11714–11727 Nucleic Acids Research, 2012, Vol. 40, No. 22 Published online 15 October 2012
doi:10.1093/nar/gks943

The novel long non-coding RNA *CRG* regulates *Drosophila* locomotor behavior

Meixia Li[1], Shengyun Wen[1,2], Xiangqian Guo[3], Baoyan Bai[3], Zhefeng Gong[1], Xiaojun Liu[4], Yijin Wang[1,2], Yanqiong Zhou[1,2], Xiaowei Chen[2,3], Li Liu[1,5,*] and Runsheng Chen[3,*]

[1]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, [2]Graduate University of the Chinese Academy of Sciences, Beijing 100039, [3]Key Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, [4]National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100005 and [5]Key Laboratory of Mental Health, Chinese Academy of Sciences, Beijing 100101, China

Received April 12, 2012; Revised September 14, 2012; Accepted September 18, 2012

ABSTRACT

Long non-coding RNAs (lncRNAs) that have no protein-coding capacity make up a large proportion of the transcriptome of various species. Many lncRNAs are expressed within the animal central nervous system in spatial- and temporal-specific patterns, indicating that lncRNAs play important roles in cellular processes, neural development, and even in cognitive and behavioral processes. However, relatively little is known about their *in vivo* functions and underlying molecular mechanisms in the nervous system. Here, we report a neural-specific *Drosophila* lncRNA, *CASK regulatory gene* (*CRG*), which participates in locomotor activity and climbing ability by positively regulating its neighboring gene *CASK* (Ca^{2+}/calmodulin-dependent protein kinase). *CRG* deficiency led to reduced locomotor activity and a defective climbing ability—phenotypes that are often seen in *CASK* mutant. *CRG* mutant also showed reduced *CASK* expression level while *CASK* over-expression could rescue the *CRG* mutant phenotypes in reciprocal. At the molecular level, *CRG* was required for the recruitment of RNA polymerase II to the *CASK* promoter regions, which in turn enhanced *CASK* expression. Our work has revealed new functional roles of lncRNAs and has provided insights to explore the pathogenesis of neurological diseases associated with movement disorders.

INTRODUCTION

Genome-wide transcriptional analyses have identified large numbers of non-protein-coding RNAs in humans and animals (1–4). Depending on their length, non-coding RNAs (ncRNAs) can be arbitrarily divided into small ncRNAs and long ncRNAs (lncRNAs). A lncRNA usually refers to a RNA transcript of longer than 200 nt without a potential protein-coding ORF; many lncRNAs show properties similar to those of mRNAs: both are transcribed by RNA polymerase II, spliced and polyadenylated.

Although the functions of most lncRNAs are still unknown, they are emerging as important regulatory factors in molecular genetic and cellular processes. lncRNAs could function as transcriptional regulators of neighboring protein-coding genes by *cis*- or *trans*-modulation (5–7). Recently, lncRNAs were found to play enhancer-like roles on the expression of nearby protein-coding genes (8). In *Drosophila melanogaster*, it has been shown that transcriptional elongation of *bithoraxoid* (*bxd*) ncRNA, the first described *Drosophila* lncRNA, can repress the expression of *Ultrabithorax* (*Ubx*) (9,10). lncRNAs have also been implicated in epigenetic gene regulation through histone or DNA modification (11–14). *Drosophila roX1* and *roX2* are functional lncRNAs that are essential for dosage compensation by hyper-activating transcription of the X chromosome in males (15). In addition, lncRNAs could act as precursors for small RNAs (16).

lncRNAs also function in the nervous system, from regulating neural development to mediating behavioral

*To whom correspondence should be addressed. Tel: +86 10 64888543; Fax: +86 10 64889892; Email: crs@sun5.ibp.ac.cn
Correspondence may also be addressed to Li Liu. Tel: +86 10 64888550; Fax: +86 10 64853625; Email: liuli@sun5.ibp.ac.cn

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

© The Author(s) 2012. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by-nc/3.0/), which permits non-commercial reuse, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com.

and cognitive processes (17–20). In *Drosophila*, transcriptomic studies have revealed large numbers of non-coding transcripts, including lncRNAs, in the different developmental stages, and many of them displayed specific expression in central or peripheral nervous system (4,21). Such preferential expression in the nervous system indicates that these lncRNAs are involved in neural development and function. For example, the lncRNA *bereft*, which is expressed specifically in the *Drosophila* peripheral nervous system, is essential for extrasensory organ development and the maintenance of the interommatidial bristles of the eye (22). Recently, the lncRNA *yellow-achaete intergenic RNA* (*yar*), which is located within a neural gene cluster, was found to be involved in *Drosophila* sleep regulation (19). These results indicate that lncRNAs play important roles in neural development and behavioral processes.

Ca^{2+}/calmodulin-dependent protein kinase (*CASK*) belongs to the membrane-associated guanylate kinase (MAGUK) protein family, the members of which are characterized by the presence of PDZ, SH3 and guanylate kinase domains at their C terminus (23). *CASK* is a scaffolding protein, and its binding to partner proteins is involved in synapse formation and plasticity, gene expression and neural development (24–26). In *Drosophila*, *CASK* (known as *cmg* or *caki*) was discovered because of its CaM-kinase domain at the N terminus (27). *CASK* is enriched in the nervous system during development from the embryonic stage to the adult (27). The newly released annotation of the *Drosophila* genome (version 5.43) suggested that the *CASK* transcriptional unit is ~40 kb long. The protein exists mainly in two isoforms: the longer includes the N-terminal CaM-kinase-like and L27 domains whereas the shorter does not include such domains. Limited mutation of the longer isoform of *CASK* caused both initiation and motor maintenance defects in *Drosophila*, implicating its involvement in movement disorders (28). *CASK* was found to be associated with both pre- and post-synaptic membranes (25). Its interaction with *Drosophila* neurexin, which is involved in synaptic transmission, might underlie the molecular mechanism of the locomotor defect (29).

In this study, we demonstrated that a novel lncRNA, *CRG* (*CASK regulatory gene*), which is expressed in the nervous system, is involved in the regulation of *Drosophila* locomotor behavior. The behavioral defects in *CRG* mutant could be rescued by over-expression of the adjacent movement-related protein-coding gene *CASK*. Furthermore, *CRG* could recruit RNA polymerase II to the *CASK* promoter regions, thereby increasing *CASK* expression. Our results provided another functional mode of lncRNA on the nearby protein-coding gene which was implicated in behavior regulation and further enriched the biological significance of lncRNA.

MATERIALS AND METHODS

Fly stocks

All flies were raised at 25°C on standard corn meal/molasses medium in a 12 h light/12 h dark cycle at 60% humidity (30). The following fly strains were used: wild-type (WT) *Canton-S* (*CS*), w^{1118}, *elav*-Gal4, *tub*-Gal4, mb247-Gal4, G7-Gal4, OK6-Gal4, 6793-Gal4, UAS-*CASK* (24), X-307, X-313, $CRG^{\Delta 1877}$ and UAS-*CRG*.

Coding potential prediction

MAF files for *CRG* locus in 12 *Drosophila* species were downloaded from Galaxy, followed by calculating the potential of non-coding RNAs with PhyloCSF (31,32). The PhyloCSF score is a log-likelihood ratio, with a positive score representing that the alignment is more likely to occur in a coding region than in a non-coding one, and negative score otherwise.

Generation of $CRG^{\Delta 1877}$

The ends-in gene targeting method (33) was used to generate a *CRG* deficiency line. The donor transgenic flies were constructed as follows: a 3 kb genomic fragment upstream and a 5 kb genomic fragment downstream of the target region, and the DNA fragment used to introduce the I-SceI cleavage sequence, were sequentially cloned into the pTARG vector to generate the gene targeting construct pTARG-*CRG*; then, the recombinant plasmid was micro-injected into w^{1118} embryos to generate the donor transgenic line.

The primers used to amplify the upstream 3 kb fragment were:

5′-TATTAACGCGTCTTCAGAGGACCGGTTATGCCG-3′,
5′-GCGCCAGATCTTACCGAATTTTAAATACATAAG-3′.

The primers used to amplifying the downstream 5 kb fragment were:

5′-CCGGCTAGCTGTGTTGTGTATACATATTTTCT-3′,
5′-TTAGGCGCGCCGGTTTTGGATGCCCTGTTTGG-3′.

The oligos used to introduce the I-SceI cleavage sequence were:

5′-CTAGTAGGGATAACAGGGTAAT-3′,
5′-CTAGATTACCCTGTTATCCCTA-3′.

Donor transgenic flies carrying the targeting construct on the second chromosome were crossed to flies containing *hs-I-SceI* and *hs-FLP* transgenes. One hour of heat-shock treatment (38°C) was applied on days 2 and 3 after egg laying. Heat-shocked virgin female flies with mosaic eyes were singly crossed with *yw;ey-FLP;MKRS/TM2,y*$^{+}$ males. The w^{+} offspring were selected and crossed with w^{1118}*;hs-I-CreI,Sb/TM6*. One hour of heat-shock treatment (36°C) was applied at the second instar larval stage to remove the WT copy with the 1877 bp *CRG* segment. The w^{-} offspring males were crossed individually to *yw;sp/CyO;MKRS/TM2,y*$^{+}$ to establish stocks.

The *CRG* deficiency was verified by DNA sequencing. The PCR primers used to confirm fragment deletion were:

5′-GTCTTTGCAGCATGCATGATTCA-3′
5′-TAGACACAAGGAGACACAACAGC-3′.

Generation of UAS-*CRG* flies

The WT cDNA fragment corresponding to the full sequence of *CRG* was obtained by RT-PCR using the primers,

5′-GCTGCAGATCTTTTTATTCAGCATTTAGTC-3′
5′-GCACGCTCGAGTTTCAGTGTTTACTCGTTT-3′.

