Appendix: DDX3X and Specific Initiation Factors Modulate *FMR1* Repeat-Associated Non-AUG initiated translation

Appendix Table of contents: **Page #:**

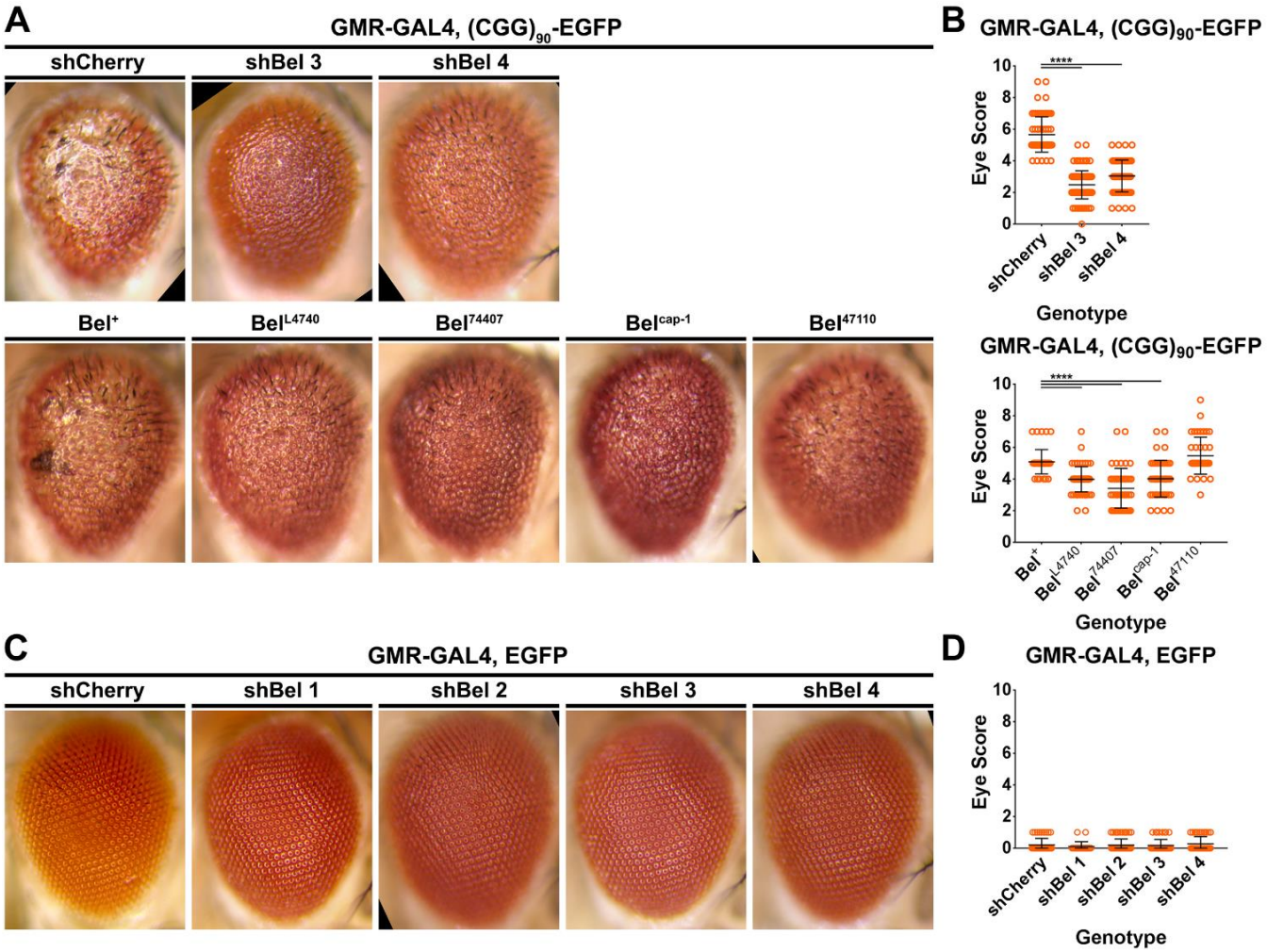
		GMR-GAL4, $(CGG)_{90}$-EGFP	GMR-GAL4
shCherry			
Belle	shBelle 1		
	shBelle 2		
	shBelle 3		
	shBelle 4		
	$Belle^{6}$		
	$Belle^{EKE}$		
	$Belle^{L4740}$		
	$Belle^{74407}$		
	$Belle^{cap-1}$		
	$Belle^{47110}$		
eIF4B	shEIF4B		
	UAS-EIF4B		
eIF4H1 , 2	shEIF4H1 1		
	shEIF4H1 2		
	shEIF4H1 3		
	shEIF4H2 1		
	shEIF4H2 2		
eIF1	shEIF1 1		
	shEIF1 2		
	$EIF1^{EY02210}$		
	shEIF1 3		
	shEIF1 4		
	shEIF1 5		
eIF1A	shEIF1A 1		
	shEIF1A 2		
	shEIF1A 3		
	shEIF1A 4		
	$EIF1A^{645}$		
	$EIF1A^{2232}$		
	$EIF1A^{c04533}$		
	$EIF1A^{EP935}$		
	UAS-EIF1A		

		GMR-GAL4, $(CGG)_{90}$-EGFP	GMR-GAL4
eIF2α	shEIF2α 1		
	shEIF2α 2		
	shEIF2α 3		
	$EIF2α^{815-29}$		
	UAS-EIF2α		
eIF3B	$EIF3B^{EY14430}$		
eIF4A	shEIF4A		
	UAS-EIF4A		
	$EIF4A^{1013}$		
eIF5, eIF5B	shEIF5 1		
	shEIF5 2		
	UAS-EIF5		
	$EIF5B^{09143}$		
	$EIF5B^{EY01401}$		
RNA Helicases	shRHAU		
	UAS-RHAU		
	shDHX57		
	shBGCN 1		
	shBGCN 2		
	shBGCN 3		
	$BGCN^{KG08129}$		
Ribosomal Proteins	shRPS25 1		
	shRPS25 2		
RNA-Binding Proteins	shSF2 1		
	shSF2 2		

Suppressor	
Enhancer	
No Effect	

Appendix Figure S1. Summary of the Candidate-Based Screen For Modifiers of $(CGG)_{90}$-Elicited Toxicity.

