, the deglycosylation event is present on NKCC1 that is still
33   localized to a membrane and the glycan is no longer in the high mannose state, indicating the
34   protein is not being retrotranslocated out of the ER. Together these data indicate that NGLY1 has
35   deglycosylation activity that is independent of both the retrotranslocation and the degradation
36   components of the ERAD pathway.

37   NKCC1 contains two canonical N-linked glycosylation sites (Payne et al., 1995) similar to those
38   validated in NKCC2 (Paredes et al., 2006). As expected, both sites are in an extracellular loop,
39   and therefore these sites face the lumen of the ER during protein translation and maturation.
40   Because of this arrangement, these two sites are not predicted to be accessible by NGLY1, which
41   is localized to the cytoplasm. Through amino acid analysis, we have identified one other
42   asparagine residue within the necessary N-X-S/T sequence for N-linked glycosylation (N168), but
43   it has not been considered based on its location in the amino-terminal cytoplasmic tail of the
44   protein. The recent Cryo-EM structure determined that the amino-terminal tail is disordered and

that the carboxy-terminal tail acts as a regulatory domain (Chew et al., 2019). Given our data, we hypothesize that N168 may be glycosylated by an unknown cytoplasmic enzyme and then deglycosylated by NGLY1 as some sort of regulation. Although rare, there have been reports of cytosolic N-linked glycosylation, including on the dog kidney Na+, K+-ATPase pump (reviewed in Hart & West, 2017). In all of these cases, and in our case here, the direct structural data has not yet been determined. Validating this putative glycosylation site and determining the method of glycosylation are a priority for future studies.

While the possibility of NKCC1 being regulated by a deglycosylation event is exciting, we cannot eliminate the possibility that this NKCC1 regulation is secondary. NGLY1 may be regulating an intermediary protein that in turn, regulates NKCC1. This was recently found to be the case for aquaporins in NGLY1 deficient cells (Tambe, Ng, & Freeze, 2019). NGLY1 was found to regulate the abundance of transcription factors Atf1/Creb1 which then regulate the transcription of several aquaporin subunits. However, the regulation of Atf1/Creb1 by NGLY1 was independent of its enzymatic activity. This seems to be in contrast to NKCC1, which directly shows a molecular weight shift in the absence of NGLY1, suggesting a more direct interaction with the deglycosylating enzyme. Nevertheless it has been shown that NGLY1 can regulate substrates in two distinctly different manners: regulating abundance through transcription factor intermediaries (Tambe et al., 2019), and regulating activity through deglycosylation and conversion of asparagine residues to aspartic acids (Lehrbach et al., 2019). While the exact method of NGLY1 regulating NKCC1 activity remains to be determined, we hypothesize that the deglycosylation activity is required due to the change in the molecular weight of NKCC1 when NGLY1 is absent.

Identification of targets and modifier genes should provide insight into the pathogenesis of a disease and help explain some of the patient phenotypes. When NRF1 was identified as the first target of NGLY1, it provided insight into some of the molecular defects observed in NGLY1 deficient cells, including deficits in proteasomal function and expression. These cellular phenotypes, however, did not translate well into insight into the complex patient symptoms. In contrast, decreased NKCC1 activity may explain some of the prominent features of NGLY1 deficiency. NKCC1 functions in many secretory epithelia, such as the salivary, sweat and lacrimal glands, to allow basolateral ion uptake and subsequent secretion (Delpire & Gagnon, 2018). Therefore, a decrease in NKCC1 activity could well explain the alacrima and reduced saliva and sweat production seen in NGLY1 deficiency. Strikingly, a recent clinical report describes a patient with a homozygous deletion in *NKCC1* (null) who has many overlapping features with NGLY1 deficiency patients, including absence of saliva, tears, and sweat (Kilquist syndrome) (Macnamara et al., 2019). The *NKCC1* null mouse also displays defects in salivation (Evans et al., 2000). Other notable, but perhaps less specific, features in the NKCC1 deficient child, including developmental delay and gastrointestinal problems, also overlap with those observed in NGLY1 deficiency. The NKCC1 deficient patient also had severe hearing loss, cochlear defects, and abnormal auditory brainstem responses (ABRs), yet NGLY1 deficiency patients only have abnormal ABRs. This difference may be explained by a 50% reduction in NKCC1 activity, rather than complete loss of activity. Nevertheless, the overlap in a majority of the symptoms between this new syndrome and NGLY1 deficiency strengthen the case for both a genetic and functional connection between NGLY1 and NKCC1.

NKCC1 may be a promising target for the development of NGLY1 deficiency therapies. Given that it is a transporter and partially exposed to the extracellular space, NKCC1 could be

1   particularly amenable to modulation by small molecules. Based on our work reported here, we
2   predict that increasing function of NKCC1 may ameliorate some symptoms. Quercetin, a
3   flavonoid, is a readily available molecule that has been shown to enhance NKCC1 activity (Asano
4   et al., 2009; Nakajima, Niisato, & Marunaka, 2011). These studies demonstrated that quercetin
5   significantly increased ^{86}Rb uptake in cell culture and that this increase was bumetanide-sensitive,
6   indicating specificity to NKCC1. This is similar to other studies showing flavonoids increasing the
7   activity of channels, such as the flavonoid genistein increasing activity of the cystic fibrosis
8   transmembrane conductance regulator (CFTR) (Sugawara & Nikaido, 2014), and several
9   flavonoids targeting cardiovascular channels (Scholz, Zitron, Katus, & Karle, 2010). These
10   quercetin studies, however, are in the context of normal functioning NKCC1 protein. In our study,
11   NKCC1 is misglycosylated and it is unknown whether quercetin would be able to modulate
12   misglycosylated NKCC1 activity. Nonetheless, quercetin was recently discovered to provide
13   benefit to NGLY1 deficient *C. elegans* in a screen of potential therapeutics (Iyer et al., 2019).
14   These results coupled with our discovery of NKCC1 as a NGLY1 substrate offer an exciting new
15   avenue of treatment for NGLY1 deficiency patients. Targeted studies are needed to determine if
16   quercetin or other molecules could specifically enhance NKCC1 function in the context of NGLY1
17   deficiency.

