CREB-Binding Protein Modulates Repeat Instability in a *Drosophila* Model for PolyQ Disease

Joonil Jung[1] and Nancy Bonini[1,2]*

Although expansion of trinucleotide repeats accounts for over 30 human diseases, mechanisms of repeat instability remain poorly understood. We show that a *Drosophila* model for the CAG/polyglutamine (polyQ) disease spinocerebellar ataxia type 3 recapitulates key features of human CAG-repeat instability, including large repeat changes and strong expansion bias. Instability is dramatically enhanced by transcription and modulated by nuclear excision repair and a regulator of DNA repair adenosine 3′,5′-monophosphate (cAMP) response element–binding protein (CREB)–binding protein—a histone acetyltransferase (HAT) whose decreased activity contributes to polyQ disease. Pharmacological treatment to normalize acetylation suppressed instability. Thus, toxic consequences of pathogenic polyQ protein may include enhancing repeat instability.

Expansion of repeat sequences causes more than 30 human diseases, including the CAG-repeat polyQ diseases spinocerebellar ataxia type 3 (SCA3) and Huntington's disease (HD), as well as noncoding repeat expansion diseases like the CGG 5′ untranslated region (5′UTR)–repeat disease fragile X syndrome (*1*, *2*). Although the repeats are normally polymorphic in size, disease occurs only when the number of repeats expands beyond a critical threshold characteristic for each disease. Repeat instability also underlines anticipation, as expanded repeats have a strong tendency to further expand, causing the disease to occur earlier with greater severity in successive generations.

To better understand mechanisms of instability, we determined whether *Drosophila* displays intergenerational repeat instability in a CAG-repeat model of human SCA3. Transgenes retaining at least some native genomic context surrounding the CAG repeat manifest instability, as for the Httexon1 transgenic mouse (*3*). The SCA3tr-Q78 transgene in our fly model (*4*) is derived from exon 10 of the human SCA3 gene and contains two polymorphic nucleotides linked to the majority of high-normal and expanded repeat alleles in patients (*5*), which suggests that it may contain genomic elements important for instability.

The rate of repeat instability was initially determined in three independent transgenic lines, following a nine-generation protocol (fig. S1). These studies showed that CAG repeats were stable with only minimal repeat changes (Fig. 1, A and D) (*6*). However, most human trinucleotide-repeat genes, including *SCA3*, are expressed in germ cells, where dramatic instability occurs (*7*). In *Escherichia coli* and mouse, transcription may play a role (*8*). Therefore, we then expressed the SCA3 transgenes in germ cells using a *nos-GAL4* driver line (fig. S2). With germline transcription, all lines now bore a striking percentage of progeny with altered repeat lengths (Fig. 1, B and D). Different transgenic lines showed different degrees of instability, which was not correlated with the level of transgene expression (Fig. 1, D and E). This highlights a potential influence of the genomic context of the transgenic insertion site per se in modulating the degree of instability, as noted in HD and muscular dystrophy models (*9*).

Although all transgenic lines showed instability with transcription (Fig. 1 and fig. S3A), we focused on line 2 for detailed mechanistic studies, confirming all key findings with at

[1]Department of Biology, University of Pennsylvania, Philadelphia, PA 19104, USA. [2]Howard Hughes Medical Institute, University of Pennsylvania, Philadelphia, PA 19104, USA.