The product was sequence-confirmed and sub-cloned into the pUAST vector (containing the GFP sequence). The recombinant plasmid was germline transformed to generate the UAS-*CRG* transgenic line. The validity of the UAS-*CRG* line was confirmed (Supplementary Figure S10).

In situ hybridization

A 228 bp cDNA fragment of *CRG* was amplified from the total RNA of WT adult head by RT-PCR with the 5′ primer 5′-AAGGAGACCCAAAACCGAAT-3′ and 3′ primer 5′-AATACTGAGCGCTGGTTGCT-3′.

A 265 bp DNA fragment of GFP was amplified from a recombinant pUAST vector containing the GFP sequence by PCR with the 5′ primer 5′-CACATGAAGCAGCACG ACTT-3′ and 3′ primer 5′-AGTTCACCTTGATGCCGT TC-3′.

Each fragment was cloned into a pGEM-T vector (Promega). Anti-sense and sense RNA probes were prepared with SP6 and T7 RNA polymerase, respectively, using a DIG RNA Labeling Kit (Roche) according to the manufacturer's instructions.

Embryos were prepared according to protocols described by Kosman *et al*. (34). Larval brains at the late third instar stage were dissected and transferred to 4% paraformaldehyde (PFA) in PBS and fixed for 45 min at room temperature. Adult brains were isolated and transferred to 8% formaldehyde (FA) in PBS and fixed for 1 h at room temperature. The prepared samples were then processed following standard procedures for whole mount *in situ* hybridization, except that a hybridization temperature of 60°C was used (21,35). All images were acquired with a Leica DM2500 microscope in DIC mode.

5′ and 3′ RACE

5′ and 3′ RACE were performed with 5′-Full RACE and 3′-Full RACE Core Set kits (Takara) according to the manufacturer's instructions.

The primers used in 5′ RACE for *CRG* were:

outer primer: 5′-CGAAACTGAATATCTCCCATGT-3′,
inner primer: 5′-GCGACTTAGGATCTTATTCG-3′.

The primers used in 5′ RACE for *CASK* were:

outer primer: 5′-GATGCCTCGACTCAATTGCT-3′,
inner primer: 5′-CGAATATTGATGAGTTGTGTCCT-3′.

The primers used in 3′ RACE for *CRG* were:

outer primer: 5′-AGGATGAGGAGATTGCAGGT-3′,
inner primer: 5′-CGCAAAACAGGCAGATAAACGT-3′.

The primers used in 3′ RACE for *CASK* were:

outer primer: 5′-CAGATGCGCACGACGATGCAGAT AC-3′,
inner primer: 5′-CCACTACGAGAGATGACAGT-3′.

The PCR products were purified and cloned into a pGEM-T vector (Promega) for sequencing.

Northern blotting

The specific fragment used for probe synthesis was amplified by RT-PCR from WT adult head cDNA. The primers used in the reaction were:

upper: 5′-AGCAACCAGCGCTCAGTATT-3′,
lower: 5′-AGCTCAAATCGCGTCAATCT-3′.

The 1.3 kb fragment was cloned into a pGEM-T vector to generate recombinant pGEM-T plasmids for sequencing. The probe was then synthesized and labeled by *in vitro* transcription of the recombinant pGEM-T plasmids with T7 RNA polymerase and DIG RNA Labeling Kit (Roche). Northern blotting was performed using total RNA isolated from the adult head and from the whole fly with DIG Wash and Block Buffer Set kit and CDP-Star kit (Roche) according to the manufacturer's instructions, except that blots were hybridized in ULTRAhyb (Ambion) at 68°C overnight. Chemiluminescent signals were visualized using X-ray film.

Western blotting

Proteins from the adult head were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 10% polyacrylamide) and electrophoretically transferred to polyvinylidene difluoride membranes. Membranes were blocked in phosphate-buffered saline containing 5% non-fat milk powder and 0.1% Tween 20 (PBST) for 1 h at room temperature, and then incubated with the indicated primary antibodies (rabbit anti-CASK serum, 1:1000) overnight at 4°C, washed three times in PBST (10 min per time), and incubated with goat anti-rabbit horseradish peroxidase conjugated IgG (1:500, Zhongshanjinqiao) for 1 h at room temperature. Membranes were subjected to three additional PBST washes. Proteins were visualized by enhanced chemiluminescence (Pierce) and quantified by densitometry (Bandscan, Glyko).

RT-PCR and quantitative RT-PCR

RT-PCR was carried out to confirm that *CRG* and *GFP* were co-transcribed in transgenic UAS-*CRG* flies. One PCR primer corresponding to *CRG* (5′-GGCCGCGGT GTTAGTAAATA-3′) and one corresponding to *GFP* (5′-ACTGGGTGCTCAGGTAGTGG-3′) were used.

Quantitative RT-PCR was carried out to quantify the RNA levels of *CRG* and *CASK* following a protocol described previously with minor modifications (36). Relative differences in specific gene expression levels were quantified using the equation: expression level = $2^{\Delta Ct}$, where ΔCt is the Ct value of the gene of interest subtracted from the Ct value of *actin*. The primers used for qPCR were:

qPCR-*CRG* up: 5′-TATCTGCCTGTTTTGCGTTG-3′;
qPCR-*CRG* low: 5′-TAATTGCAGCCGCGTTTAAT-3′;

qPCR-*CASK* up: 5′-GCGGCACATTAAGAACACC T-3′;
qPCR-*CASK* low: 5′-ACGACATAGGGCGTGAACT C-3′.

qPCR-*actin* up: 5′-CAGGCGGTGCTTTCTCTCTA-3′;
qPCR-*actin* low: 5′-AGCTGTAACCGCGCTCAGTA-3′;

Luciferase assay

The potential *CASK* promoter regions from the *CASK* transcription start site, 2 kb-up (−2018/−1), 1 kb-up1 (−1000/−1), 1 kb-up2 (−2018/−1000), 0.5 kb-up (−500/−1), 0.2 kb-up (−200/−1), 2 kb-down (−190/+1891), 1 kb-down1 (+1/+1000), 1 kb-down2 (+1000/+1891), 0.5 kb-down (+1/+500) and 0.2 kb-down (+1/+200) stream were amplified from *Drosophila* genomic DNA and sub-cloned into empty pGL3-Basic luciferase reporter plasmids to generate 10 recombinant reporter plasmids. The primers used to amplify these fragments were:

2 kb-up
5′-CAATAACGCGTCACCTGAGTGGCTGAG CCTT-3′
5′-CAATACTCGAGTGATTGGCGAAGAACTGG TTG-3′,
1 kb-up1
5′-GGTGGGGTACCGCTGTGGGGTAATAAT AATA-3′
5′-GTTAGGCTAGCTGATTGGCGAAGAACTGG TTG-3′,
1 kb-up2
5′-ATATAACGCGTCACCTGAGTGGCTGAG CCTT-3′
5′-AGGCGCTCGAGTATTATTATTACCCCA CAGC-3′,
0.5 kb-up
5′-GGTGGGGTACCAAGGAGCAGGAAAATA TAAT-3′
5′-GTTAGGCTAGCTGATTGGCGAAGAACTGG TTG-3′,
0.2 kb-up
5′-GGAGAGGTACCTAACGAAGTAAGTTGC TGCA-3′
5′-GTTAGGCTAGCTGATTGGCGAAGAACTGG TTG-3′,
2 kb-down
5′-GATTAGGTACCAGTTGCTGCAGTACCACC CTT-3′
5′-GGTCGGCTAGCATATTGATGAGTTGTGTC CTG-3′,
1 kb-down1
5′-ATTATGGTACCATTCCTCACGGGCCGTGTT TG-3′
5′-ATTATGCTAGCGTTCGCGGGGAAAGTTCG CAG-3′,
1 kb-down2
5′-ATATAACGCGTCTGCGAACTTTCCCCGCG AAC-3′
5′-AGGCGCTCGAGATATTGATGAGTTGTGTC CTG-3′.
0.5 kb-down
5′-ATTATGGTACCATTCCTCACGGGCCGTGTT TG-3′
5′-GATGAGCTAGCTGCCGCTGTTGTAAAA TGCC-3′,
0.2 kb-down
5′-ATTATGGTACCATTCCTCACGGGCCGTGTT TG-3′
5′-GGAGTGCTAGCCACGAACAATTCGCAT ATAG-3′,

The full *CRG* cDNA sequence was amplified from adult WT flies by RT-PCR and sub-cloned into expression plasmid pcDNA3.1 to generate an expression plasmid for *CRG*.

HeLa cells were grown at 37°C, 5% CO_2 in Dulbecco's modified Eagle's medium (DMEM, Invitrogen) supplemented with 10% fetal bovine serum, 1% (v/v) penicillin-streptomycin. Cell transfection was performed with cells plated on 24-well plates using a VigoFect kit (Vigorous) according to the manufacturer's instructions. 200 ng reporter plasmids, with or without a 200 ng expression plasmid, were co-transfected with 10 ng of the internal control plasmid pRL-TK. Luciferase activity was measured with Dual-Luciferase Reporter Assay System (Promega) 24 h after the transfectants were incubated. All experiments were performed in triplicate.