18   In this study we took a series of unbiased approaches in *Drosophila* to identify modifiers of NGLY1
19   deficiency. This resulted in a number of new insights into the potential pathogenesis of NGLY1
20   that we hope others will also investigate. With rare diseases like NGLY1 deficiency, unbiased and
21   forward genetic approaches are an efficient method for expanding possible avenues of
22   investigation and therapeutic development. This study also highlights the power of using model
23   organisms like *Drosophila* to uncover pathways and genes that can be validated in mammalian
24   systems and targeted for therapeutic development.

25   **Materials and Methods**
26
27   ***Drosophila* lines**
28   Flies were maintained at 25°C on a 12-hour light/dark cycle and raised on a standard diet based
29   on the Bloomington Stock Center standard medium with malt. All flies were aged 3-5 days old
30   for experiments. For the DGRP screen, the following *D. melanogaster* stocks were used:
31   *Pngl*^{RNAi} (Bloomington *Drosophila* Stock Center: 54853) and *Tubulin*-GAL4 driver (5138). The
32   *Tubulin*-GAL80 strain was provided by Dr. Carl Thummel (University of Utah). The DGRP
33   strains are available at the Bloomington *Drosophila* Stock Center. To measure circadian rhythm,
34   the following stocks were used: a *w*- Berlin control strain, a *w*; *Pdf*-GAL4 strain (outcrossed to
35   w- Berlin), and a *yv*; UAS-*Pngl*^{RNAi} strain (Bloomington stock center #42592). For bang
36   sensitivity assays, the following stocks were used: UAS-*Pngl*^{RNAi} (BL #54853), UAS-*Ncc69*^{RNAi}
37   (BL #28682), *repo*-GAL4 were obtained from the Bloomington Stock Center and Adrian
38   Rothenfluh (University of Utah) respectively.
39
40   **DGRP screen**
41   Virgin females from the DGRP strains were fed yeast overnight and then crossed with males
42   from the donor strain UAS-*Pngl*^{RNAi}/*Cyo*,*Tubulin*-GAL80; *Tubulin*-GAL4/*Sb* in two replicate
43   bottles. Progeny were collected and scored for the four balancer classes: *CyO*, *Sb*, double
44   balanced, or no balancers, with the no balancer flies being the *NGLY1* knockdown. This cross
45   should produce the expected 1:1:1:1 ratio of the four genotypes. Given that there is always a

1 very low level of lethality associated with each balancer, the largest balancer class was
2 considered the closest to the expected number. We scored at least 200 flies per DGRP cross.
3 Males and females were combined for a single count. To calculate the proportion of *NGLY1*
4 knockdown flies by generating a ratio of *NGLY1* knockdown/ largest balancer class. This metric
5 was used for the GWA.
6
7 **Genome Wide Association**
8 GWA was performed as previously described (Chow, Kelsey, Wolfner, & Clark, 2016). DGRP
9 genotypes were downloaded from the website, http://dgrp.gnets.ncsu.edu/. Variants were filtered
10 for minor allele frequency (≥ 0.05), and non-biallelic sites were removed. A total of 2,007,145
11 variants were included in the analysis. The proportion of *NGLY1* knockdown flies surviving was
12 regressed on each SNP. To account for cryptic relatedness (B. Z. He et al., 2014; Huang et al.,
13 2014), GEMMA (v. 0.94) (Zhou & Stephens, 2012) was used to both estimate a centered genetic
14 relatedness matrix and perform association tests using the following linear mixed model (LMM):
15

$$y = \alpha + x\beta + u + \epsilon$$
$$u \sim MVN\_n\,(0, \lambda\tau^{\wedge}(-1)\,K)$$
$$\epsilon \sim MVN\_n\,(0, \tau^{\wedge}(-1)\,I\_n\,)$$

19
20 where, as described and adapted from (Zhou & Stephens, 2012), y is the n-vector of proportion
21 lethality for the n lines, α is the intercept, x is the n-vector of marker genotypes, β is the effect size
22 of the marker. u is a n x n matrix of random effects with a multivariate normal distribution (MVN_n)
23 that depends on λ, the ratio between the two variance components, $\tau^{\wedge}(-1)$, the variance of
24 residuals errors, and where the covariance matrix is informed by K, the calculated n x n marker-
25 based relatedness matrix. K accounts for all pairwise non-random sharing of genetic material
26 among lines. ϵ, is a n-vector of residual errors, with a multivariate normal distribution that depends
27 on $\tau^{\wedge}(-1)$ and I_n, the identity matrix. Genes were identified from SNP coordinates using the BDGP
28 R54/dm3 genome build. A SNP was assigned to a gene if it was +/- 1 kb from a gene body.
29
30
31 **Gene Set Enrichment Analysis**
32 GSEA was run to generate a rank-list of genes based on their enrichment for significantly
33 associated polymorphisms as previously described (Palu et al., 2019). Polymorphisms within 1kb
34 of more than 1 gene were assigned to one gene based on a priority list of exon, UTR, intron, and
35 upstream or downstream. Genes were assigned to GO categories, and calculation of enrichment
36 score was performed as described (Subramanian et al. 2005). Only gene sets with ≥ 5 genes,
37 >0.25 enrichment score, and a $P < 0.05$ were considered.
38
39 **Evolutionary Rate Covariation**
40 ERC is a method to examine the similarity of evolutionary histories of pairs of genes (Clark, Alani,
41 & Aquadro, 2012). The method examines the variation over time of a gene's rate of sequence
42 evolution. Using estimates of evolutionary rate over the branches of a gene's phylogenetic tree,
43 the method measures the correlation between genes of these branch-specific rates. Genes within
44 correlated rate variation tend to be functionally related and have been used to discover new genes
45 within pathways and diseases (Brunette, Jamalruddin, Baldock, Clark, & Bernstein, 2019;
46 Priedigkeit, Wolfe, & Clark, 2015; Raza et al., 2019)
47
48 ERC values in this study were taken from a compilation of ERC correlations calculated separately
49 for 3 taxonomic groups: 62 mammals, 39 non-mammalian vertebrates, and 22 *Drosophila*