*To whom correspondence should be addressed. E-mail: nbonini@sas.upenn.edu

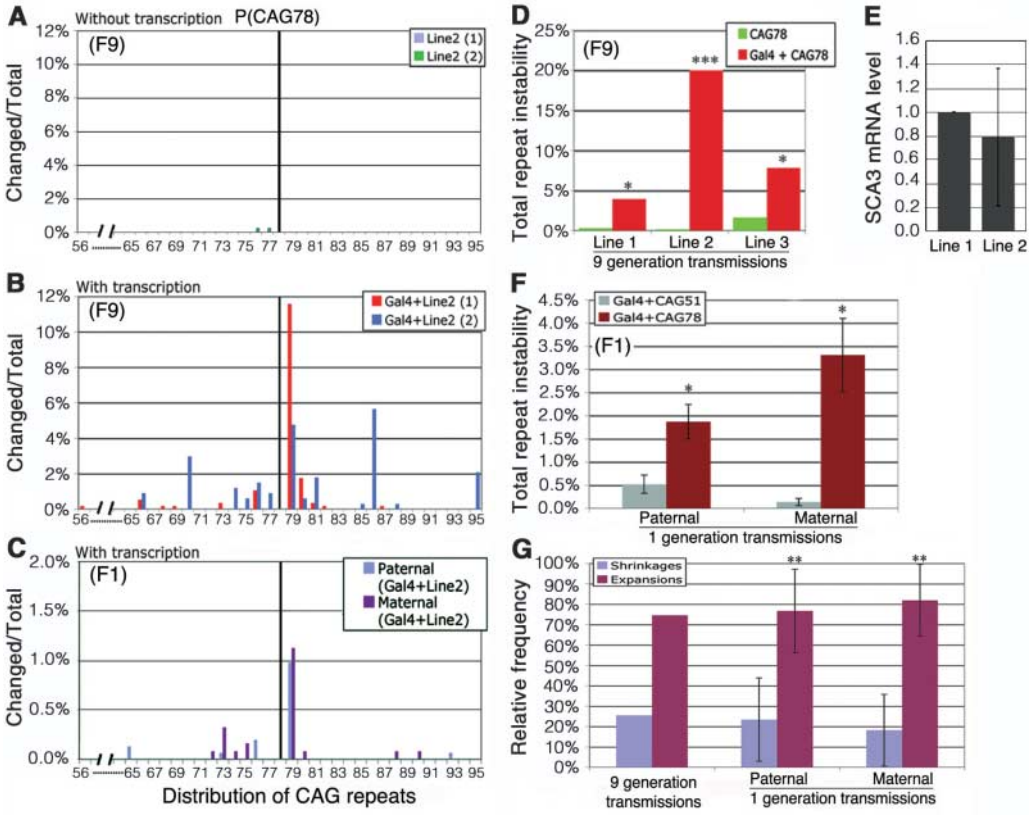

Fig. 1. CAG-repeat instability in an SCA3 model mimics human CAG-repeat instability. (A, B, and D) Germline transcription with *nos-GAL4* (Gal4) significantly enhanced instability in nine-generation studies (chi-square test, $*P < 0.05$, $***P < 0.0001$). (**A** and **B**) CAG-repeat distribution among F_9 progeny of line 2. Parental repeat is labeled P(CAG78). (**C**) CAG-repeat instability in single-generation studies. (**D** and **E**) Similar levels of SCA3 mRNA were present, but repeat instability was significantly different between lines 1 and 2. (chi-square test, $P < 10^{-6}$). (**F**) CAG51 (27 repeat shrinkage of line 2) showed a significantly lower rate of instability than the original CAG78 allele [Student's t test, $*P < 0.05$; mean ± SEM; $n = 3$ (CAG51) or 4 (CAG78)]. (**G**) Expansion bias was observed in repeat instability of line 2 (Student's t test, $**P < 0.01$; mean ± SD; $n = 2$ (F_9), 6 (F_1, paternal transmission), or 5 (F_1, maternal transmissions).