Chromatin immunoprecipitation

The chromatin immunoprecipitation CHIP assay was performed using EZ Chip Kit (Millipore). 200 adult heads from each line were treated with FA to crosslink the proteins to the genomic DNA. The tissue was then homogenized in lysis buffer to obtain a cell lysate that was subsequently sonicated to shear the genomic DNA into lengths of between 200 and 1000 bp. The sheared supernatant was immunoprecipitated using RNA polymerase II antibody (Abcam); normal mouse IgG (Millipore) was used as a negative control. The specific protein–DNA crosslinks were reversed at 65°C, and the DNA was purified to remove chromatin proteins and quantified by real-time quantitative PCR. The primers used for qPCR were:

CHIP up1
5′-TTCCTTCTGCTCTGCTTACC-3′
5′-CCTTCCTCATTCTTAAGTGG-3′
CHIP up2
5′-GCTGGCTGATACAAGTAGGT-3′
5′-AGCACCAACAACCATTTCGA-3′

CHIP up3
5′- GAAAGGTGGACATTAGTGGG -3′
5′- AACCGTAGGTGCGTAATCTT -3′
CHIP up4
5′-GTTGTTGTCGTTGGGCGTTT-3′
5′-CGCGCCATTCAACTTTATCT-3′
CHIP up5
5′-GGCGCGAAAGAGAAGTAATG-3′
5′-GCTTGAGTGAAGTCAAATGG-3′
CHIP up6
5′-ACTTGCTTGCGTACCGCATA-3′
5′-GTTTTGGGGACGAGAAAGTA-3′
CHIP up7
5′-CTGTGTCGTTATGAATGACT-3′
5′-TGGATTGGTTCGGAACTTGG-3′
CHIP down1
5′-CAATCGAGTGCTAATGCGAT-3′
5′-TGTGCATTTGGCCCTGCTTT-3′
CHIP down2
5′-TTGCCAATTGTGCGACACTC-3′
5′-ACGAAATGTGGAGAACTGAG-3′
CHIP down3
5′-ACTACCAGAAAGTCCTTCGC-3′
5′-AGGCGCTACCTTTGTCTT-3′
CHIP down4
5′-CGAACATCCTTGGAACGAGA-3′
5′-ATGGTGTCGCAACGGAAATG-3′
CHIP down5
5′-TCCTTCCTACCATTGGCACT-3′
5′-GCCACTATAAAGTTGGGCTG-3′
CHIP down6
5′-GTGTTCCCATCACACAAGTG-3′
5′-AGTTGAAGGCCGCTAATCAG-3′
CHIP down7
5′-GCGTCATCTGCTTGACAGAA-3′
5′-CTTTGAGCGTTTTCCTTCGG-3′

RNA immunoprecipitation

The RNA immunoprecipitation (RIP) assay was carried out using EZ-Magna kit (Millipore). One hundred WT flies were collected and washed three times with ice-cold PBS, then homogenized in ice-cold PBS until a single-cell suspension was obtained. The cells were collected by centrifugation and resuspended in RIP lysis buffer, incubated on ice for 5 min and stored at −80°C overnight. Magnetic beads were preincubated with 5 µg RNA polymerase II (Abcam) or with 5 µg normal mouse IgG (Millipore); the latter was used as a negative control. The frozen RIP lysate was thawed and centrifuged. The supernatant was incubated with the magnetic beads–antibody complex at 4°C overnight. The immunoprecipitates were digested with protease K to remove the proteins and release the RNAs. The RNAs were then purified and reverse-transcribed into cDNA using random hexamers. Using these cDNA as template, PCR was performed with the following primers:

CRG RIP1
5′-CAGCATTTAGTCAAAACAGC-3′
5′-GTCACCACTGTCTTCAGTTG-3′
CRG RIP2
5′-GTTTCAAGCCAAGTAAGAAC-3′
5′-GGCTTACTCATTTCTCATGG-3′
CRG RIP3
5′-GGAAAAAACATAGAGTGAGC-3′
5′-GTCTCCTTGTTGAGTTCTCG-3′
CRG RIP4
5′-CAACTCAAACTCAGGTTGGA-3′
5′-ACAGTTTGTCGGTGTCATTG-3′
CRG RIP5
5′-AGTGCAGCTTCATGATCAGC-3′
5′-GTGTTTGTGACATATCTTCG-3′
CRG RIP 6
5′-GAACACGTAGTTACTTTGGG-3′
5′-GCTATCTTCATCTCATCTTC-3′
CRG RIP 7
5′-AAGGAGAGGCAATGCAAATC-3′
5′-TGTTTTCTGGCACACTTAGC-3′
CRG RIP8
5′-CGATACCGACAGAGAGTATC-3′
5′-CTCGGTATCATATCATCATC-3′
CRG RIP9
5′-TGCTTATGATCTCCAATGTC-3′
5′-TTCTCATTCCAGTTTCTTCG-3′

Mutant rescue experiments *in vitro*

The full *CRG* transcript, and mutant *CRG* fragment with RIP4, RIP5, RIP6, RIP7, RIP4-7 deletion, respectively, were amplified from *Drosophila* cDNA and sub-cloned into the expression vector pAc5.1 to generate six recombinant expression plasmids. The primers used to amplify these fragments were:

Full length of *CRG*
5′-CCGCGGGTACCTTTTATTCAGCATTTA GTCA-3′
5′-CGACGGCGGCCGCTTTCAGTGTTTACTCG TTTT-3′
CRG with RIP4 deletion
5′-GGCGCGATATCTTTTATTCAGCATTTA GTCA-3′
5′-GGCGCTCTAGATTTCAGTGTTTACTCG TTTT-3′
5′-GCCAGGTTTCTTCATTGATATGTGTTGTGT ATAC-3′
5′-GTATACACAACACATATCAATGAAGAAAC CTGGC-3′
CRG with RIP5 deletion
5′-GGCGCGATATCTTTTATTCAGCATTTA GTCA-3′
5′-GGCGCTCTAGATTTCAGTGTTTACTCG TTTT-3′
5′- TACTAAGTTCGATCTCCGCCAGGTTTCTTC ATTG-3′
5′-CAATGAAGAAACCTGGCGGAGATCGAACT TAGTA-3′
CRG with RIP6 deletion
5′-GGCGCGATATCTTTTATTCAGCATTTA GTCA-3′

5′-GGCGCTCTAGATTTCAGTGTTTACTCG TTTT-3′
5′-TCCTTAACTTCATCGCTTTGGGACTCATCA ACTA-3′
5′-TAGTTGATGAGTCCCAAAGCGATGAAGTT AAGGA-3′

CRG with RIP7 deletion
5′-GGCGCGATATCTTTTATTCAGCATTTA GTCA-3′
5′-GGCGCTCTAGATTTCAGTGTTTACTCG TTTT-3′
5′-CATTTCATCGATTTTTGGGCTATCTTCATC TCATC-3′
5′-GATGAGATGAAGATAGCCCAAAAATCGAT GAAATG-3′

CRG with RIP4-7 deletion
5′-GGCGCGATATCTTTTATTCAGCATTTA GTCA-3′
5′-GGCGCTCTAGATTTCAGTGTTTACTCG TTTT-3′
5′-CATTTCATCGATTTTTGGATATGTGTTGTG TATAC-3′
5′-GTATACACAACACATATCCAAAAATCGAT GAAATG-3′

Drosophila S2 cells with *CRG* absence were grown at 25°C in Hyclone SFX-insect cell culture medium (Thermo). Cell transfection was performed with cells plated on 6-well plates using a Attractene transfection reagent (Qiagen) according to the manufacturer's instructions. 1.2 μg of the empty pAc5.1 plasmid and the above six recombinant plasmids were transfected into *Drosophila* S2 cell, respectively, after 2-day transfection, the recruitment of RNA polymerase II to the *CASK* potential promoter regions for each group were analysed using CHIP. All experiments were performed in triplicate.

Buridan's paradigm

The locomotor activity of the flies was measured using Buridan's paradigm as described earlier (37). Each group of flies at 1–2 days after eclosion was rapidly anesthetized with carbon dioxide (CO_2) for 5 min, during this time their wings were cut to about half their normal length. The flies were allowed to recover for 1 day before being used for the experiments. Individual flies were placed on an 80 mm diameter circular platform surrounded by a 200 mm diameter LED cylinder and allowed to walk freely. Two strips, each strip covering a 14.4° angle, were arranged on either side of the center of the cylinder. Three-minute traces were recorded with a camera at a rate of 12 Hz and then processed by Limelight (Version 2.0). Trace length was used to evaluate locomotor activity.

Climbing assay

Climbing ability was tested using the negative geotaxis assay described by Coulom and Birman (38). Three to five-day-old flies were used in the behavioral tests. For each fly strain, three groups of adult flies (10 flies per group) were anesthetized and placed in vertical glass tubes (25 cm in length, 1.5 cm in diameter). About 1 h later, when the flies had recovered from CO_2 exposure, they were gently tapped to the bottom of the vertical tube and allowed to climb up the tube wall for 10 s. The test was repeated three times for each group. The climbing score was calculated as the mean percentage of flies that reached a height of 15 cm within 10 s.

Statistical analysis

All experimental data were subjected to one-way ANOVA (SPSS Version 11.5, SPSS Inc.). All graphs represent the means ± SEM.

RESULTS

Identification of the novel lncRNA *CRG*

To study the function of lncRNAs in the nervous system of *Drosophila*, we screened 16 302 *D. melanogaster* ESTs downloaded from Unigene (Build #37, ftp://ftp.ncbi.nih.gov/repository/UniGene/) for lncRNAs that were expressed in the central nervous system (CNS). We excluded the 12 405 ESTs annotated as coding sequence, and assessed the protein-coding capacity of the remaining 3897 ESTs using GenScan (39). We aligned the remaining 2421 ESTs without protein-coding capacity with the *Drosophila* genome using SIM4 (40), and obtained 2275 ESTs with a full-length coverage threshold of ≥80% and an identity threshold within the covered regions of ≥85%. We then analysed 5 kb of genomic sequence flanking both sides of each EST locus using GenScan (39) and Genie (41) and compared the ESTs and their flanking regions against the Swiss-Prot protein sequence database (ftp://us.expasy.org/databases/swiss-prot) using BLAST to identify any remaining sequences with protein-coding capacity. Ultimately, we obtained 377 candidate lncRNAs. After checking the expression information for these candidates in the EST database (http://www.ncbi.nlm.nih.gov/nucest), we identified 107 lncRNA candidates with an average length longer than 500 nt that were expressed in the adult brain.