1 species. Mammal and non-mammalian vertebrate alignments were taken from the multiz
2 alignment available from the UCSC Genome Browser (Haeussler et al., 2019). For each
3 alignment, we filtered out low quality orthologs containing fewer than 50 non-gap amino acid sites
4 or less than 70% non-gap sites and removed alignments with fewer than 15 species. Alignments
5 were made for the *Drosophila* species after downloading protein-coding genome sequences from
6 FlyBase and NCBI. Orthologous groups were identified using Orthofinder and alignments made
7 with PRANK (Emms & Kelly, 2015; Löytynoja & Goldman, 2008). For each amino acid alignment,
8 we estimated branch lengths using aaml in the phylogenetic analysis using maximum likelihood
9 (PAML) package (Yang, 2007). ERC values (correlation coefficients) for all genes with *NGLY1*
10 were calculated using the RERconverge package (Kowalczyk et al., 2019). We report the ERC
11 results for the mammalian group as the negative log of their p-values for each gene pair
12 (Supplemental Table S5 'nlogpvbest1'). Each gene pair also incorporated results from the
13 vertebrate and *Drosophila* datasets by summing their negative log p-values, when orthologs were
14 present for their respective datasets (Supp. Table S5 'sumnlogpvbest'). The resulting
15 taxonomically integrated results of ERC with *NGLY1* were sorted and used for gene set
16 enrichment analysis (GSEA).
17

18 ### *Drosophila* circadian rhythm assay
19 Male flies with the following genotypes were used in circadian rhythm assays: *w/Y;Pdf*-GAL4/+,
20 *yv/Y;* UAS-*Pngl*^{RNAi}/+, and *yv/Y; Pdf*-GAL4/UAS-*Pngl*^{RNAi}/+. 2-5 day old flies were entrained for
21 at least 3 days to a 12 hr light: 12 hr dark regimen (LD) within a *Drosophila* Activity Monitor
22 (DAM; TriKinetics, Waltham, MA) filled with standard fly food. After entrainment, flies were
23 monitored in complete darkness (DD) for 8 days. The data was collected in 30 minute bins, and
24 analyzed for period length using ClockLab, Version 6. Graphs were generated and one-way
25 ANOVA performed, with Tukey's multiple comparison of all three genotypes, using GraphPad
26 Prism, Version 8.
27

28 ### *Drosophila* seizure assay
29 The Bang Sensitivity Assay (BSA) was performed on the following genotypes: UAS-*Pngl*^{RNAi};
30 *repo*-GAL4, UAS-*Ncc69*^{RNAi}; *repo*-GAL4, and UAS-*Pngl*^{RNAi}/+; UAS-*Ncc69*^{RNAi}/*repo*-GAL4.
31 Females 4-7 days old were assayed. Flies were not exposed to CO_2 for 3 days prior to BSA
32 testing. Flies were flipped into empty vials and allowed to rest for 2 hours. They were then
33 vortexed on a Thermo Scientific LP Vortex Mixer for 10 seconds at maximum speed. The vortexed
34 flies were filmed for 60 seconds. The video was used to score seizures at 5, 10, 30, and 60
35 seconds.
36

37 ### Mammalian cell culture and proteasome inhibition
38 Mouse embryonic fibroblasts (MEFs) were generated by Jackson Laboratory (Bar Harbor, Maine)
39 from *NGLY1* knockout mice and littermate controls (C57BL/6J-*Ngly1*^{em4Lutzy}/J, #027060). MEFs
40 were immortalized in the laboratory of Dr. Hamed Jafar-Nejad (Baylor College of Medicine) and
41 then gifted to us. MEFs were grown in DMEM (Gibco 11965) supplemented with 10% fetal bovine
42 serum (FBS) and penicillin/streptomycin in 5% CO_2 at 37°C. For proteasome inhibition, MEFs
43 were incubated with 500nM bortezomib (EMD Millipore) or an equivalent volume of DMSO as a
44 vehicle control, for 4 hours under standard conditions.
45