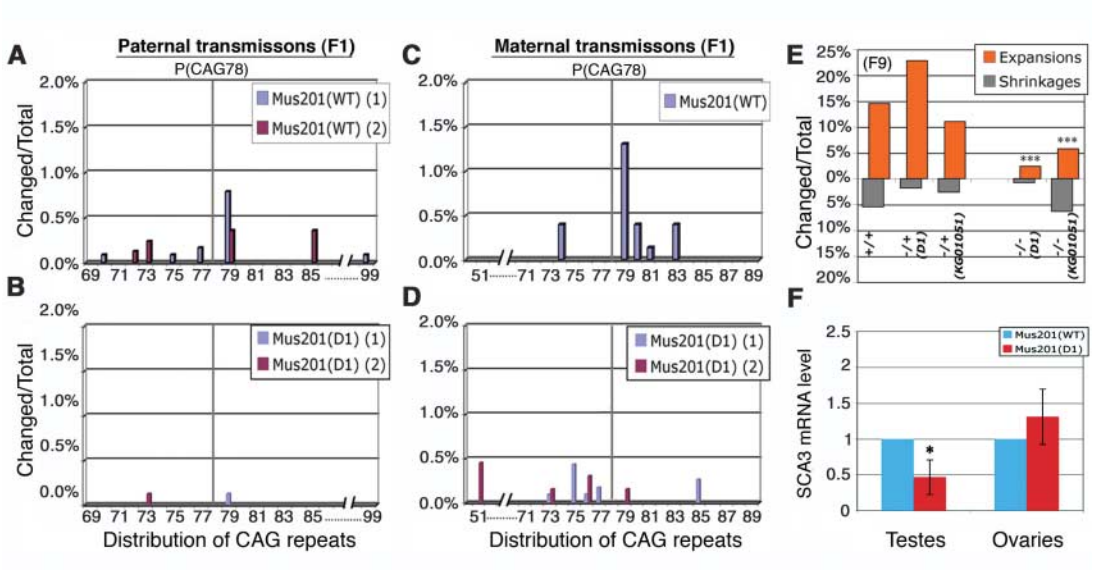

Fig. 2. Suppression of CAG-repeat expansions with loss of NER and/or TCR. (**A** and **B**) Compared with the control, *mus201*^{D1} flies showed a significantly reduced overall rate of instability in paternal transmission (Student's *t* test, $P = 0.01$). (**C** and **D**) In maternal transmission, expansions were selectively reduced with *mus201*^{D1} (Fisher's exact test, $P < 0.001$). Repeat shrinkages were unaffected. (**E**) CAG-repeat expansions were suppressed in nine-generation studies, tested with two independent alleles *mus201*^{D1} and *mus201*^{KG01051} (Fisher's exact test, $^{***}P < 0.001$). (**F**) SCA3 expression level was not affected in ovaries, but was reduced in testes (Student's *t* test, $^{*}P < 0.05$; mean ± SD; $n = 3$).

least one additional line and/or distinct disease transgene. For line 2, whereas small repeat changes involving one to three repeat expansions made up the majority of events (58% of changes), large repeat expansions and shrinkages involving ≥10 CAG repeats were also frequent (10%; Fig. 1B). Moreover, the wide range of repeat size variation (56 to 95 CAG repeats, reflecting shrinkages of 22 repeat units to expansions of 17) was similar to the range of CAG-repeat sizes seen in SCA3 patients (*10*).

In single-generation studies, the repeats also showed little instability without transcription, but strong instability upon germline transcription (Fig. 1C). Of note, changes involving ≥10 repeat expansions or shrinkages were seen. This suggests a mechanism(s) in vivo that generates long repeat changes in single or a few molecular events. Additionally, we observed length-dependent increases in the rate of instability, as are seen in mammals (Fig. 1F) (*11*). We also observed an ~3:1 bias for repeat expansions (Fig. 1G). This is similar to what is observed in humans, although expansion bias is primarily observed in paternal transmissions. In mammals, the prolonged period of meiotic arrest during oogenesis is thought to play a role in maternal repeat shrinkages (*12*, *13*); lack of such an arrest during *Drosophila* oogenesis may explain the expansion bias in maternal transmission. Thus, the SCA3 fly model with germline transcription displays several key features of CAG-repeat instability that are seen in human SCA3 patients.

Expanded trinucleotide repeats form unusual DNA structures in vitro, such as hairpins and slipped strands (*14*). These may attract DNA repair proteins and trigger repair responses leading to repeat instability in vivo (*15*, *16*). The strong relation between repeat instability and transcription in the fly raised the possibility that DNA repair mechanisms—transcription-coupled repair (TCR) in particular—may be involved (*17*). The fly Mus201 protein is an ortholog of human Rad2/XPG, and loss of Mus201 function leads to complete loss of nucleotide excision repair (NER) and/or TCR in flies. We therefore examined the effect of eliminating TCR on repeat instability.

In the *mus201*^{D1} (a null mutant) background, we observed a dramatic reduction in instability. Control flies showed the typical 1 to 3% instability with strong bias for repeat expansions (Fig. 2, A and C). However, with loss of Mus201 function, repeat expansions were suppressed in both maternal and paternal transmissions (Fig. 2, B and D), as well as in nine-generation studies (Fig. 2E). SCA3 mRNA levels were reduced in the testes, which suggested a dependence on the TCR pathway for transgene expression, but were not reduced in ovaries (Fig. 2F). This indicates a specific effect