After further screening using RT-PCR and whole mount *in situ* hybridization, we found one lncRNA *CRG* showed relatively restricted expression in the *Drosophila* CNS from the embryonic to the adult stages and was selected for further functional investigation.

Although the EST for *CRG* was only 650 nt long, 5′ and 3′ RACE showed that *CRG* transcription yielded a 2672-nt, non-spliced and polyadenylated lncRNA that was transcribed in the same direction as the protein-coding *CASK* gene that is immediately upstream of it (Figure 1A and Supplementary Figure S1). ORF analysis of the *CRG* sequence found no ORFs longer than 200 nt, and because there was no match to any protein sequence in the Swiss-Prot database corresponding to an ORF longer than 50 nt, we supposed that *CRG* is probably a ncRNA. Computational method was used to further analyse the protein-coding potential of *CRG* (31,32). The PhyloCSF score for the *CRG* locus was −9060.55 decibans, indicating that *CRG* is 1.137×10^{906} times more likely to be a non-coding transcript than protein-coding. In addition, *in vitro* translation assay

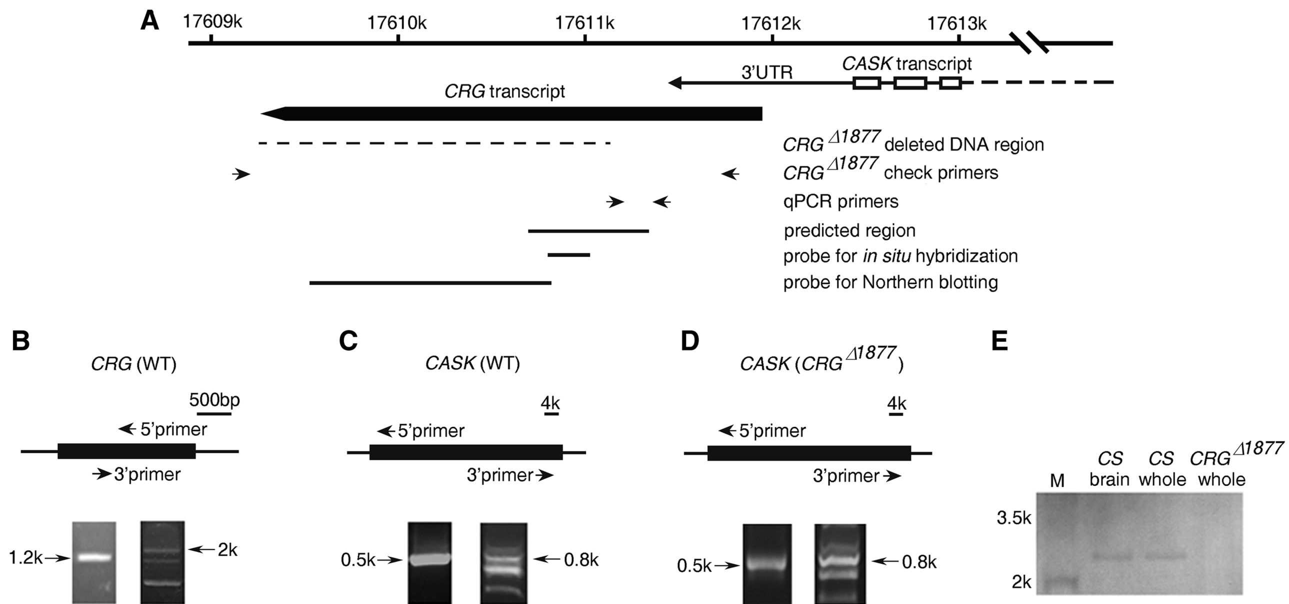

Figure 1. Identification of *CRG* as a lncRNA. (**A**) Schematic representation of the *CRG* locus and transcript. The region of DNA deleted in $CRG^{\Delta1877}$, the primers used to check the null mutant by PCR and qPCR, the predicted region, and the probes used for *in situ* hybridization and northern blotting are shown. (**B**) 5′ and 3′ RACE of *CRG* in WT fly. (**C**) 5′ and 3′ RACE of *CASK* in WT fly. (**D**) 5′ and 3′ RACE of *CASK* in $CRG^{\Delta1877}$ flies. $CRG^{\Delta1877}$, *CRG* null mutant. (**E**) Northern blotting of *CRG* in WT and $CRG^{\Delta1877}$ flies. A ~2.6-kb *CRG* transcript was observed in the adult head and whole fly of the WT, *Canton-S*, but not in the whole fly of the $CRG^{\Delta1877}$ flies. M, RNA marker. $CRG^{\Delta1877}$, *CRG* null mutant.

demonstrated that full *CRG* expression did not produce any protein products which further confirmed its non-protein-coding capacity (Supplementary Figure S2).

Since there is a 464-bp overlap between the 5′ end of *CRG* and the 3′ UTR of *CASK*, we compared 5′ and 3′ RACE products between *CRG* and *CASK* in WT fly (Figure 1B and C). The 5′ and 3′ end sequences of the *CRG* transcript were completely different from those of the *CASK* transcript, suggesting that the *CRG* and *CASK* transcripts are independent. To further confirm the *CRG* and *CASK* were two independent transcripts, we also detected the 5′ and 3′ end sequences of the *CASK* transcript in *CRG* null mutant. We found the *CRG* absence did not affect the full length of *CASK* transcript (Figure 1D), which was the same as that of WT fly (Figure 1C). The independent *CRG* transcript was also supported by northern blotting, by which the size of the *CRG* was consistent with the result of 5′ and 3′ RACE in WT fly (Figure 1E). Thus, we confirmed that *CRG* is indeed a lncRNA.

CRG is a neuro-specific lncRNA

We evaluated the expression of *CRG* in the CNS of flies at different developmental stages by *in situ* hybridization. In adult brain tissues, *CRG* expression was concentrated in the central brain and in the regions between the central brain and the optic lobes (Figure 2). At the third instar larval stage, the *CRG* expression level in central brain lobes was higher than that in the optic lobe or the ventral nerve cord (Figure 2). *CRG* showed a dynamic expression pattern in embryos. It first appeared in the CNS at embryonic stage 14, and was increasingly expressed during embryonic development (Figure 2 and Supplementary Figure S3). In addition, using quantitative RT-PCR, we quantified the distribution of *CRG* in the head, thorax and abdomen of adult fly. The results showed that *CRG* expressed highly in adult head, quite low in thorax part and nearly absent in abdomen segments of adult fly (Supplementary Figure S4). These results indicate that *CRG* is a CNS-specific lncRNA, and strongly suggest that *CRG* plays a role in the CNS.

CRG is required for locomotor behavior in *Drosophila*

To explore the function of this lncRNA further, we constructed a *CRG* deficiency line using a gene-targeting approach. Because highly conserved regions may be important functional units, we analysed *CRG* genomic sequence conservation in 12 *Drosophila* species and honeybee, mosquito and beetle (Supplementary Figure S5) and identified a 1877-bp fragment that was highly conserved across the 12 *Drosophila* species. We deleted this fragment to generate the *CRG* deficiency line $CRG^{\Delta1877}$ (Figure 1A).

The deficiency was first confirmed by PCR using primers flanking the deficiency region and subsequent sequencing (Supplementary Figure S6). At the RNA level, no *CRG* band was observed in $CRG^{\Delta1877}$ by northern blotting, and no expression of *CRG* was detected in $CRG^{\Delta1877}$ by quantitative RT-PCR (Figures 1E and 3A). Moreover, *in situ* hybridization revealed that *CRG* expression was absent in $CRG^{\Delta1877}$ at various developmental stages from the embryo to the adult (Figure 2 and Supplementary Figure S3). Taken together, these results demonstrated that $CRG^{\Delta1877}$ was a *CRG*-null mutant line.

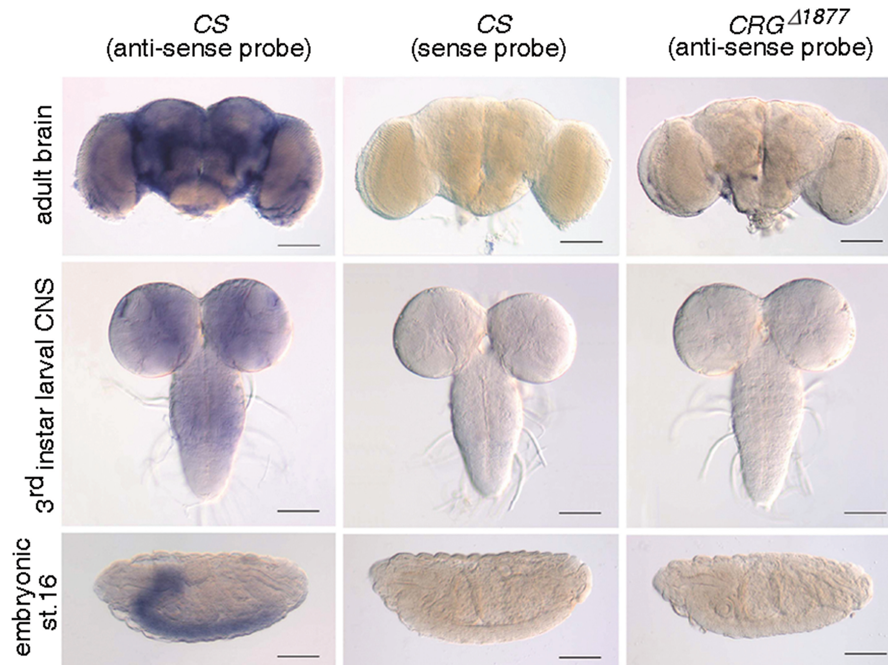

Figure 2. Expression of *CRG* by *in situ* hybridization. The left column shows the expression of *CRG* in the CNS of adults, third instar larvae, and stage 16 embryos in WT *CS* flies, detected with the *CRG* anti-sense probe. No signal was detected with the *CRG* sense probe in *CS* flies (middle column) or with the *CRG* anti-sense probe in *CRG*$^{\Delta 1877}$ flies (right column). Scale bars, 100 µm.