46 ### Western blotting

1   MEFs were grown to 80-90% confluency then collected. Cell pellets were weighed and then
2   resuspended in a proportional volume of phosphate buffered saline (PBS). Equivalent volumes of
3   resuspension were always used for each lysis. Cells were lysed using a cell fractionation kit (Cell
4   Signaling Technologies, #9038) with each buffer supplemented with 1mM PMSF and 1x protease
5   inhibitor cocktail (Cell Signaling Technologies, #5871).
6
7   Lysates were separated by SDS-PAGE on 3-8% Tris-acetate gels (BioRad #3450129) for 2.5
8   hours at 150V, then transferred to PVDF membrane by wet transfer at 50V for 1 hour. Membranes
9   were blocked in either 5% milk or 5% BSA according to the recommendations of the primary
10 antibody manufacturer. Primary antibodies were as follows: anti-NKCC1 (Cell Signaling
11 Technologies #14581), anti-TCF11/NRF1 (Cell Signaling Technologies #8052). Membranes were
12 incubated in primary antibody at 1:1000 in blocking buffer overnight. IRDye secondary antibody
13 (Abcam #216773) was used for infrared detection at 1:10,000 dilution in blocking buffer for 1 hour.
14 Membranes were scanned on an Odyssey CLx (Li-cor) and analyzed with the accompanying
15 software, Image Studio™.
16
17 **Deglycosylation reactions**
18 MEFs were lysed in the same manner as described for Western blotting. The membrane fraction
19 was then incubated with one of the three deglycosylation enzymes: O-Glycosidase (New England
20 Biolabs, #P0733), PNGase F (New England Biolabs, #P0704), Endoglycosidase F (New England
21 Biolabs, #P0702) according to the manufacturer's directions. Reactions were incubated at 37°C
22 for 1 hour. Controls were treated with all the same buffers and reaction conditions but without the
23 added enzyme.
24
25 **Rb$^+$ flux assay**
26 20,000 cells/well of immortalized mouse embryonic fibroblasts from control or *Ngly1 -/-* mice were
27 seeded into a 96-well plate. The following day, media was removed and the cells were washed
28 with 1x PBS. 100 μl of pre-incubation medium (in mM, 135 Na gluconate, 5 K gluconate, 1 Ca
29 gluconate, 1 Mg gluconate, 15 HEPES pH 7.4, 5 glucose) was added to each well and the cells
30 were incubated for 30 minutes at 37 °C. Next, 100 μl of pre-incubation medium containing either
31 DMSO, bumetanide or ouabain was added to each well to achieve final concentrations of 0.1%
32 (DMSO), 10 μM (bumetanide) or 0.1 mM (ouabain) and incubated for 30 minutes at room
33 temperature. Then, 1 ml of medium containing DMSO (0.1%), bumetanide (10 μM) or ouabain
34 (0.1 mM), as well as ^{86}Rb (10 mCi/μl), was added to each well. Three different media were used.
35 Isotonic media contained (in mM): 140 NaCl, 5 KCl, 2 CaCl$_2$, 1 MgCl$_2$, 5 glucose, 15 HEPES pH
36 7.4. Hypertonic medium was the same as isotonic medium, with the addition of 75 mM sucrose.
37 For hypotonic medium, isotonic medium was diluted 1:2 in water. The cells were incubated for 7
38 minutes at room temperature. Medium was removed and the cells were washed 3 times with ice-
39 cold 1x PBS. Cells were lysed in 100 μl 2% SDS and incubated for 15 minutes at room
40 temperature. Radioactivity was measured in a liquid scintillation counter.
41
42 **Acknowledgements**

Competing Interests

None.

References

Ahlers, L. R. H., Trammell, C. E., Carrell, G. F., Mackinnon, S., Torrevillas, B. K., Chow, C. Y., … Goodman, A. G. (2019). Insulin Potentiates JAK/STAT Signaling to Broadly Inhibit Flavivirus Replication in Insect Vectors. *Cell Reports*, *29*(7), 1946-1960.e5. https://doi.org/10.1016/j.celrep.2019.10.029

Asano, J., Niisato, N., Nakajima, K., Miyazaki, H., Yasuda, M., Iwasaki, Y., … Marunaka, Y. (2009). Quercetin stimulates Na+/K+/2Cl- cotransport via PTK-dependent mechanisms in human airway epithelium. *American Journal of Respiratory Cell and Molecular Biology*, *41*(6), 688–695. https://doi.org/10.1165/rcmb.2008-0338OC

Bebök, Z., Mazzochi, C., King, S. A., Hong, J. S., & Sorscher, E. J. (1998). The mechanism underlying cystic fibrosis transmembrane conductance regulator transport from the endoplasmic reticulum to the proteasome includes Sec61beta and a cytosolic, deglycosylated intermediary. *The Journal of Biological Chemistry*, *273*(45), 29873–29878. https://doi.org/10.1074/jbc.273.45.29873

Brunette, G. J., Jamalruddin, M. A., Baldock, R. A., Clark, N. L., & Bernstein, K. A. (2019). Evolution-based screening enables genome-wide prioritization and discovery of DNA repair genes. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(39), 19593–19599. https://doi.org/10.1073/pnas.1906559116

Chew, T. A., Orlando, B. J., Zhang, J., Latorraca, N. R., Wang, A., Hollingsworth, S. A., … Feng, L. (2019). Structure and mechanism of the cation-chloride cotransporter NKCC1. *Nature*, *572*(7770), 488–492. https://doi.org/10.1038/s41586-019-1438-2

Chow, C. Y., Kelsey, K. J. P., Wolfner, M. F., & Clark, A. G. (2016). Candidate genetic modifiers of retinitis pigmentosa identified by exploiting natural variation in Drosophila. *Human Molecular Genetics*, *25*(4), 651–659. https://doi.org/10.1093/hmg/ddv502

Chow, C. Y., Wolfner, M. F., & Clark, A. G. (2013a). Large neurological component to genetic differences underlying biased sperm use in Drosophila. *Genetics*, *193*(1), 177–185. https://doi.org/10.1534/genetics.112.146357

Chow, C. Y., Wolfner, M. F., & Clark, A. G. (2013b). Using natural variation in Drosophila to discover previously unknown endoplasmic reticulum stress genes. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(22), 9013–9018. https://doi.org/10.1073/pnas.1307125110