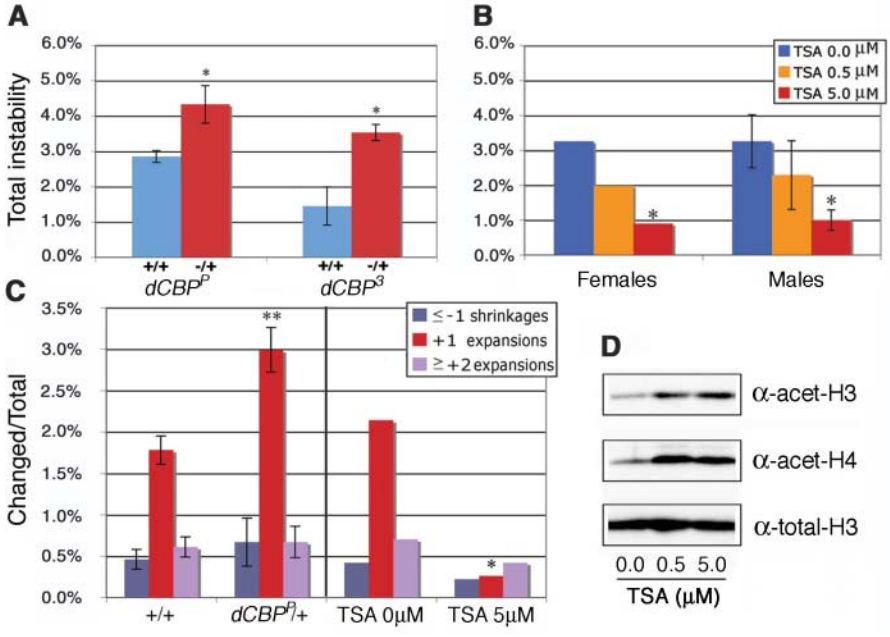

Fig. 3. dCBP and the HDAC inhibitor TSA modulate repeat instability. (**A**) Significant increase in repeat instability was observed in *dCBP* mutants (Student's *t* test, $^{*}P < 0.05$; $n = 3$). Fly genotypes: *dCBP*^{P}/*FM7c* with CAG78/CAG78 (left) and *dCBP*^{3}/+ with CAG78/CAG46 (right). (**B**) Treatment of flies bearing CAG78/CAG78 with TSA suppressed instability [Student's *t* test, $^{*}P < 0.05$; $n = 2$ (females) or 3 (males)]. (**C**) Decrease in dCBP (*dCBP*^{P}/+) preferentially increased +1 expansions (two-way ANOVA Bonferroni post test, $^{**}P < 0.01$), whereas TSA treatment (5 µM) had the opposite effect (Fisher's exact test, $^{*}P < 0.05$). All values, means ± SEM. (**D**) TSA-treated flies had increased levels of acetylated histones H3 and H4.

of Mus201 activity on repeat instability in females. Effects of NER have been reported in *Escherichia coli* and cell lines with conflicting results; our findings indicate that Mus201/TCR modulates CAG-repeat expansions during germline transmissions in vivo. This potentially contributes to the genetic "anticipation" phenomenon prevalent in human patients.

In polyQ diseases, expanded CAG repeats are translated into pathogenic polyQ tracts to form nuclear inclusions. Key regulators of DNA repair, such as adenosine 3′,5′-monophosphate (cAMP) response element–binding protein (CREB)–binding protein (CBP) (*18*, *19*), are found in polyQ inclusions, which leads to a reduction in soluble CBP levels (*20*, *21*). Soluble pathogenic polyQ protein may also increase degradation of CBP (*22*), as well as inhibit CBP/p300 histone acetyltransferase (HAT) activity through direct binding (*23*). We therefore tested whether decreased CBP levels could alter DNA repair activity in flies. Decreasing *Drosophila* CBP (*dCBP*) gene dosage significantly modified the survival rate of methyl methanesulfonate–treated larvae (a measure of DNA repair) (fig. S4). This raised the possibility that pathogenic polyQ protein may indirectly enhance CAG-repeat instability in polyQ disease by inhibiting CBP functions important for DNA repair.