The *CRG*$^{\Delta 1877}$ fly was viable and fertile, and had no visible defects. However, the mutant fly was defective in locomotor activity and climbing ability when tested with the Buridan's paradigm and a negative geotaxis assay, respectively (Figure 3B and C). To determine whether the defects were caused by *CRG* deficiency, we performed rescue experiments by over-expressing *CRG* in the *CRG*$^{\Delta 1877}$ background. The *CRG* level was fully restored by *elav*-Gal4-driven over-expression (Figure 3A), and the locomotor activity defect in *CRG*$^{\Delta 1877}$ was fully rescued by *elav*-Gal4-driven over-expression (Figure 3B). The climbing ability defect in *CRG*$^{\Delta 1877}$ was also restored by *CRG* over-expression driven by *elav*-Gal4 (Figure 3C). In addition, we chose three specific Gal4 lines to narrow down the regions where *CRG* are required for function. The results showed that the behavioral defects of *CRG* null mutant flies in the two behavior paradigms could not be rescued by *CRG* over-expression driven by muscle-specific Gal4 (G7-Gal4), motoneuron-specific Gal4 (OK6-Gal4) and cholinergic neuron-specific Gal4 (6793-Gal4), respectively (Supplementary Figure S7), which suggested that the *CRG* may function in the central brain, but not in the motor neurons and muscles.

CRG regulates *CASK* participating in locomotor behavior

Because *CRG* is localized immediately downstream of *CASK*, which is a movement disorder-related protein-coding gene that is implicated in the walking behavior of *Drosophila* (27,28) (Supplementary Figure S8), we assumed an interaction between *CRG* and *CASK*. First, we examined *CASK* RNA levels in the adult head of *CRG*$^{\Delta 1877}$ flies by quantitative RT-PCR. *CASK* RNA levels were significantly decreased in the *CRG* deficiency line compared with its levels in WT flies, and could be rescued by pan-neuronal expression of *CRG* (Figure 3D). Next, we examined *CASK* protein levels in *CRG*$^{\Delta 1877}$ flies by western blotting. Compared with those in the WT control, *CASK* protein levels were reduced in the adult head of the *CRG* deficiency line. This defect could be rescued by *CRG* over-expression driven by *elav*-Gal4 (Figure 3E). This suggests that *CRG* may act as a regulator of *CASK* expression. In fact, we named the lncRNA *CRG* because of this likely interaction.

To determine whether *CASK* down-regulation evoked by *CRG* deletion was responsible for the aberrant behavior of *CRG*$^{\Delta 1877}$ flies, we attempted to rescue the defective phenotypes in the *CRG* mutant by over-expressing *CASK*. We found that over-expression of *CASK* in the nervous system of *CRG*$^{\Delta 1877}$ flies could effectively rescue both the locomotor activity and the climbing ability defects, as revealed by Buridan's paradigm and the negative geotaxis assay, respectively (Figure 3F and G). Thus, *CASK* is a relatively direct effector that mediates *CRG*-regulated fly locomotor activity and climbing ability.

CRG positively regulates *CASK* expression

Because lncRNAs have been reported to be important for regulating transcription of their neighboring protein-coding genes, we performed luciferase assays to determine whether *CRG* transcripts affected the *CASK* promoter regions to regulate *CASK* transcription. First, because the longer *CASK* transcript is thought to be involved in locomotor behavior (28), we chose segments 2 kb

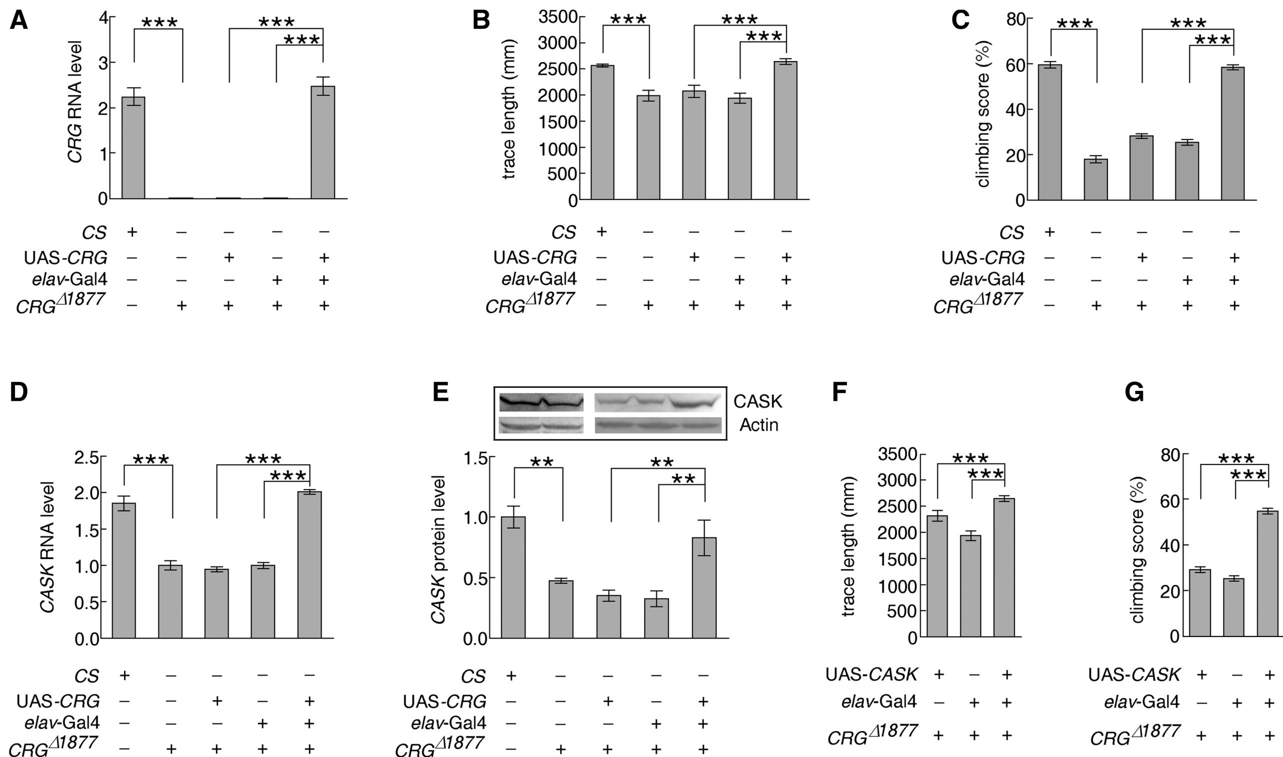

Figure 3. *CRG* affects *CASK*-dependent locomotor activity and climbing ability. (**A**) The expression level of *CRG* measured by qPCR. *CRG* transcription was abolished in $CRG^{\Delta 1877}$ but was fully rescued by pan-neuronal expression of *CRG* driven by *elav*-Gal4. (**B**) In Buridan's paradigm, the 3-min traces of $CRG^{\Delta 1877}$ mutant flies were significantly shorter than those of WT *CS* flies, and were restored by the over-expression of *CRG* driven by *elav*-Gal4. $N = 45$ for each group. (**C**) Climbing ability in the geotaxis assay was reduced in $CRG^{\Delta 1877}$ compared with *CS* flies and rescued by restoration of *CRG* driven by *elav*-Gal4. $N = 9$ for each group. (**D**) The level of *CASK* RNA was reduced in $CRG^{\Delta 1877}$ adult head and was rescued by *CRG* over-expression driven by *elav*-Gal4 as determined by qPCR. (**E**) Levels of *CASK* protein decreased in $CRG^{\Delta 1877}$ adult head and were rescued by *CRG* over-expression driven by *elav*-Gal4 as determined by western blotting. (**F**) The locomotor activity defect in $CRG^{\Delta 1877}$ was restored by *CASK* expression driven by *elav*-Gal4 as determined by Buridan's paradigm. $N = 45$ for each group. (**G**) The defective climbing ability in $CRG^{\Delta 1877}$ flies was rescued by *CASK* over-expression driven by *elav*-Gal4. $N = 9$ for each group. $^{**}P < 0.01$; $^{***}P < 0.001$. Error bars indicate the SEM.

upstream (2 kb-up) and 2 kb downstream (2 kb-down) of the start site of the longer transcript. We tested the two sequences in the luciferase reporter system and confirmed them to be *CASK* promoters (Figure 4A and B). Then, we selected six fragments from the candidate promoter regions: three upstream of the *CASK* transcription start site (2 kb-up, 1 kb-up1, and 1 kb-up2) and three downstream of the *CASK* transcription start site (2 kb-down, 1 kb-down1, 1 kb-down2) (Figure 4C and D). In a luciferase-based reporter system, our results showed that all six fragments functioned as promoters to induce luciferase expression, moreover, the *CRG* transcript substantially enhanced *CASK* 2 kb-up and 2 kb-down promoter-induced luciferase expression, but had no effect on the luciferase expression induced by any of the other fragments (Figure 4C and D). Furthermore, similar assay for the other four fragments surrounding the *CASK* transcription start site, 0.5 kb-up, 0.2 kb-up and 0.5 kb-down, 0.2 kb-down was also performed (Supplementary Figure S9). Among these fragments, all the fragments except for 0.2 kb-down functioned as promoters to induce luciferase expression, and *CRG* transcript also had no effect on the luciferase expression induced by the four fragments (Supplementary Figure S9). Taken together, these results suggested that the two 2 kb-up and 2 kb-down promoter regions were the *CRG*-targeted segments.