Clark, N. L., Alani, E., & Aquadro, C. F. (2012). Evolutionary rate covariation reveals shared

functionality and coexpression of genes. *Genome Research*, *22*(4), 714–720. https://doi.org/10.1101/gr.132647.111

Darman, R. B., & Forbush, B. (2002). A regulatory locus of phosphorylation in the N terminus of the Na-K-Cl cotransporter, NKCC1. *The Journal of Biological Chemistry*, *277*(40), 37542–37550. https://doi.org/10.1074/jbc.M206293200

Delpire, E., & Gagnon, K. B. (2018). Na+ -K+ -2Cl- Cotransporter (NKCC) Physiological Function in Nonpolarized Cells and Transporting Epithelia. *Comprehensive Physiology*, *8*(2), 871–901. https://doi.org/10.1002/cphy.c170018

Delpire, E., Wolfe, L., Flores, B., Koumangoye, R., Schornak, C. C., Omer, S., … Adams, D. R. (2016). A patient with multisystem dysfunction carries a truncation mutation in human SLC12A2, the gene encoding the Na-K-2Cl cotransporter, NKCC1. *Cold Spring Harbor Molecular Case Studies*, *2*(6), a001289. https://doi.org/10.1101/mcs.a001289

Emms, D. M., & Kelly, S. (2015). OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*, 157. https://doi.org/10.1186/s13059-015-0721-2

Enns, G. M., Shashi, V., Bainbridge, M., Gambello, M. J., Zahir, F. R., Bast, T., … Goldstein, D. B. (2014). Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *16*(10), 751–758. https://doi.org/10.1038/gim.2014.22

Evans, R. L., Park, K., Turner, R. J., Watson, G. E., Nguyen, H. Van, Dennett, M. R., … Melvin, J. E. (2000). Severe impairment of salivation in Na+/K+/2Cl- cotransporter (NKCC1)-deficient mice. *The Journal of Biological Chemistry*, *275*(35), 26720–26726. https://doi.org/10.1074/jbc.M003753200

Fujihira, H., Masahara-Negishi, Y., Tamura, M., Huang, C., Harada, Y., Wakana, S., … Suzuki, T. (2017). Lethality of mice bearing a knockout of the Ngly1-gene is partially rescued by the additional deletion of the Engase gene. *PLoS Genetics*, *13*(4), e1006696. https://doi.org/10.1371/journal.pgen.1006696

Haas, M., & Forbush, B. (1998). The Na-K-Cl cotransporters. *Journal of Bioenergetics and Biomembranes*, *30*(2), 161–172. https://doi.org/10.1023/a:1020521308985

Haeussler, M., Zweig, A. S., Tyner, C., Speir, M. L., Rosenbloom, K. R., Raney, B. J., … Kent, W. J. (2019). The UCSC Genome Browser database: 2019 update. *Nucleic Acids Research*, *47*(D1), D853–D858. https://doi.org/10.1093/nar/gky1095

Hart, G., & West, C. (2017). *Essentials of Glycobiology.* (A. Varki, R. Cummings, J. Esko, P. Stanley, G. Hart, M. Aebi, … P. Seberger, Eds.) (3rd Editio). Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press. Retrieved from https://www.ncbi.nlm.nih.gov/books/NBK310274/

He, B. Z., Ludwig, M. Z., Dickerson, D. A., Barse, L., Arun, B., Vilhjálmsson, B. J., … Kreitman, M. (2014). Effect of genetic variation in a Drosophila model of diabetes-associated misfolded human proinsulin. *Genetics*, *196*(2), 557–567. https://doi.org/10.1534/genetics.113.157800

He, P., Grotzke, J. E., Ng, B. G., Gunel, M., Jafar-Nejad, H., Cresswell, P., … Freeze, H. H. (2015). A congenital disorder of deglycosylation: Biochemical characterization of N-glycanase 1 deficiency in patient fibroblasts. *Glycobiology*, *25*(8), 836–844. https://doi.org/10.1093/glycob/cwv024

Hirsch, C., Blom, D., & Ploegh, H. L. (2003). A role for N-glycanase in the cytosolic turnover of glycoproteins. *The EMBO Journal*, *22*(5), 1036–1046. https://doi.org/10.1093/emboj/cdg107

Hosomi, A., Fujita, M., Tomioka, A., Kaji, H., & Suzuki, T. (2016). Identification of PNGase-dependent ERAD substrates in Saccharomyces cerevisiae. *The Biochemical Journal*, *473*(19), 3001–3012. https://doi.org/10.1042/BCJ20160453

Huang, W., Massouras, A., Inoue, Y., Peiffer, J., Ràmia, M., Tarone, A. M., … Mackay, T. F. C. (2014). Natural variation in genome architecture among 205 Drosophila melanogaster Genetic Reference Panel lines. *Genome Research*, *24*(7), 1193–1208. https://doi.org/10.1101/gr.171546.113

Iyer, S., Mast, J. D., Tsang, H., Rodriguez, T. P., DiPrimio, N., Prangley, M., … Perlstein, E. O. (2019). Drug screens of NGLY1 deficiency in worm and fly models reveal catecholamine, NRF2 and anti-inflammatory-pathway activation as potential clinical approaches. *Disease Models & Mechanisms*, *12*(11). https://doi.org/10.1242/dmm.040576

Kario, E., Tirosh, B., Ploegh, H. L., & Navon, A. (2008). N-linked glycosylation does not impair proteasomal degradation but affects class I major histocompatibility complex presentation. *The Journal of Biological Chemistry*, *283*(1), 244–254. https://doi.org/10.1074/jbc.M706237200