To test this, we determined whether loss of dCBP enhanced repeat instability. Because dCBP is essential, heterozygous flies were tested. Such partial loss of CBP function may reflect well the typical prolonged disease situation, as severe CBP depletion causes cytotoxic cell loss. We examined two different alleles of *dCBP*: $dCBP^{P}$, a hypomorphic allele, and $dCBP^{3}$, a null microdeletion. Both alleles enhance polyQ toxicity in heterozygous animals in *Drosophila* (*24*). The effects of the $dCBP^{P}$ allele were examined on CAG78/CAG78 flies; however, the $dCBP^{3}$ allele was deleterious in that background, so instability was examined with CAG78 in trans to a less toxic CAG46 allele (CAG78/CAG46). CAG78/CAG78 control flies showed an instability rate of 2.85 ± 0.17%; the CAG78/CAG46 flies showed a reduced rate, with the CAG78 repeat showing 1.45 ± 0.54% instability (Fig. 3A). In both situations, however, a significant increase in instability of the CAG78 repeat was observed [4.33 ± 0.53% ($dCBP^{P}$) and 3.35 ± 0.23% ($dCBP^{3}$)] (Fig. 3A). Similar increases in instability were confirmed in independent transgenic lines (fig. S3A). Although we did not observe an obvious change in the overall pattern of histone H3 or H4 acetylation in ovaries of mutant flies, localized decreases in acetylation of histones or other protein targets may occur. Significant changes in SCA3 transgene expression were not observed (fig. S5A). We further tested the role of CBP on repeat instability by up-regulating *dCBP* in males using an overexpression allele $dCBP^{EP1179}$; this resulted in a reduced rate of instability (fig. S3B).

The inhibitory effect of pathogenic polyQ proteins on CBP is thought to be due to loss of its HAT activity. Treating flies or mouse models of polyQ disease with histone deacetylase (HDAC) inhibitors to normalize acetylation levels protects against polyQ protein pathogenesis (*25*). Our data indicated that HDAC inhibitors may also protect against the increased rate of repeat instability. We therefore treated flies expressing the pathogenic polyQ protein in germ cells with an HDAC inhibitor trichostatin A (TSA), and measured repeat instability.

Placebo-treated flies showed normal instability [3.26 ± 0.29% and 3.26 ± 0.76% for maternal and paternal transmissions, respectively] (Fig. 3B). In contrast, there was a sharp drop in rate of instability of flies fed TSA. With 5 μM TSA, CAG-repeat instability decreased to 0.90 ± 0.02% and 1.00 ± 0.29% for maternal and paternal transmissions, respectively (Fig. 3B). TSA treatment increased the level of acetylated histones in flies (Fig. 3D), and SCA3 transcription was not affected (fig. S5B). Although TSA is known to affect multiple pathways in vivo, analysis of the pattern of repeat changes by TSA suggested that reduced instability was, at least in part, due to compensation for decreased dCBP and/or HAT activity, as both *dCBP* gene changes and TSA treatment preferentially modulated +1 repeat changes relative to other events (Fig. 3C).

Several lines of evidence suggested that the features we characterized for repeat instability of SCA3 transgenes may have broader implications. For example, the majority of triplet-repeat disease genes are expressed in germ cells, and a number of polyQ proteins (Huntingtin, androgen receptor, ataxin-7, and ataxin-3) interact with or inhibit CBP or other HAT proteins (*26*). Therefore, to extend these findings, we examined the effects of transcription and the contribution of CBP to HD models; we also examined instability in a model for fragile X CGG premutation expansions, an unstable noncoding trinucleotide repeat (*27*). As with SCA3, we found a consistent enhancement of repeat instability with germline transcription for both HD and CGG models (fig. S6). We also tested the role of CBP on HD. As with SCA3tr-Q78, reducing the level of *dCBP* ($dCBP^{P}$/+) enhanced repeat instability of an HD Httexon1Q93 transgene (fig. S7). Thus, features of repeat instability, including transcriptional dependence, may be fundamental aspects of trinucleotide instability displayed in multiple disease models.

These studies reveal that *Drosophila* recapitulates several central features of human CAG-repeat instability, including the wide range of repeat changes and strong bias for expansions. Moreover, germline transcription along with NER and/or TCR dramatically enhanced instability. A consequence of polyQ gene expression in germ cells is that polyQ protein pathology—namely, inhibition of CBP activity via sequestration or inhibitory interactions—may further enhance repeat instability. In HD, dramatic expansion of CAG repeats occurs not only intergenerationally, but also somatically, in affected brain regions, which develop striking inclusions containing CBP (*28*, *29*). Known modifiers like Msh2 modulate CAG-repeat instability similarly in both the germline and somatic tissues, which suggests that our findings with germline instability may be applicable to somatic instability.

Trinucleotide-repeat instability has been viewed largely as a matter of DNA metabolism; however, our data suggest that repeat instability may be influenced by aspects of polyQ protein toxicity; thus, treatments to curb polyQ protein pathology may also be effective means to help clamp repeat instability. Among these, HDAC inhibitors are in clinical trials (*30*).