CRG recruits RNA polymerase II to the *CASK* promoters

To determine how *CRG* enhances the *CASK* promoters, we tested whether *CRG* affected the association between the transcription initiation complex and the *CASK* promoter regions. Because RNA polymerase II (Pol (II)) is the central component of the transcription initiation complex, we examined the effect of *CRG* transcripts on the interaction between Pol (II) and the *CASK* promoter region. We selected 14 regions to examine the interaction influenced by *CRG*: 7 within the upstream *CRG*-targeted segment (CHIP up1 to CHIP up7) and 7 within the downstream *CRG*-targeted segment (CHIP down1 to CHIP down7) (Figure 5A). Using CHIP assays, we found that the occupancy of Pol (II) on the two *CASK* promoter regions in CHIP up3 and CHIP down4 was decreased in the *CRG*-null mutant, and the decrement could be fully restored by *CRG* over-expression driven by pan-neuronal *elav*-Gal4 (Figure 5B and C). Thus, the *CRG* transcript regulates *CASK* expression by enhancing the association

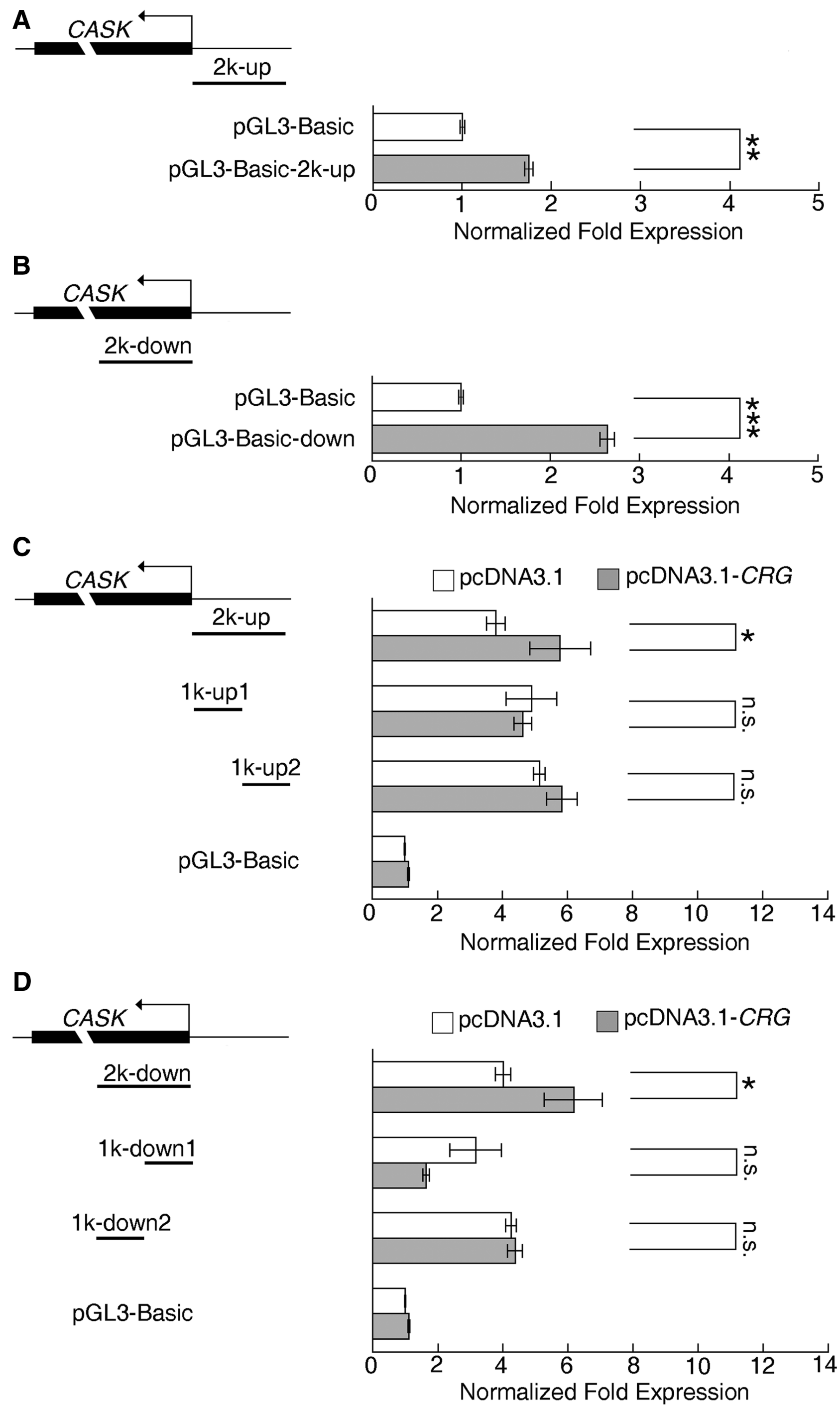

Figure 4. *Trans*-acting mechanism involved in *CRG*-mediated regulation of *CASK* expression. (**A** and **B**) Introduction of the *CASK* 2 kb-up and 2 kb-down fragments into the pGL3-Basic luciferase reporter plasmid-induced luciferase expression, suggesting that they contain *CASK* promoter regions. (**C** and **D**) Additional expression of the *CRG* transcript in pcDNA3.1 promoted the *CASK* 2 kb-up- and 2 kb-down-induced luciferase expression, but had no effect on the luciferase expression of the empty pGL3-Basic, 1 kb-up1, 1 kb-up2, 1 kb-down1, 1 kb-down2 constructs. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$; and n.s., not significant. Error bars indicate the SEM.

between the transcription initiation complex and the *CASK* promoter regions.

Next, we examined whether the *CRG* transcript exerts its function by participating in the transcription initiation complex, and if so, which regions of the *CRG* transcript are required. We performed RIP assays to identify the candidate *CRG* fragments that interacted with Pol (II). Four consecutive truncated *CRG* fragments (RIP4 to RIP7) that jointly span a 1148-nt segment in the center of *CRG* were amplified from immunoprecipitate of Pol (II) extracted from WT tissues (Figure 6A–E). The next question is whether these four truncated *CRG* fragments

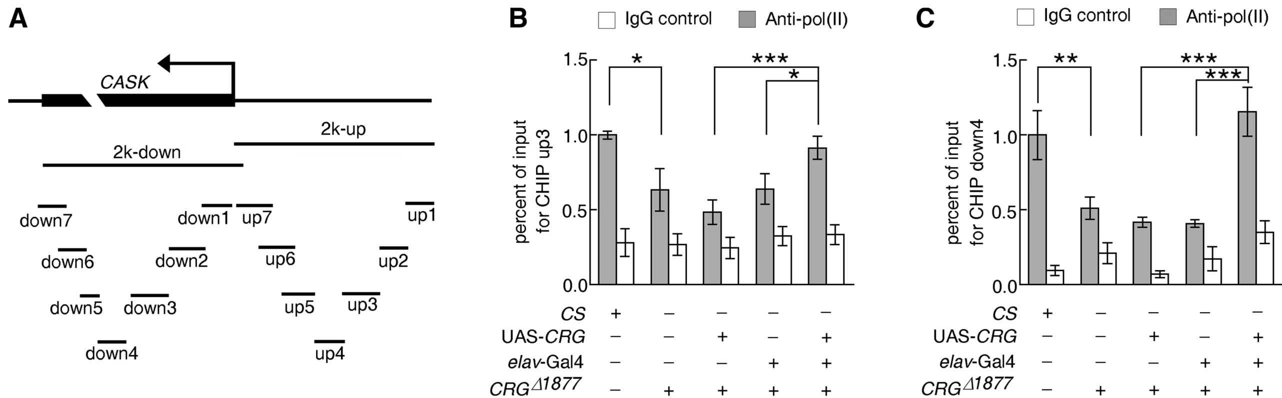

Figure 5. *CRG* transcript enhances the recruitment of Pol (II) to the *CASK* promoter regions. (**A**) Schematic representation of the *CASK* gene showing the transcription start site (arrow). The detected occupancy sites CHIP up1-7 and CHIP down1-7 of the *CASK* promoter regions are shown. (**B** and **C**) *In vivo* occupancy of Pol (II) on the *CASK* promoters was down-regulated in $CRG^{\Delta 1877}$ flies, but was restored by over-expression of the *CRG* transcript, indicating that such occupancy is positively regulated by the *CRG* transcript. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Error bars indicate the SEM.