Katiyar, S., Joshi, S., & Lennarz, W. J. (2005). The retrotranslocation protein Derlin-1 binds peptide:N-glycanase to the endoplasmic reticulum. *Molecular Biology of the Cell*, *16*(10), 4584–4594. https://doi.org/10.1091/mbc.e05-04-0345

Kobayashi, A., Tsukide, T., Miyasaka, T., Morita, T., Mizoroki, T., Saito, Y., … Yamamoto, M. (2011). Central nervous system-specific deletion of transcription factor Nrf1 causes progressive motor neuronal dysfunction. *Genes to Cells : Devoted to Molecular & Cellular Mechanisms*, *16*(6), 692–703. https://doi.org/10.1111/j.1365-2443.2011.01522.x

Kondo, H., Rabouille, C., Newman, R., Levine, T. P., Pappin, D., Freemont, P., & Warren, G. (1997). p47 is a cofactor for p97-mediated membrane fusion. *Nature*, *388*(6637), 75–78. https://doi.org/10.1038/40411

Koumangoye, R., Omer, S., & Delpire, E. (2018). Mistargeting of a truncated Na-K-2Cl cotransporter in epithelial cells. *American Journal of Physiology. Cell Physiology*, *315*(2), C258–C276. https://doi.org/10.1152/ajpcell.00130.2018

Kowalczyk, A., Meyer, W. K., Partha, R., Mao, W., Clark, N. L., & Chikina, M. (2019). RERconverge: an R package for associating evolutionary rates with convergent traits. *Bioinformatics (Oxford, England)*, *35*(22), 4815–4817. https://doi.org/10.1093/bioinformatics/btz468

Lam, C., Ferreira, C., Krasnewich, D., Toro, C., Latham, L., Zein, W. M., … Wolfe, L. (2017).

Prospective phenotyping of NGLY1-CDDG, the first congenital disorder of deglycosylation. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *19*(2), 160–168. https://doi.org/10.1038/gim.2016.75

Landolt-Marticorena, C., & Reithmeier, R. A. (1994). Asparagine-linked oligosaccharides are localized to single extracytosolic segments in multi-span membrane glycoproteins. *The Biochemical Journal*, *302 (Pt 1*(1), 253–260. https://doi.org/10.1042/bj3020253

Lavoy, S., Chittoor-Vinod, V. G., Chow, C. Y., & Martin, I. (2018). Genetic Modifiers of Neurodegeneration in a Drosophila Model of Parkinson's Disease. *Genetics*, *209*(4), 1345–1356. https://doi.org/10.1534/genetics.118.301119

Lee, C. S., Lee, C., Hu, T., Nguyen, J. M., Zhang, J., Martin, M. V., … Chan, J. Y. (2011). Loss of nuclear factor E2-related factor 1 in the brain leads to dysregulation of proteasome gene expression and neurodegeneration. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(20), 8408–8413. https://doi.org/10.1073/pnas.1019209108

Lehrbach, N. J., Breen, P. C., & Ruvkun, G. (2019). Protein Sequence Editing of SKN-1A/Nrf1 by Peptide:N-Glycanase Controls Proteasome Gene Expression. *Cell*, *177*(3), 737-750.e15. https://doi.org/10.1016/j.cell.2019.03.035

Lehrbach, N. J., & Ruvkun, G. (2016). Proteasome dysfunction triggers activation of SKN-1A/Nrf1 by the aspartic protease DDI-1. *ELife*, *5*(AUGUST), 1–19. https://doi.org/10.7554/eLife.17721

Leiserson, W. M., Forbush, B., & Keshishian, H. (2011). Drosophila glia use a conserved cotransporter mechanism to regulate extracellular volume. *Glia*, *59*(2), 320–332. https://doi.org/10.1002/glia.21103

Löytynoja, A., & Goldman, N. (2008). Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science (New York, N.Y.)*, *320*(5883), 1632–1635. https://doi.org/10.1126/science.1158395

Mackay, T. F. C., Richards, S., Stone, E. A., Barbadilla, A., Ayroles, J. F., Zhu, D., … Gibbs, R. A. (2012). The Drosophila melanogaster Genetic Reference Panel. *Nature*, *482*(7384), 173–178. https://doi.org/10.1038/nature10811

Macnamara, E. F., Koehler, A. E., D'Souza, P., Estwick, T., Lee, P., Vezina, G., … Tifft, C. J. (2019). Kilquist syndrome: A novel syndromic hearing loss disorder caused by homozygous deletion of SLC12A2. *Human Mutation*, *40*(5), 532–538. https://doi.org/10.1002/humu.23722

McNeill, H., Knebel, A., Arthur, J. S. C., Cuenda, A., & Cohen, P. (2004). A novel UBA and UBX domain protein that binds polyubiquitin and VCP and is a substrate for SAPKs. *The Biochemical Journal*, *384*(Pt 2), 391–400. https://doi.org/10.1042/BJ20041498

Misaghi, S., Pacold, M. E., Blom, D., Ploegh, H. L., & Korbel, G. A. (2004). Using a small molecule inhibitor of peptide: N-glycanase to probe its role in glycoprotein turnover. *Chemistry & Biology*, *11*(12), 1677–1687. https://doi.org/10.1016/j.chembiol.2004.11.010

Nakajima, K., Niisato, N., & Marunaka, Y. (2011). Quercetin stimulates NGF-induced neurite outgrowth in PC12 cells via activation of Na(+)/K(+)/2Cl(-) cotransporter. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *28*(1), 147–156. https://doi.org/10.1159/000331723