Candidate modifier genes for RAN translation were crossed to either GMR-Gal4; $(CGG)_{90}$-eGFP flies or GMR-Gal4 alone and their eye phenotypes were assessed. Blue indicates suppression of the rough eye phenotype. Salmon indicates exacerbation or induction of a rough eye phenotype. Each candidate modifier was screened against both GMR-Gal4; $(CGG)_{90}$-eGFP flies or GMR-Gal4 alone across a minimum of two independent crosses with at least 25 flies evaluated per cross.



Appendix Figure S2. *Belle* Disruption Mitigates $(CGG)_{90}$-Elicited Toxicity.

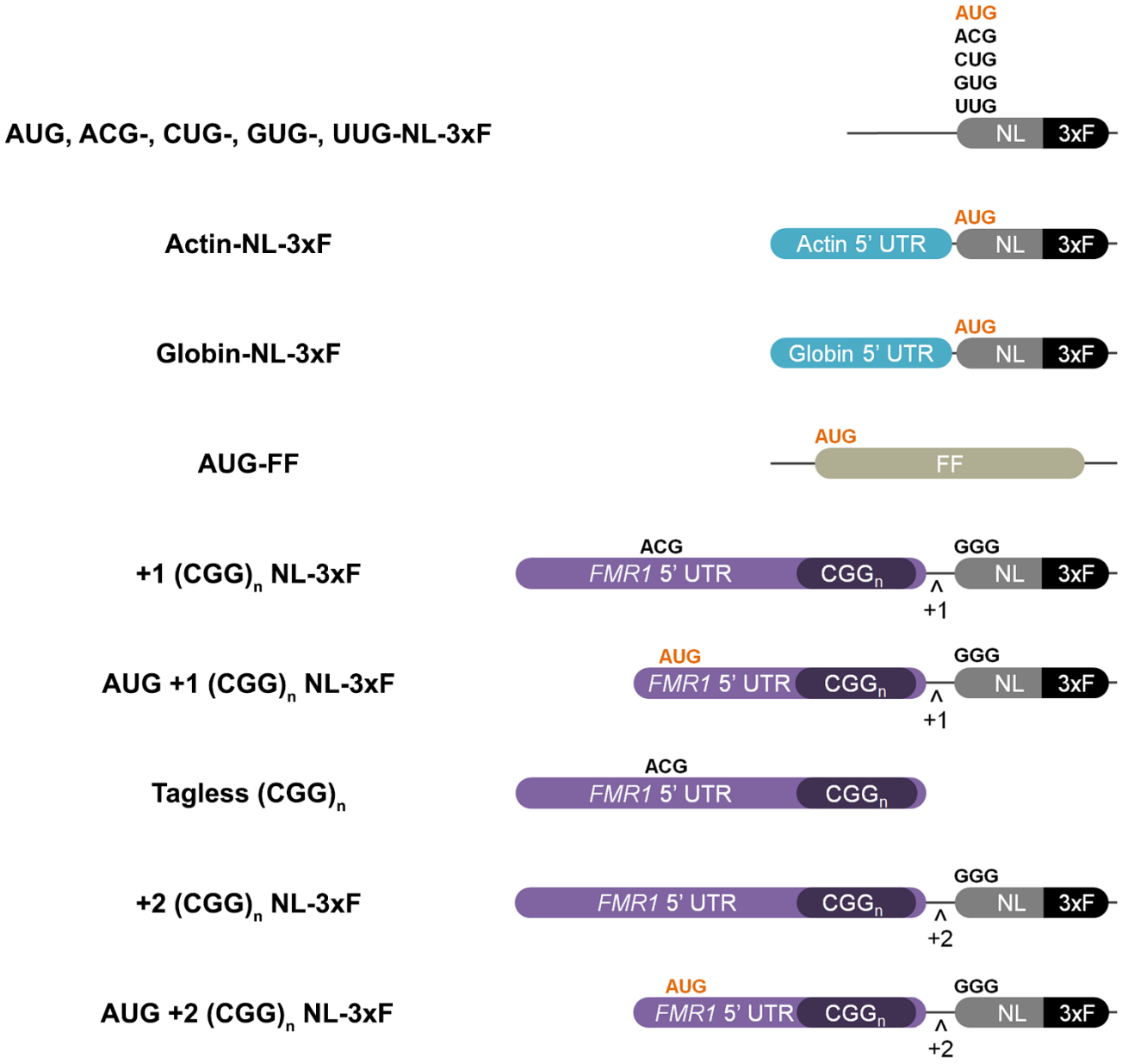
A Representative photographs of fly eyes expressing GMR-GAL4, $(CGG)_{90}$-EGFP with additional *belle* disruptions.

B Quantitation of GMR-GAL4, $(CGG)_{90}$-EGFP eye phenotypes with additional *belle* disruptions (Mann-Whitney U test with Bonferonni corrections for multiple comparisons; *n*=42-98/genotype)

C Representative photographs of fly eyes expressing GMR-GAL4 and an AUG-initiated EGFP, along with *belle* shRNAs.