References and Notes

1. J. R. Gatchel, H. Y. Zoghbi, *Nat. Rev. Genet.* **6**, 743 (2005).
2. C. E. Pearson, K. N. Edamura, J. D. Cleary, *Nat. Rev. Genet.* **6**, 729 (2005).
3. L. Mangiarini *et al.*, *Nat. Genet.* **15**, 197 (1997).
4. J. M. Warrick *et al.*, *Cell* **93**, 939 (1998).
5. P. Maciel *et al.*, *Eur. J. Hum. Genet.* **7**, 147 (1999).
6. S. M. Jackson *et al.*, *Gene* **347**, 35 (2005).
7. Y. Kawaguchi *et al.*, *Nat. Genet.* **8**, 221 (1994).
8. F. Tanaka *et al.*, *Am. J. Hum. Genet.* **65**, 966 (1999).
9. Y. Zhang, D. G. Monckton, M. J. Siciliano, T. H. Connor, M. L. Meistrich, *Hum. Mol. Genet.* **11**, 791 (2002).
10. Y. Takiyama *et al.*, *Hum. Mol. Genet.* **6**, 1063 (1997).
11. V. C. Wheeler *et al.*, *Hum. Mol. Genet.* **8**, 115 (1999).
12. T. Sato *et al.*, *Hum. Mol. Genet.* **8**, 99 (1999).
13. M. D. Kaytor, E. N. Burright, L. A. Duvick, H. Y. Zoghbi, H. T. Orr, *Hum. Mol. Genet.* **6**, 2135 (1997).
14. R. D. Wells, R. Dere, M. L. Hebert, M. Napierala, L. S. Son, *Nucleic Acids Res.* **33**, 3785 (2005).
15. K. Manley, T. L. Shirley, L. Flaherty, A. Messer, *Nat. Genet.* **23**, 471 (1999).
16. C. Savouret *et al.*, *EMBO J.* **22**, 2264 (2003).
17. Y. Lin, V. Dion, J. H. Wilson, *Nat. Struct. Mol. Biol.* **13**, 179 (2006).
18. S. Hasan, M. O. Hottiger, *J. Mol. Med.* **80**, 463 (2002).
19. M. Tini *et al.*, *Mol. Cell* **9**, 265 (2002).
20. F. C. Nucifora Jr. *et al.*, *Science* **291**, 2423 (2001).
21. Y. Chai, J. Shao, V. M. Miller, A. Williams, H. L. Paulson, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9310 (2002).
22. H. Jiang, F. C. Nucifora Jr., C. A. Ross, D. B. DeFranco, *Hum. Mol. Genet.* **12**, 1 (2003).
23. G. Schaffar *et al.*, *Mol. Cell* **15**, 95 (2004).
24. J. P. Taylor *et al.*, *Genes Dev.* **17**, 1463 (2003).
25. J. S. Steffan *et al.*, *Nature* **413**, 739 (2001).
26. L. Bodai, J. Pallos, L. M. Thompson, J. L. Marsh, *Curr. Med. Chem.* **10**, 2577 (2003).
27. P. Jin *et al.*, *Neuron* **39**, 739 (2003).
28. L. Kennedy *et al.*, *Hum. Mol. Genet.* **12**, 3359 (2003).
29. M. DiFiglia *et al.*, *Science* **277**, 1990 (1997).
30. J. Marx, *Science* **310**, 43 (2005).

31. We thank R. Pittman, H. Paulson, A. Cashmore, D. Lessing, L. Li, M. Watson for comments; S. DiNardo, L. Thompson, L. Goldstein, and P. Jin for reagents; and K. Xu and X. Teng for technical assistance. Supported by grants from the Hereditary Disease Foundation (to J.J.), the Packard Foundation, and NINDS (to N.M.B.). N.M.B. is an investigator of the Howard Hughes Medical Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1139517/DC1
Materials and Methods
Figs. S1 to S7
References

4 January 2007; accepted 9 February 2007
Published online 1 March 2007;
10.1126/science.1139517
Include this information when citing this paper.

CREB-Binding Protein Modulates Repeat Instability in a *Drosophila* Model for PolyQ Disease

Joonil Jung and Nancy Bonini (March 1, 2007)
Science **315** (5820), 1857-1859. [doi: 10.1126/science.1139517]
originally published online March 1, 2007

Editor's Summary