Need, A. C., Shashi, V., Hitomi, Y., Schoch, K., Shianna, K. V., McDonald, M. T., … Goldstein, D. B. (2012). Clinical application of exome sequencing in undiagnosed genetic conditions. *Journal of Medical Genetics*, *49*(6), 353–361. https://doi.org/10.1136/jmedgenet-2012-100819

Owings, K. G., Lowry, J. B., Bi, Y., Might, M., & Chow, C. Y. (2018). Transcriptome and functional analysis in a Drosophila model of NGLY1 deficiency provides insight into therapeutic approaches. *Human Molecular Genetics*, *27*(6), 1055–1066. https://doi.org/10.1093/hmg/ddy026

Palu, R. A. S., Ong, E., Stevens, K., Chung, S., Owings, K. G., Goodman, A. G., & Chow, C. Y. (2019). Natural Genetic Variation Screen in Drosophila Identifies Wnt Signaling, Mitochondrial Metabolism, and Redox Homeostasis Genes as Modifiers of Apoptosis. *G3*, *9*(12), 3995–4005. https://doi.org/10.1534/g3.119.400722

Paredes, A., Plata, C., Rivera, M., Moreno, E., Vázquez, N., Muñoz-Clares, R., … Gamba, G. (2006). Activity of the renal Na+-K+-2Cl- cotransporter is reduced by mutagenesis of N-glycosylation sites: role for protein surface charge in Cl- transport. *American Journal of Physiology. Renal Physiology*, *290*(5), F1094-102. https://doi.org/10.1152/ajprenal.00071.2005

Park, H., Suzuki, T., & Lennarz, W. J. (2001). Identification of proteins that interact with mammalian peptide:N-glycanase and implicate this hydrolase in the proteasome-dependent pathway for protein degradation. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11163–11168. https://doi.org/10.1073/pnas.201393498

Payne, J. A., Xu, J. C., Haas, M., Lytle, C. Y., Ward, D., & Forbush, B. (1995). Primary structure, functional expression, and chromosomal localization of the bumetanide-sensitive Na-K-Cl cotransporter in human colon. *The Journal of Biological Chemistry*, *270*(30), 17977–17985. https://doi.org/10.1074/jbc.270.30.17977

Priedigkeit, N., Wolfe, N., & Clark, N. L. (2015). Evolutionary signatures amongst disease genes permit novel methods for gene prioritization and construction of informative gene-based networks. *PLoS Genetics*, *11*(2), e1004967. https://doi.org/10.1371/journal.pgen.1004967

Qi, L., Tsai, B., & Arvan, P. (2017). New Insights into the Physiological Role of Endoplasmic Reticulum-Associated Degradation. *Trends in Cell Biology*, *27*(6), 430–440. https://doi.org/10.1016/j.tcb.2016.12.002

Raza, Q., Choi, J. Y., Li, Y., O'Dowd, R. M., Watkins, S. C., Chikina, M., … Kwiatkowski, A. V. (2019). Evolutionary rate covariation analysis of E-cadherin identifies Raskol as a regulator of cell adhesion and actin dynamics in Drosophila. *PLoS Genetics*, *15*(2), e1007720. https://doi.org/10.1371/journal.pgen.1007720

Renn, S. C. P., Park, J. H., Rosbash, M., Hall, J. C., & Taghert, P. H. (1999). A pdf

neuropeptide gene mutation and ablation of PDF neurons each cause severe abnormalities of behavioral circadian rhythms in Drosophila. *Cell*, *99*(7), 791–802. https://doi.org/10.1016/s0092-8674(00)81676-1

Rusan, Z. M., Kingsford, O. A., & Tanouye, M. A. (2014). Modeling glial contributions to seizures and epileptogenesis: cation-chloride cotransporters in Drosophila melanogaster. *PloS One*, *9*(6), e101117. https://doi.org/10.1371/journal.pone.0101117

Scholz, E. P., Zitron, E., Katus, H. A., & Karle, C. A. (2010). Cardiovascular ion channels as a molecular target of flavonoids. *Cardiovascular Therapeutics*, *28*(4), e46-52. https://doi.org/10.1111/j.1755-5922.2010.00212.x

Simon, D. B., & Lifton, R. P. (1996). The molecular basis of inherited hypokalemic alkalosis: Bartter's and Gitelman's syndromes. *The American Journal of Physiology*, *271*(5 Pt 2), F961-6. https://doi.org/10.1152/ajprenal.1996.271.5.F961

Singh, R., Almutairi, M. M., Pacheco-Andrade, R., Almiahuob, M. Y. M., & Di Fulvio, M. (2015). Impact of Hybrid and Complex N-Glycans on Cell Surface Targeting of the Endogenous Chloride Cotransporter Slc12a2. *International Journal of Cell Biology*, *2015*, 505294. https://doi.org/10.1155/2015/505294

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., … Mesirov, J. P. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/pnas.0506580102

Sugawara, E., & Nikaido, H. (2014). Properties of AdeABC and AdeIJK efflux systems of Acinetobacter baumannii compared with those of the AcrAB-TolC system of Escherichia coli. *Antimicrobial Agents and Chemotherapy*, *58*(12), 7250–7257. https://doi.org/10.1128/AAC.03728-14

Tambe, M. A., Ng, B. G., & Freeze, H. H. (2019). N-Glycanase 1 Transcriptionally Regulates Aquaporins Independent of Its Enzymatic Activity. *Cell Reports*, *29*(13), 4620-4631.e4. https://doi.org/10.1016/j.celrep.2019.11.097

Tomlin, F. M., Gerling-Driessen, U. I. M., Liu, Y.-C., Flynn, R. A., Vangala, J. R., Lentz, C. S., … Bertozzi, C. R. (2017). Inhibition of NGLY1 Inactivates the Transcription Factor Nrf1 and Potentiates Proteasome Inhibitor Cytotoxicity. *ACS Central Science*, *3*(11), 1143–1155. https://doi.org/10.1021/acscentsci.7b00224