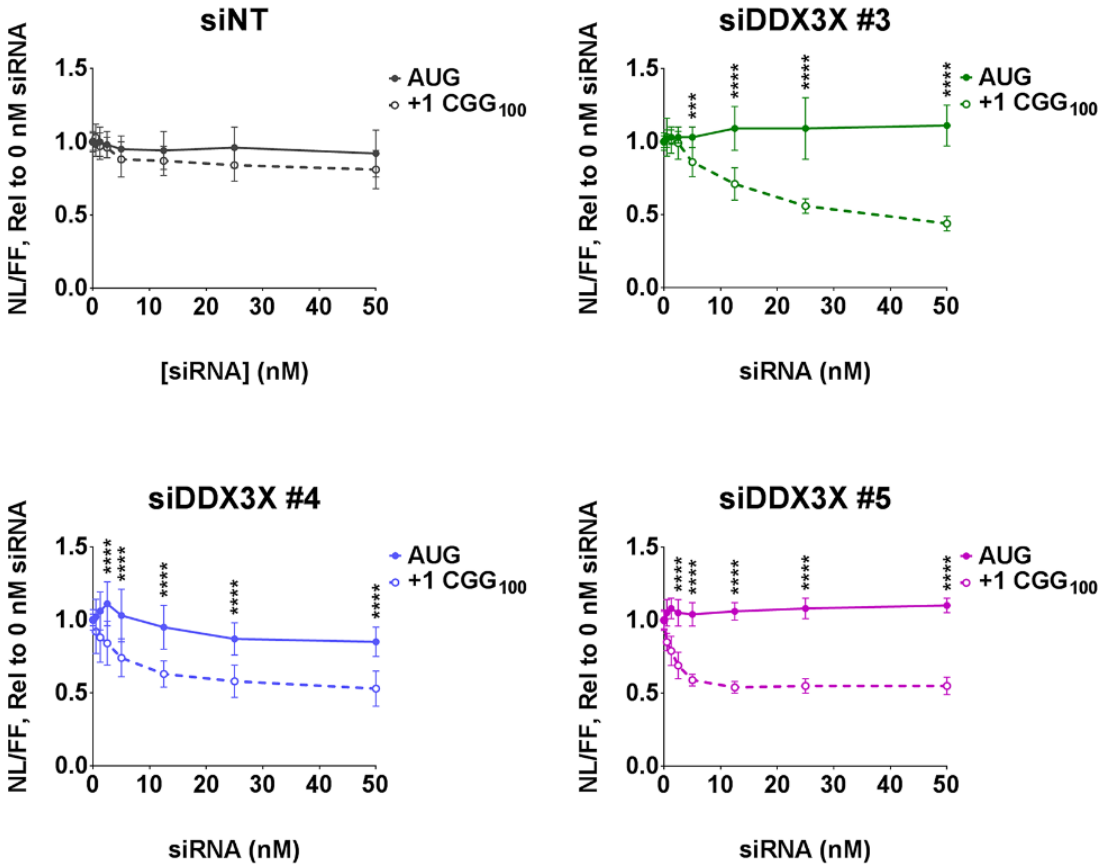
D Quantitation of GMR-GAL4, EGFP eye phenotypes with *belle* shRNAs (*n*=20-64/genotype).

Data Information: For all panels, **** P≤0.0001 for the specified statistical test (compiled from ≥3 replicates).



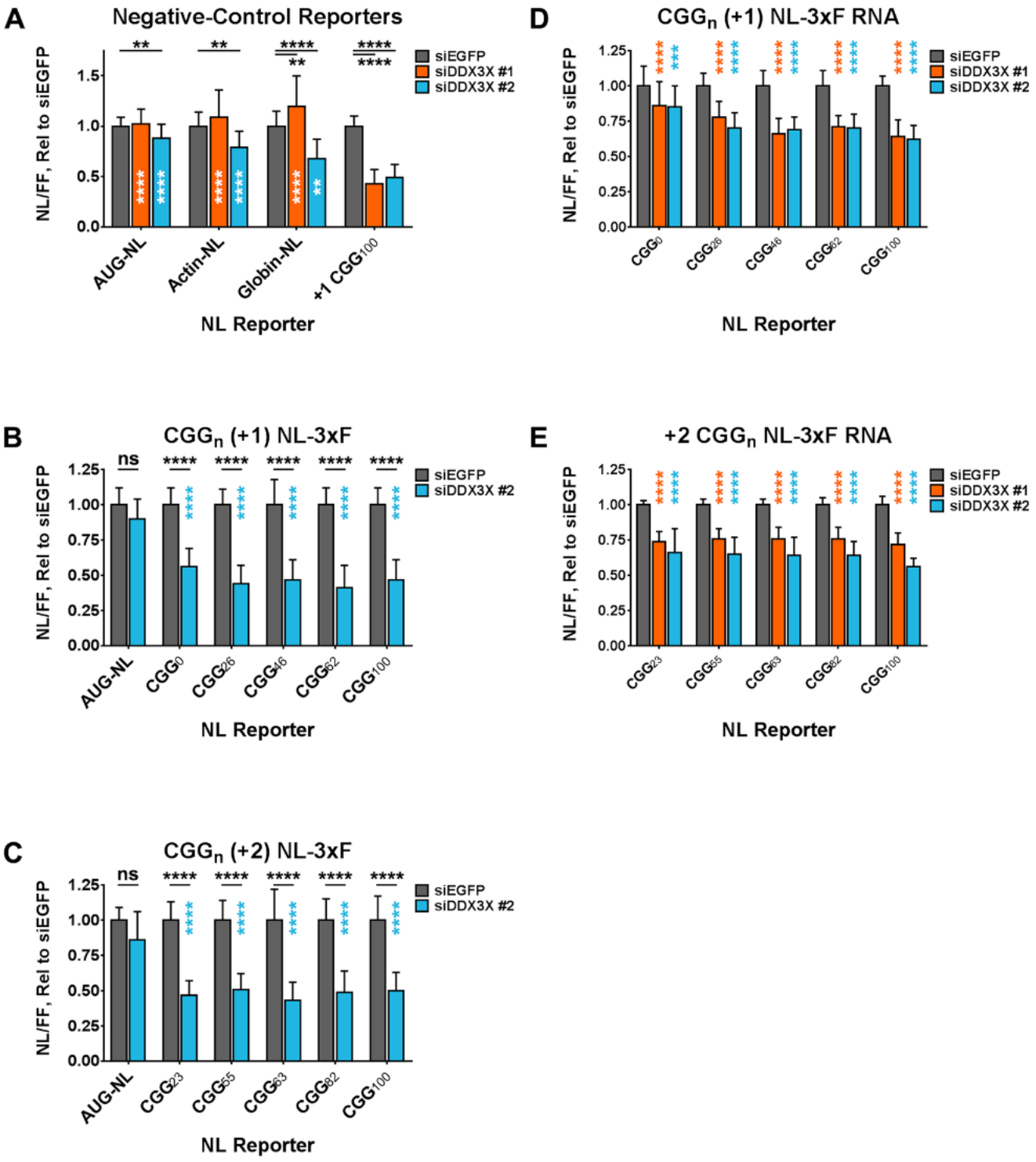
Appendix Figure S3. NL-3XF and FF Reporter Constructs Used in This Study.

The tag (e.g., NL-3xF, FF, tagless), 5' leader sequence (e.g., *β Actin* or *β Globin* 5' UTR), start codon (e.g., AUG, ACG, GGG), and tag reading frame relative to the CGG repeat (e.g., +1, +2) are indicated.



Appendix Figure S4. Knockdown of *DDX3X* by 3 Additional siRNAs Selectively Inhibits RAN Translation of +1 $(CGG)_{100}$ NL-3xF.