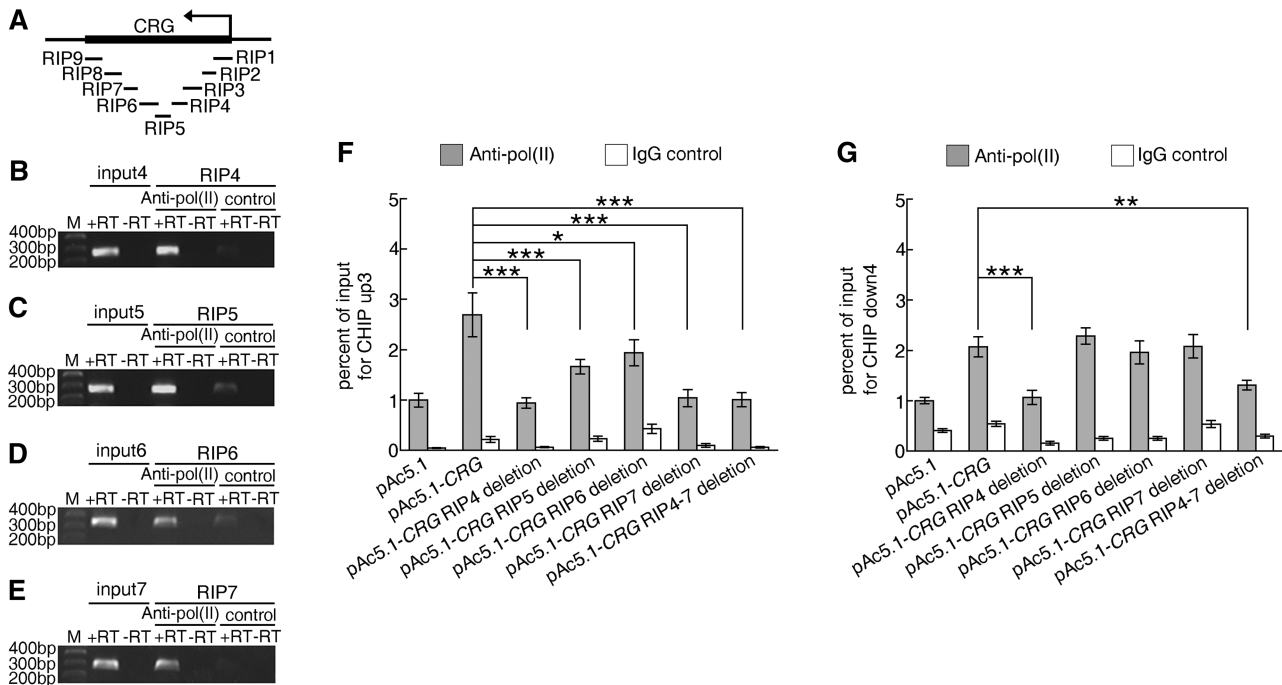

Figure 6. The *CRG* regions important for the recruitment of Pol (II) to the *CASK* promoter regions were identified. (**A**) Schematic representation of the *CRG* gene showing the transcription start site (arrow). All the detected truncated co-precipitated *CRG* fragments were shown. (**B–E**) *In vivo* identification of the *CRG* transcript fragments associated with Pol (II). RIP was performed with an anti-Pol (II) antibody; normal mouse IgG was used as the negative control. M, DNA marker. +RT, reverse transcription. −RT, control for DNA contamination. (**F** and **G**) *In vitro* determination of *CRG* regions important for the Pol (II) recruitment to the *CASK* promoter regions. For the occupancy site CHIP up3, all the *CRG* RIP4, 5, 6, 7 were involved in the recruitment of Pol (II), whereas for CHIP dow4, only *CRG* RIP4 participated in the process. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Error bars indicate the SEM.

are indeed functionally important for recruiting Pol (II) and regulate *CASK* expression. To answer this question, we performed rescue experiments in *Drosophila* S2 cell by mutant *CRG* with the RIP4, 5, 6, 7 fragments deletion, respectively. Using CHIP method, we found that for *CASK* CHIP up3, all the *CRG* RIP4, 5, 6, 7 fragments were involved in the recruitment of Pol (II), whereas for *CASK* CHIP down4, only *CRG* RIP4 played role in the process (Figure 6F and G). Taken together, these results suggested that the *CRG* transcript regulated *CASK* expression by recruiting Pol (II) to the *CASK* promoter regions.

DISCUSSION

Here, we have described a novel neural-specific lncRNA, *CRG*. *CRG* has important biological significance—it is involved in locomotor behavior in *Drosophila*—which is

attributed to its regulation of the transcription of the adjacent protein-coding gene *CASK*.

Although the functions of the most lncRNAs are still unknown, many are expressed in nervous system in *Drosophila* embryos (21) or in mouse (17,18). Our study traced the expression patterns of *CRG* from embryonic and larval stages to adult. The spatiotemporal expression pattern of *CRG* revealed that *CRG* was confined to specific regions of nervous system, which suggested that it could play important roles in neural functions of *Drosophila*. The next question is whether the functional site of *CRG* in CNS or peripherally at the neuromuscular junction (NMJ). To solve this question, electrophysiological experiments can be performed to analyse the excitatory junction potentials (EJPs) in third instar larval muscle in WT comparing with *CRG* nulls. If no differences were observed then it would suggest a central defect. Here, we performed the behavioral rescue experiments which *CRG* was over-expressed driven by motoneuron-specific OK6-Gal4 line or muscle-specific G7-Gal4 line, respectively, in the *CRG* null mutant background. The results showed that defective phenotype of *CRG*-deficiency line could be rescued by the pan-neuronal *CRG* restoration, not by the peripheral *CRG* restoration. So *CRG* was suggested to play central effects in this study.

Most lncRNAs were located in intergenic or intragenic/intronic configurations with protein-coding genes (6,8,14). We validated the full-length transcript of *CRG* using 5′ and 3′ RACE and northern blotting. To check the phylogenetic conservation of *CRG*, we annotated the genomes of 12 *Drosophila*-related species. Strong conservation among all the annotated genomes of the *Drosophila* species suggested *CRG* may be a functional lncRNA. In addition, *CRG* partially overlaps with the 3′ UTR of the adjacent upstream protein-coding gene *CASK*, but *CRG* does not belong to 3′ UTR-associated RNAs which are contiguous with the upstream protein-coding region in the same mRNA (42). First, *CRG* starts in but extends out of the 3′ UTR of *CASK*. Second, *CRG* and *CASK* were transcribed separately. Third, the reduced *CASK* expression in $CRG^{\Delta1877}$ mutant could be rescued by over-expression of WT *CRG*. Thus, *CRG* is independent of *CASK* and not a *CASK* 3′-UTR-derived ncRNA. As for mammal species, such as *Homo sapiens*, *Mus musculus* or *Cattle genus*, despite also host the gene *CASK*, as homologues of *CASK* in *D. melanogaster*, they share poor sequence similarity. Moreover, no EST sequence similar to *CRG* was detected in Database of Expressed Sequence Tags (dbEST) in NCBI using blast. Thus, no similar phenomenon is detected in mammal species.

CASK belonging to a conservative protein family from *Caenorhabditis elegans*, *Drosophila* to mammal, is wildly distributed in nervous system (23,24). As a scaffolding protein, *CASK* interacts with other proteins suggesting the diversity roles of *CASK* in neural activity, development and neurological disease (43). In *Drosophila*, *CASK* which is involved in locomotor behavior was considered as a susceptibility gene for movement disorder (27,28). Both homozygous and heterozygous mutant of *CASK* caused behavior defects in Buridan's paradigm (27) (Supplementary Figure S8). From our observations, *CRG* mutants and *CASK* mutants share similar locomotion defects. *CASK* over-expression rescued the defective phenotype in the *CRG*-deficiency line, while *CRG* positively regulate *CASK*'s expression. Thus, it is highly possible that *CRG* is involved in locomotor behavior by regulating *CASK* expression, which provides a new insight into the pathogenesis of neurological diseases associated with movement disorders.

Increasing evidences suggested that lncRNA could display the biological function by mediating the nearby protein-coding genes in transcriptional regulation and epigenetic gene regulation (6,11,14). In our study, we found that *CRG* promoted the recruitment of RNA Pol (II) to the *CASK* promoter regions to enhance *CASK* transcription, and the detailed *CRG* functional regions responsible for the process were identified, but there are still several questions left open. For example, the details about the transcription initiation complex whether *CRG* recruit RNA Pol (II) by direct interaction, whether *CRG* is involved in epigenetic regulation, are all not clear. Further experiments are needed to explore the regulation mechanism of *CRG*. Interestingly, a previous study reported that a lncRNA could promote the dissociation of the transcription initiation complex from the neighboring *DHFR* promoter through the formation of a complex between the lncRNA, the *DHFR* promoter, and transcription factor IIB (44). The different mechanisms through which ncRNAs exert their regulatory effects highlight their functional diversity and complexity.

To determine whether the expression of any other genes is potentially regulated by *CRG*, we made a whole-body comparison of genome-wide expression profile between $CRG^{\Delta1877}$ mutants versus *CRG* WT flies using microarrays and identified 491 genes down-regulated and 329 genes up-regulated significantly in response to *CRG* mutants (Supplementary Table S1). This has important implications for *CRG* which has wide-ranging effects on gene expression in *Drosophila*. However, *CASK,* which was experimentally identified to be essential target gene of *CRG* regulating in *Drosophila* CNS in our study, was not included in the down-regulated gene list (Supplementary Table S2). The reason might be, in part, owing to the relative low amount of neural-specific *CASK* in the whole body of the two flies. It is worth noting *tsl*, about 1 kb downstream of *CRG*, was significantly down regulated in $CRG^{\Delta1877}$ mutants. However, $CRG^{\Delta1877}$ did not show similar phenotypes as *tsl* mutants (45–47). It indicated that *CRG* might regulate *tsl* in response to other unknown function. The transcriptomic information of *CRG* regulating leads to new questions about the nature of *CRG* for further study in *Drosophila*.

In conclusion, our study suggests that the lncRNA *CRG* recruits Pol (II) to *CASK* promoter regions, which in turn promotes *CASK* expression and thereby influences *Drosophila* locomotor behavior. Further studies are needed to elucidate the detailed molecular mechanisms of *CRG*-mediated regulation of *CASK* expression.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1–2 and Supplementary Figures 1–10.

ACKNOWLEDGEMENTS

We thank Dr Jean-Rene Martin for providing the UAS-*CASK* flies, Dr Yongqing Zhang for providing the G7 and OK6-Gal4 lines, Dr Wei Xie for providing the anti-*CASK* antibody, and Hong Luo and Ying Liu for assistance with *in situ* hybridization experiments. We acknowledge Haiyun Gong, Tengfei Xiao and Xudong Zhao of the IBP core facility centre for technical support. We also thank Joy Fleming and Geir Skogerbø for valuable comments on the manuscript.