Wolfe, N. W., & Clark, N. L. (2015). ERC analysis: web-based inference of gene function via evolutionary rate covariation. *Bioinformatics (Oxford, England)*, *31*(23), 3835–3837. https://doi.org/10.1093/bioinformatics/btv454

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, *44*(7), 821–824. https://doi.org/10.1038/ng.2310

1   Zhu, B., Jiang, L., Huang, T., Zhao, Y., Liu, T., Zhong, Y., … Xu, H. (2017). ER-associated
2   degradation regulates Alzheimer's amyloid pathology and memory function by modulating
3   γ-secretase activity. *Nature Communications, 8*(1), 1472. https://doi.org/10.1038/s41467-
4   017-01799-4

5   Zolekar, A., Lin, V. J. T., Mishra, N. M., Ho, Y. Y., Hayatshahi, H. S., Parab, A., … Wang, Y.-C.
6   (2018). Stress and interferon signalling-mediated apoptosis contributes to pleiotropic
7   anticancer responses induced by targeting NGLY1. *British Journal of Cancer, 119*(12),
8   1538–1551. https://doi.org/10.1038/s41416-018-0265-9

9

10
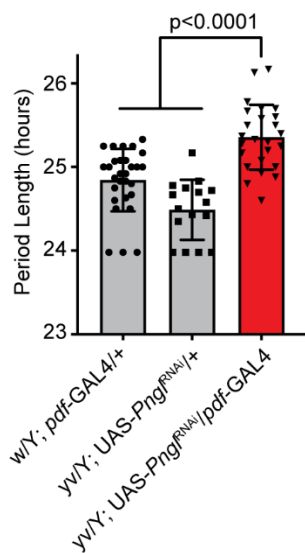
1   **Supplement**

2

3   **Supplemental Figure 1. *NGLY1* knockdown causes circadian rhythm defect.** Activity monitor was used to analyze 2-5 day old
4   flies for 1 week in complete darkness after three days of entrainment in a 12-hour light, dark cycle. At least 15 flies were analyzed for
5   each genotype. Period length was calculated from activity using ClockLab. One-way ANOVA gave an overall P < 0.0001. Subsequently
6   Tukey's test was used to calculate individual adjusted p-values between genotypes shown on the graph.

A   **B**
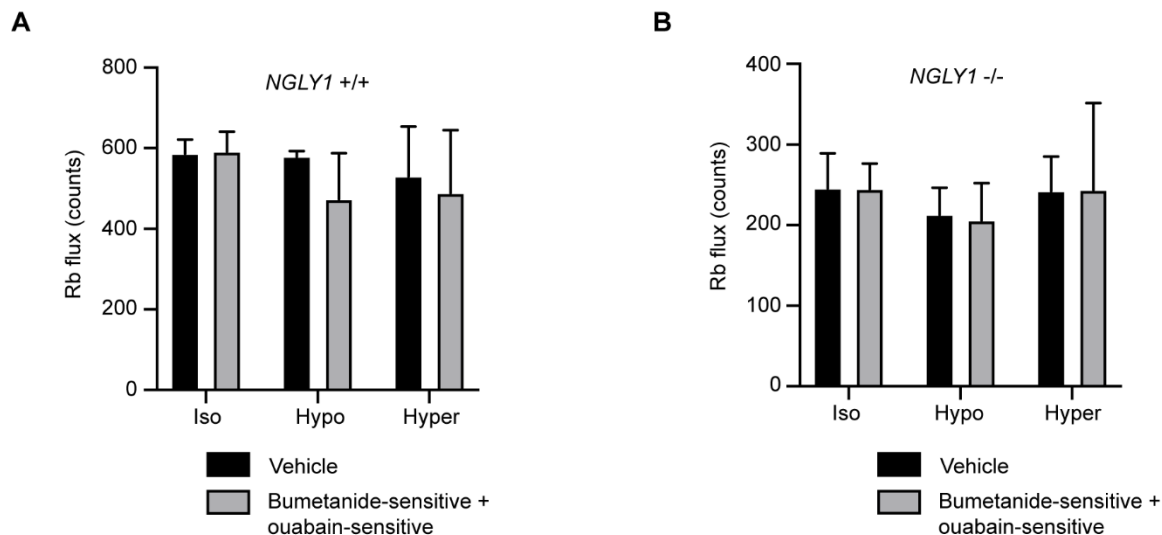
7

8   **Supplemental Figure 2. ^{86}Rb uptake in MEFs occurs through bumetanide-sensitive and ouabain-sensitive pathways**. Either
9   NGLY1 +/+ (**A**) or NGLY1 -/- (**B**) MEFs were incubated with DMSO as a vehicle control or with the NKCC1 inhibitor, bumetanide, or
10   the Na$^+$/K$^+$-ATPase inhibitor, ouabain. ^{86}Rb flux was measured in 3 conditions: isotonic (iso), hypotonic (hypo), and hypertonic (hyper).
11   The sum of bumetanide-sensitive and ouabain-sensitive flux was compared to vehicle control. There were no significant effects of
12   genotype (p=0.3267), condition (p=0.3602) or an interaction (p=0.6244) in NGLY1 +/+ cells, nor of genotype (p=0.9422), condition
13   (p=0.4987) or an interaction (p=0.9909) in NGLY1 -/- cells by two-way ANOVA, indicating that in both cell types ^{86}Rb flux was
14   comprised of the bumetanide-sensitive and ouabain-sensitive activities.