Asterisks indicate comparisons between relative AUG-NL-3xF and +1 $(CGG)_{100}$ NL-3xF expression (two-way ANOVA with Sidak's multiple comparisons test; *n*=9-21/condition). *** P≤0.001, **** P≤0.0001 for the specified statistical test. Points represent means ± SD (compiled from ≥3 replicates).



Appendix Figure S5. *DDX3X* Knockdown Inhibits RAN Translation at CGG Repeats Selectively.

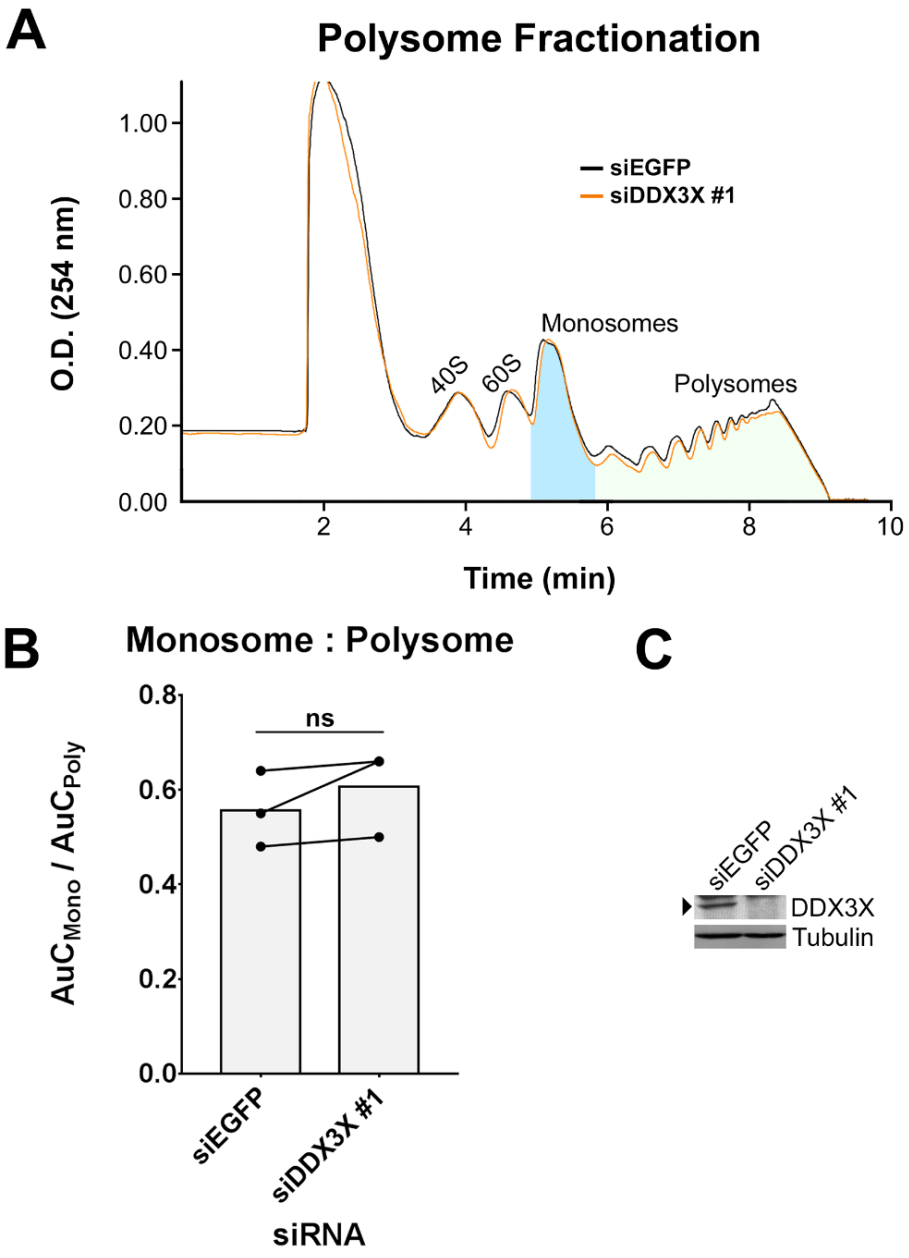
A Expression in HeLa cells of transfected AUG-NL-3xF and +1 $(CGG)_{100}$ NL-3xF reporter plasmids, with and without *DDX3X* knockdown, compared to the expression of AUG-initiated NL-3xF reporters bearing the short, minimally-structured 5' UTRs of human *β actin* and *β globin* (two-way ANOVA with Tukey's multiple comparisons test; *n*=15-39/condition). Black asterisks refer to comparisons between

siEGFP-treated and siDDX3X-treated cells; white asterisks refer to comparisons between siDDX3X-treated cells expressing +1 $(CGG)_{100}$ NL-3xF and those expressing a different reporter.

B, C $(CGG)_n$ +1 and $(CGG)_n$ +2 NL-3xF expression with and without *DDX3X* knockdown across a range of CGG repeat sizes (two-way ANOVA with Tukey's multiple comparisons test; *n*=12-18/condition). Black asterisks refer to comparisons between siDDX3X- and siEGFP-treated cells; blue asterisks refer to comparisons between siDDX3X-treated cells expressing AUG-NL-3xF and those expressing a different reporter.

D, E Expression of *in vitro*-transcribed $(CGG)_n$ +1 and $(CGG)_n$ +2 NL-3xF reporter RNAs with and without *DDX3X* knockdown across a range of CGG repeat sizes (two-way ANOVA with Tukey's multiple comparisons test; *n*=12-24/condition). Orange and blue asterisks refer to comparisons between siDDX3X- and siEGFP-treated cells.

Data Information: For all panels, ns=non-significant, ** P≤0.01, **** P≤0.0001 for the specified statistical test. All panels depict data as means ± SD (compiled from ≥3 replicates).

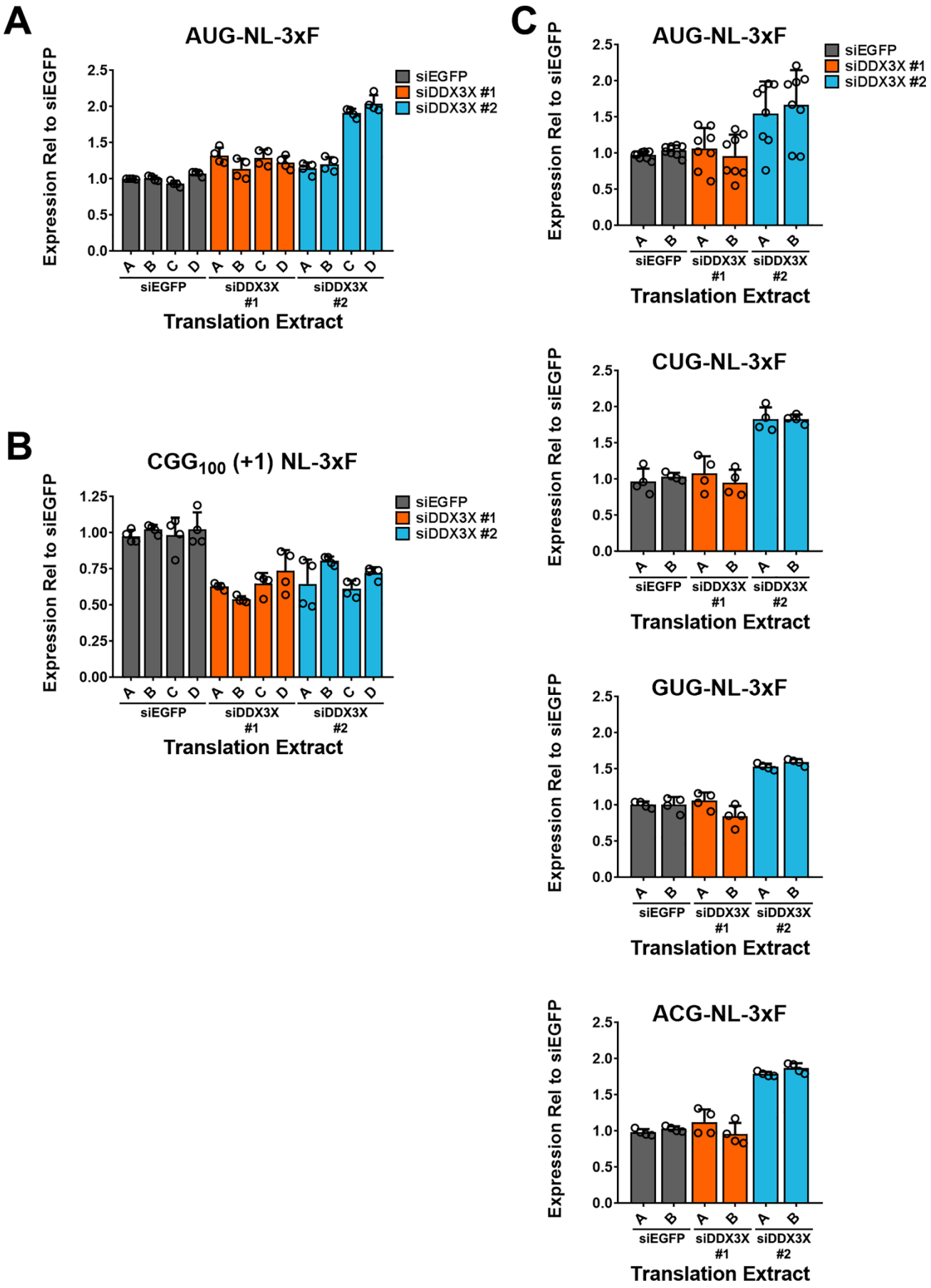


Appendix Figure S6. Knockdown of *DDX3X* Does Not Inhibit Global Translation.

A Representative polysome-fractionation profiles of HeLa-cell lysates transfected with siDDX3X #1 or siEGFP. The areas-under-the-curve (AuC) for monosomes and polysomes are shaded blue and green, respectively.

B Mean ratios (±SD) of the AuCs of monosomes to polysomes across three replicates (Student's paired t-test, *n*=3/condition). ns=non-significant.

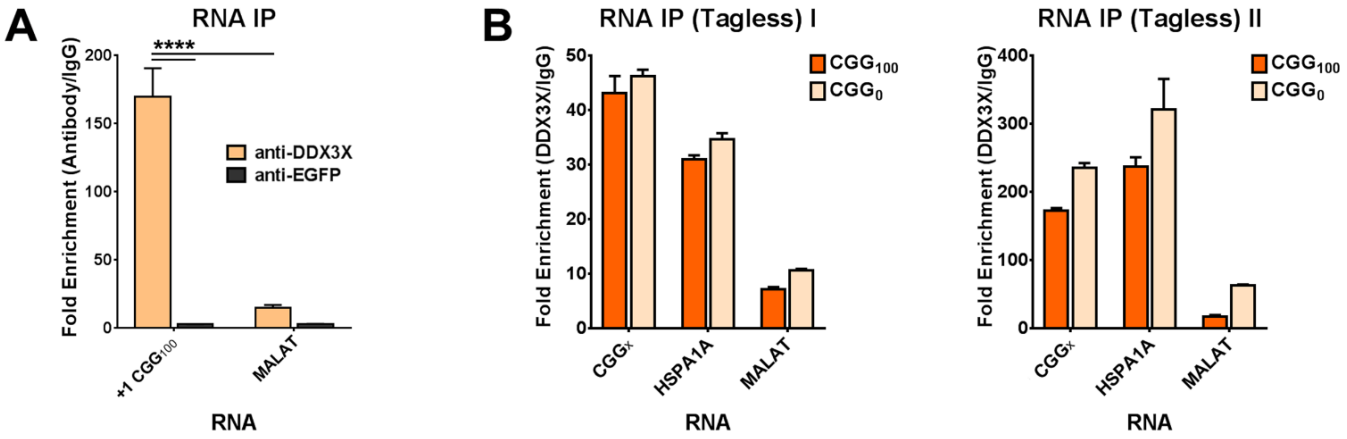
C Anti-DDX3X western blot of HeLa lysates processed for polysome fractionation.


A
AUG-NL-3xF
Expression Rel to siEGFP
siEGFP
siDDX3X #1
siDDX3X #2
Translation Extract
B
CGG100 (+1) NL-3xF
C
CUG-NL-3xF
GUG-NL-3xF
ACG-NL-3xF

Appendix Figure S7. Individual replicates from Independently-Prepared Translation Extracts.

A, B Expression of *in vitro*-transcribed AUG-NL-3xF (A) and +1 $(CGG)_{100}$ NL-3xF RNAs (B) in *in vitro* translation extracts, collected from HeLa cells treated with siRNAs against EGFP or *DDX3X*. Four replicate lysates ("A-D") were generated per siRNA. Panels depict pooled data (mean ± SD) gathered across two replicates.

C Expression of *in vitro* transcribed near-AUG reporter RNAs in *in vitro* translation extracts, collected from HeLa cells treated with siRNAs against EGFP or *DDX3X*. As above, panels depict pooled data gathered across two replicates. Reporter RNAs were tested in duplicate lysates (2 per siRNA).



Appendix Figure S8. Anti-DDX3X RNA Immunoprecipitation (RIP) Co-Precipitates +1 $(CGG)_{100}$ NL-3xF mRNA, Independent of the NL-3xF Tag or NRE Size.

A Enrichment of +1 $(CGG)_{100}$ NL-3xF mRNA following anti-DDX3X RIP, relative to incubation with isotype control IgG. *MALAT* RNA, in contrast, is not enriched (Student's t test, *n*=3). In addition, +1 $(CGG)_{100}$ NL-3xF mRNA is not enriched following anti-EGFP RIP from cells expressing EGFP (Student's t test, *n*=3). This experiment is a replicate of that presented in Figure 4D.

B Enrichment of $(CGG)_{100}$, $(CGG)_0$, and *HSPA1A*, but not MALAT RNA, following anti-DDX3X IP in two independent replicates. In contrast with panel A and Figure 4D, the $(CGG)_{100}$ and $(CGG)_0$ constructs do not have a NL-3xF tag.

Data Information: **** P≤0.0001.

+1 $(CGG)_n$ NL-3xF

TAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTA
AGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATTCGTTA
CAGATCTGCTCAGCTCCGTTTCGGTTTCACTTCCGGTGGAGGGCCG
CCTCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGT
G**ACG**GAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCG(**CGG**)$_n$CTGG
GCCTCGAGGATATCAAGATCTGGCCTCGGCGGCCAAGCTTGGCAATCCGGT
ACTGTTGGTAAAGCCACCGGGGTCTTCACACTC...

AUG +1 $(CGG)_{100}$ NL-3xF

TAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTA
AGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATTCGTTA
CACC**ATG**GCGCCGCTGCCAGGGGGCGTGCGGCAGCG(**CGG**)$_{100}$CTGGG
CCTCGAGGATATCAAGATCTGGCCTCGGCGGCCAAGCTTGGCAATCCGG
TACTGTTGGTAAAGCCACCGGGGTCTTCACACTC...

+2 $(CGG)_n$ NL-3xF

TAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTA
AGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATTCGTTA
ACAGATCTGCTCAGCTCCGTTTCGGTTTCACTTCCGGTGGAGGGCCG
CCTCTGAGCGGGCGGCGGGCCGACGGCGAGCGCGGGCGGCGGCGGT
GACGGAGGCGCCGCTGCCAGGGGGCGTGCGGCAGCG(**CGG**)$_n$CTGGG
CCATCAAGATCTGGCCTCGGCGGCCAAGCTTGGCAATCCGGTACTGTTG
GTAAAGCCACCGGGGTCTTCACACTC...

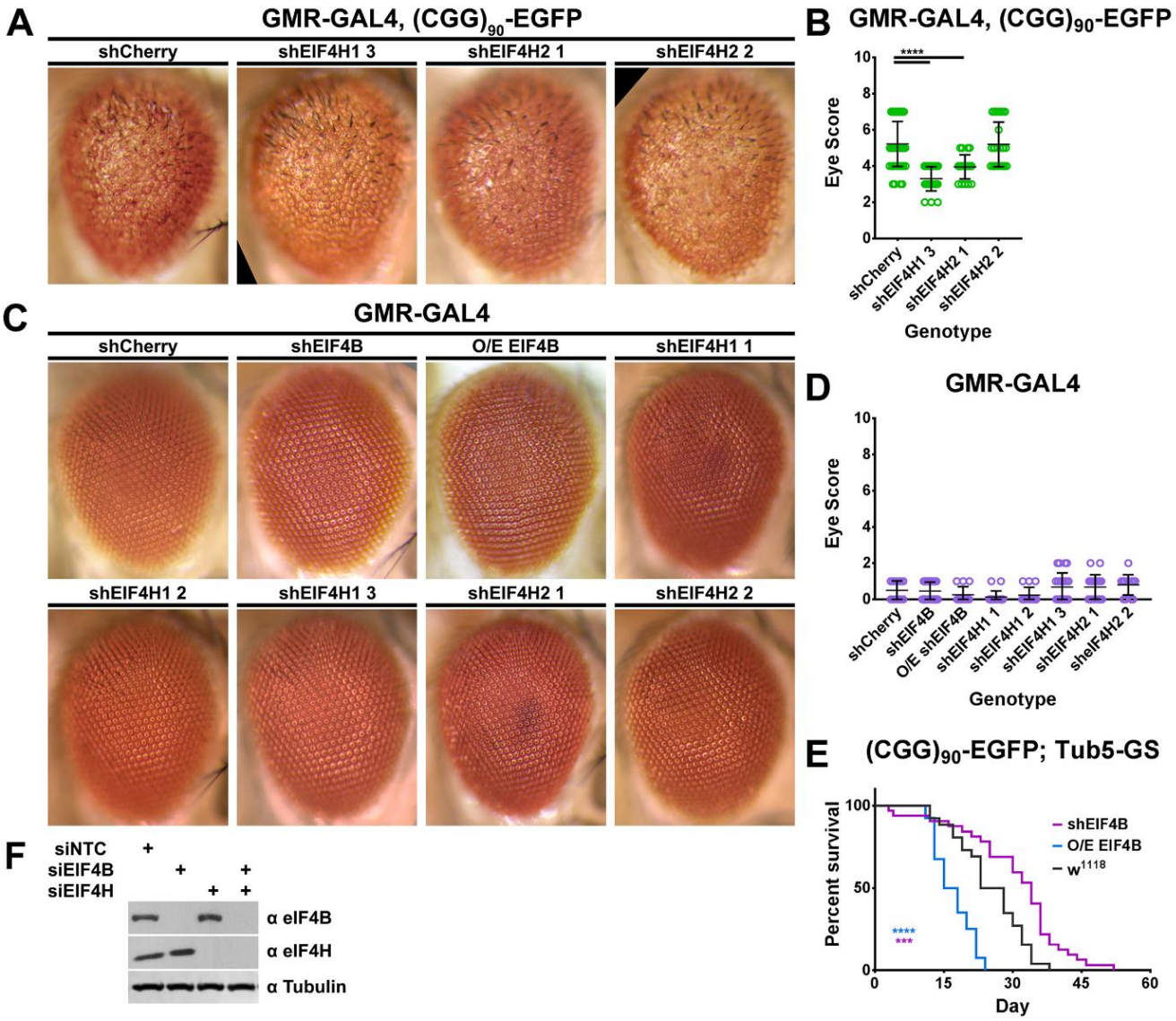
AUG +2 $(CGG)_{100}$ NL-3xF

TAATACGACTCACTATAGGGAGACCCAAGCTGGCTAGCGTTTAAACTTA
AGCTTGGTACCGAGCTCGGATCCACTAGTCCAGTGTGGTGGAATTCGTT
ACACC**ATG**GGGCGCCGCTGCCAGGGGGCGTGCGGCAGCG(**CGG**)$_{100}$CTG
GGCCTCGAGATTCGATCGTCGTGATATCAAGATCTGGCCTCGGCGG
CCAAGCTTGGCAATCCGGTACTGTTGGTAAAGCCACCGGGGTCTTC
ACACTC...

T7 Promotor Site
FMR1 5' UTR
NL Open Reading Frame

AUG Start Codon
Non-AUG Start Codon
CGG NRE

Appendix Figure S9. Sequence Maps of the +1 and +2 (CGG)100 NL-3xF Reporter Constructs, With and Without an AUG Inserted 5' to the CGG NRE.



Appendix Figure S10. Disruption of *eIF4B* or *eIF4H* Mitigate $(CGG)_{90}$-Elicited Toxicity.

A Representative photographs of fly eyes expressing GMR-GAL4, $(CGG)_{90}$-EGFP with additional *eIF4H1* and *eIF4H2* shRNAs.

B Quantitation of GMR-GAL4, $(CGG)_{90}$-EGFP eye phenotypes with additional *eIF4H1* and *eIF4H2* shRNAs (Mann-Whitney U test with Bonferonni corrections for multiple comparisons; *n*=21-55/genotype).

C Representative photographs of fly eyes expressing GMR-GAL4, without $(CGG)_{90}$-EGFP, along with *eIF4H1* and *eIF4H2* shRNAs.

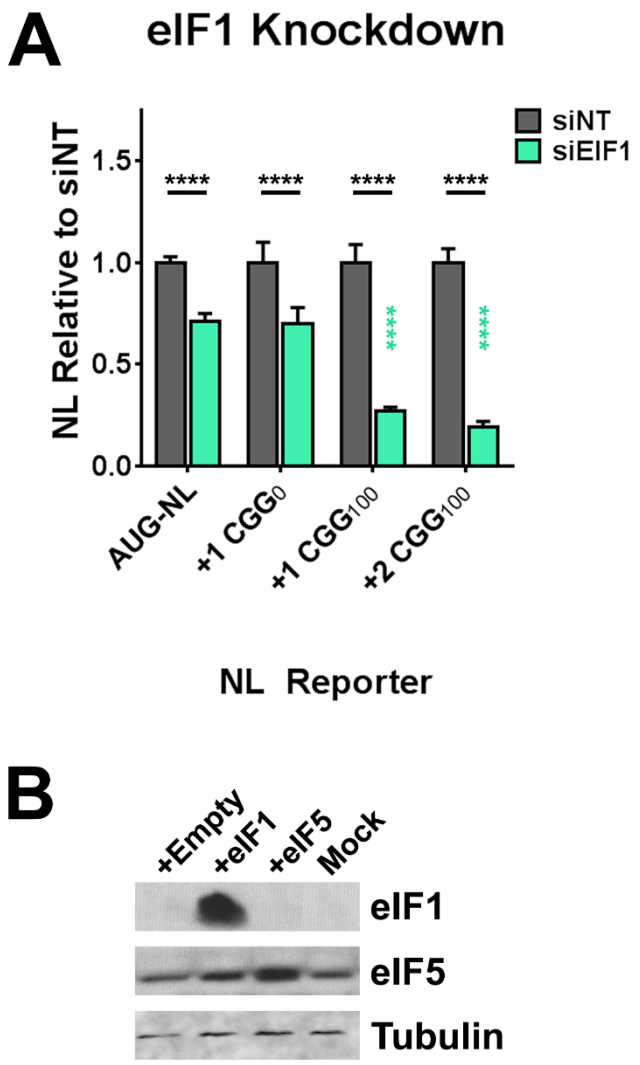
D Quantitation of GMR-GAL4 eye phenotypes with *eIF4B*, *eIF4H1*, and *eIF4H2* shRNAs (*n*=12-

33/genotype).

E Longevity assays of $(CGG)_{90}$-EGFP; Tub5-GS flies with knockdown and over-expression of *eIF4B* (Log-rank Mantel-Cox test with Bonferroni corrections for multiple comparisons; *n*=26-32/genotype).

F Anti-eIF4B and anti-eIF4H western blot of HeLa cells transfected with siRNAs against *EIF4B* or *EIF4H*.

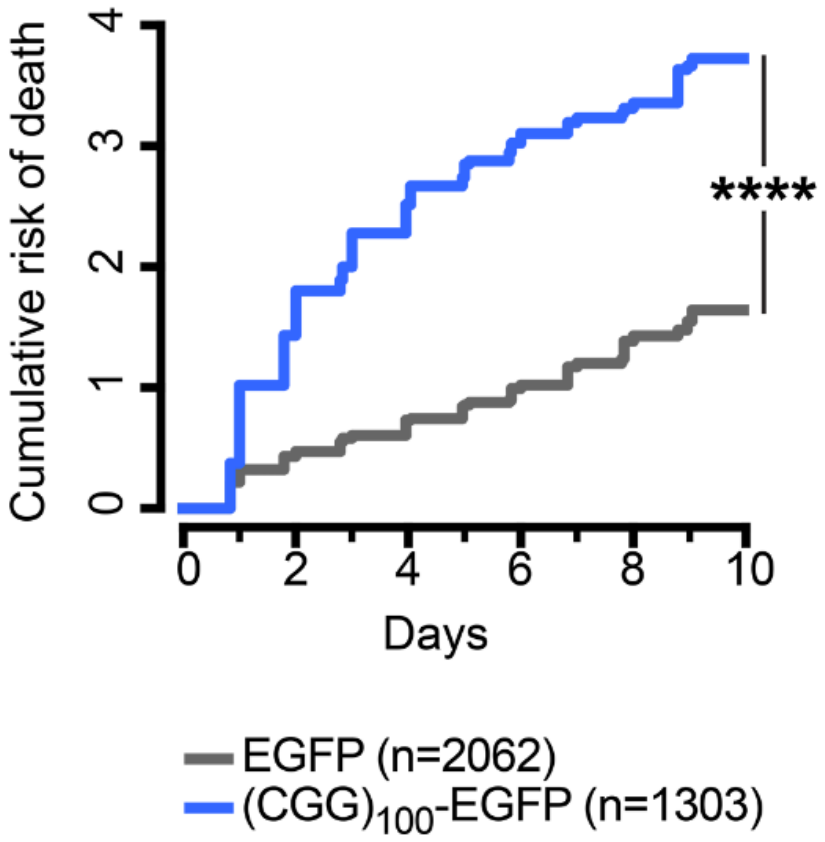
Data Information: For all panels, ns=non-significant, *** P≤0.001, **** P≤0.0001 for the specified statistical test. All panels depict data as means ± SD (compiled from ≥3 replicates).



Appendix Figure S11. *EIF1* Knockdown Inhibits Translation Reporters Globally.

A Expression of plasmid-based NL-3xF reporters in HEK293 cells with and without knockdown of *EIF1* (two-way ANOVA with Sidak's multiple comparisons test; *n*=12/condition). Black asterisks refer to comparisons between non-targeting siRNA (siNT)-transfected and siEIF1-transfected cells; green asterisks refer to comparisons between siEIF1-transfected cells expressing AUG-NL-3xF and those expressing a different reporter. Graph depicts data as mean ± SD. **** P≤0.0001. Panel depicts data as means ± SD (compiled from ≥3 replicates).

B Anti-eIF1 and –eIF5 western blot demonstrating over-expression of eIF1 and eIF5 after transfection of the respective plasmid constructs.



Appendix Figure S12. +1 $(CGG)_{100}$ EGFP is More Toxic to Primary Rodent Neurons Than an AUG-Initiated EGFP Construct by Longitudinal Automated Fluorescence Microscopy. Transfection of +1 $(CGG)_{100}$ EGFP plasmid-based reporters increased the cumulative risk of death in primary rodent neurons, relative to transfection of EGFP reporters. Cox proportional hazard analysis; *n*=1303-2062 cells/condition. **** P<0.0001. Data compiled over 3 replicates.

Fly Stock	Source	Catalog #
shCherry	BDSC	35785
shBel 1	BDSC	35185
shBel 2	BDSC	35302
shBel 3	VDRC	6299
shBel 4	BDSC	28049
w^{1118}	BDSC	5905
Bel^{6}	BDSC	4024
Bel^{EKE}	Deng Lab	-
Bel^{L7470}	Deng Lab	-
Bel^{74407}	Deng Lab	-
Bel^{cap-1}	BDSC	1178
Bel^{47110}	Deng Lab	-
shEIF4B	BDSC	31364
UAS-EIF4B	Todd Lab	-
shEIF4H1 1	VDRC	100817
shEIF4H1 2	VDRC	34301
shEIF4H1 3	VDRC	48119
shEIF4H2 1	VDRC	102825
shEIF4H2 2	VDRC	32192
shEIF1 1	VDRC	29216
shEIF1 2	BDSC	55232
$EIF1^{EY02210}$	BDSC	15406
shEIF1 3	BDSC	57174
shEIF1 4	VDRC	29215
shEIF1 5	VDRC	105763
shEIF1A 1	VDRC	100611
shEIF1A 2	BDSC	29316
shEIF1A 3	BDSC	31185
shEIF1A 4	VDRC	26022
$EIF1A^{645}$	BDSC	23925
$EIF1A^{2232}$	BDSC	23941
$EIF1A^{c04533}$	BDSC	11495
$EIF1A^{EP935}$	BDSC	17203
UAS-EIF1A	FlyORF	F000848

Fly Stock	Source	Catalog #
shEIF2α 1	VDRC	7799
shEIF2α 2	VDRC	104562
shEIF2α 3	VDRC	7798
$EIF2α^{815-29}$	BDSC	4926
UAS-EIF2α	FlyORF	F000983
$EIF3B^{EY14430}$	BDSC	20931
shEIF4A	BDSC	33970
UAS-EIF4A	Xie Lab	-
$EIF4A^{1013}$	BDSC	8647
shEIF5 1	BDSC	34841
shEIF5 2	VDRC	29070
UAS-EIF5	BDSC	22132
$EIF5B^{09143}$	BDSC	11735
$EIF5B^{EY01401}$	BDSC	19641
shRHAU	VDRC	44984
St r	BDSC	20040
shDHX57	BDSC	55373
shBGCN 1	VDRC	108334
shBGCN 2	VDRC	25590
shBGCN 3	BDSC	36636
$BGCN^{KG08129}$	BDSC	14687
shRPS25 1	VDRC	101342
shRPS25 2	VDRC	52602
shSF2 1	BDSC	29522
shSF2 2	BDSC	32367
GMR-GAL4	BDSC	8605
Tub5-GS	Pletcher Lab	-
ElaV-GS	Pletcher Lab	-
CGG_{90}-EGFP	Jin Lab	-
EGFP	BDSC	6874

Appendix Table S1: Fly stocks used in this study and their sources.

Primer Name	Primer Sequence (5’ to 3’)
EGFP (Forward)	TCTTCTTCAAGGACGACGGCAACTAC
EGFP (Reverse)	GTACTCCAGCTTGTGCCCCAGGATGT
Belle (Forward)	CAGTAGCTTGTGGAACGTAAGAAGTTT
Belle (Reverse)	TTACTCATATTATCCTCCAATCAGTTGC
RPL32 Dmel (Forward)	GTTGTGCACCAGGAACTTCTTGAATCCG
RPL32 Dmel (Reverse)	CTTCCAGCTTCAAGATGACCATCCGC
Nanoluciferase (Forward)	GGTGGTGTACCCTGTGGATG
Nanoluciferase (Reverse)	AACCCCGTCGATTACCAGTG
Firefly luciferase (Forward)	GCAGTACCGGATTGCCCAAG
Firefly luciferase (Reverse)	GTCGGGGATGATCTGGTTGC
MALAT (Forward)	TGGTGATGAAGGTAGCAGGC
MALAT (Reverse)	GGCATGCTGGTCTAGGATCC

Appendix Table S2. Primers used in this study for qRT-PCR.