FUNDING

Ministry of Science and Technology of China [2009CB918702, 2009CB825400 and 2012CB825504]; National Natural Sciences Foundation of China [31030037, 30921064 and 31070944]; Chinese Academy of Sciences [GJHZ1005]. Funding for open access charge: National Natural Sciences Foundation of China.

Conflict of interest statement. None declared.

REFERENCES

1. Birney,E., Stamatoyannopoulos,J.A., Dutta,A., Guigó,R., Gingeras,T.R., Margulies,E.H., Weng,Z., Snyder,M., Dermitzakis,E.T., Thurman,R.E. *et al*. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **447**, 799–816.
2. Ravasi,T., Suzuki,H., Pang,K.C., Katayama,S., Furuno,M., Okunishi,R., Fukuda,S., Ru,K., Frith,M.C., Gongora,M.M. *et al*. (2006) Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res*., **16**, 11–19.
3. Guttman,M., Amit,I., Garber,M., French,C., Lin,M.F., Feldser,D., Huarte,M., Zuk,O., Carey,B.W., Cassady,J.P. *et al*. (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, **458**, 223–227.
4. Graveley,B.R., Brooks,A.N., Carlson,J.W., Duff,M.O., Landolin,J.M., Yang,L., Artieri,C.G., van Baren,M.J., Boley,N., Booth,B.W. *et al*. (2011) The developmental transcriptome of *Drosophila melanogaster*. *Nature*, **471**, 473–479.
5. Feng,J., Bi,C., Clark,B.S., Mady,R., Shah,P. and Kohtz,J.D. (2006) The *Evf-2* noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev*., **20**, 1470–1484.
6. Huarte,M., Guttman,M., Feldser,D., Garber,M., Koziol,M.J., Kenzelmann-Broz,D., Khalil,A.M., Zuk,O., Amit,I., Rabani,M. *et al*. (2010) A large intergenic non-coding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, **142**, 409–419.
7. Ng,S.Y., Johnson,R. and Stanton,L.W. (2012) Human long non-coding RNAs promote pluripotency and neuronal differentiation by association with chromatin modifiers and transcription factors. *EMBO J*., **31**, 522–533.
8. Ørom,U.A., Derrien,T., Beringer,M., Gumireddy,K., Gardini,A., Bussotti,G., Lai,F., Zytnicki,M., Notredame,C., Huang,Q. *et al*. (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell*, **143**, 46–58.
9. Lipshitz,H.D., Peattie,D.A. and Hogness,D.S. (1987) Novel transcripts from the *Ultrabithorax* domain of the bithorax complex. *Genes Dev*., **1**, 307–322.
10. Petruk,S., Sedkov,Y., Riley,K.M., Hodgson,J., Schweisguth,F., Hirose,S., Jaynes,J.B., Brock,H.W. and Mazo,A. (2006) Transcription of *bxd* noncoding RNAs promoted by trithorax represses *Ubx* in *cis* by transcriptional interference. *Cell*, **127**, 1209–1221.
11. Khalil,A.M., Guttman,M., Huarte,M., Garber,M., Raj,A., Rivea Morales,D., Thomas,K., Presser,A., Bernstein,B.E., van Oudenaarden,A. *et al*. (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc. Natl Acad. Sci. USA*, **106**, 11667–11672.
12. Gupta,R.A., Shah,N., Wang,K.C., Kim,J., Horlings,H.M., Wong,D.J., Tsai,M.C., Hung,T., Argani,P., Rinn,J.L. *et al*. (2010) Long non-coding RNA *HOTAIR* reprograms chromatin state to promote cancer metastasis. *Nature*, **464**, 1071–1076.
13. Yu,W., Gius,D., Onyango,P., Muldoon-Jacobs,K., Karp,J., Feinberg,A.P. and Cui,H. (2008) Epigenetic silencing of tumour suppressor gene *p15* by its antisense RNA. *Nature*, **451**, 202–206.
14. Wang,K.C., Yang,Y.W., Liu,B., Sanyal,A., Corces-Zimmerman,R., Chen,Y., Lajoie,B.R., Protacio,A., Flynn,R.A., Gupta,R.A. *et al*. (2011) A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature*, **472**, 120–124.
15. Ilik,I. and Akhtar,A. (2009) *roX* RNAs: non-coding regulators of the male X chromosome in flies. *RNA Biol*., **6**, 113–121.
16. Fejes-Toth,K., Sotirova,V., Sachidanandam,R., Assaf,G., Hannon,G.J., Kapranov,P., Foissac,S., Willingham,A.T., Duttagupta,R., Dumais,E. *et al*. (2009) Post-transcriptional processing generates a diversity of 5′-modified long and short RNAs. *Nature*, **457**, 1028–1032.
17. Mercer,T.R., Dinger,M.E., Sunkin,S.M., Mehler,M.F. and Mattick,J.S. (2008) Specific expression of long noncoding RNAs in the mouse brain. *Proc. Natl Acad. Sci. USA*, **105**, 716–721.
18. Ponjavic,J., Oliver,P.L., Lunter,G. and Ponting,C.P. (2009) Genomic and transcriptional co-localization of protein-coding and long non-coding RNA pairs in the developing brain. *PLoS Genet*., **5**, e1000617.
19. Soshnev,A.A., Ishimoto,H., McAllister,B.F., Li,X., Wehling,M.D., Kitamoto,T. and Geyer,P.K. (2011) A conserved long noncoding RNA affects sleep behavior in *Drosophila*. *Genetics*, **189**, 455–468.
20. Qureshi,I.A. and Mehler,M.F. (2011) Non-coding RNA networks underlying cognitive disorders across the lifespan. *Trends Mol. Med*., **17**, 337–346.
21. Inagaki,S., Numata,K., Kondo,T., Tomita,M., Yasuda,K., Kanai,A. and Kageyama,Y. (2005) Identification and expression analysis of putative mRNA-like non-coding RNA in *Drosophila*. *Genes Cells*, **10**, 1163–1173.
22. Hardiman,K.E., Brewster,R., Khan,S.M., Deo,M. and Bodmer,R. (2002) The *bereft* gene, a potential target of the neural selector gene *cut*, contributes to bristle morphogenesis. *Genetics*, **161**, 231–247.
23. Hata,Y., Butz,S. and Südhof,T.C. (1996) CASK: a novel *dlg*/PSD95 homolog with an N-terminal calmodulin-dependent protein kinase domain identified by interaction with neurexins. *J. Neurosci*., **16**, 2488–2494.
24. Zordan,M.A., Massironi,M., Ducato,M.G., Te Kronnie,G., Costa,R., Reggiani,C., Chagneau,C., Martin,J.R. and Megighian,A. (2005) *Drosophila* CAKI/CMG protein, a homolog of human CASK, is essential for regulation of neurotransmitter vesicle release. *J. Neurophysiol*., **94**, 1074–1083.
25. Chen,K. and Featherstone,D.E. (2011) Pre and postsynaptic roles for *Drosophila* CASK. *Mol. Cell. Neurosci*., **48**, 171–182.
26. Lopes,C., Gassanova,S., Delabar,J.M. and Rachidi,M. (2001) The CASK/Lin-2 *Drosophila* homologue, *Camguk*, could play a role in epithelial patterning and in neuronal targeting. *Biochem. Biochem. Biophys. Res. Commun*., **284**, 1004–1010.
27. Martin,J.R. and Ollo,R. (1996) A new *Drosophila* Ca^{2+}/calmodulin-dependent protein kinase (Caki) is localized in the central nervous system and implicated in walking speed. *EMBO J*., **15**, 1865–1876.

28. Slawson,J.B., Kuklin,E.A., Ejima,A., Mukherjee,K., Ostrovsky,L. and Griffith,L.C. (2011) Central regulation of locomotor behavior of *Drosophila melanogaster* depends on a CASK isoform containing CaMK-like and L27 domains. *Genetics*, **187**, 171–184.
29. Sun,M., Liu,L., Zeng,X., Xu,M., Liu,L., Fang,M. and Xie,W. (2009) Genetic interaction between Neurexin and CAKI/CMG is important for synaptic function in *Drosophila* neuromuscular junction. *Neurosci. Res.*, **64**, 362–371.
30. Guo,A., Li,L., Xia,S.Z., Feng,C.H., Wolf,R. and Heisenberg,M. (1996) Conditioned visual flight orientation in *Drosophila*: dependence on age, practice, and diet. *Learn. Mem.*, **3**, 49–59.
31. Lin,M.F., Jungreis,I. and Kellis,M. (2011) PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics*, **27**, i275–i282.
32. Cabili,M.N., Trapnell,C., Goff,L., Koziol,M., Tazon-Vega,B., Regev,A. and Rinn,J.L. (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.*, **25**, 1915–1927.
33. Rong,Y.S., Titen,S.W., Xie,H.B., Golic,M.M., Bastiani,M., Bandyopadhyay,P., Olivera,B.M., Brodsky,M., Rubin,G.M. and Golic,K.G. (2002) Targeted mutagenesis by homologous recombination in *D. melanogaster*. *Genes Dev.*, **16**, 1568–1581.
34. Kosman,D., Mizutani,C.M., Lemons,D., Cox,W.G., McGinnis,W. and Bier,E. (2004) Multiplex detection of RNA expression in *Drosophila* embryos. *Science*, **305**, 846.
35. Tautz,D. and Pfeifle,C. (1989) A non-radioactive *in situ* hybridization method for the localization of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene hunchback. *Chromosoma*, **98**, 81–85.
36. Wang,Z., Pan,Y., Li,W., Jiang,H., Chatzimanolis,L., Chang,J., Gong,Z. and Liu,L. (2008) Visual pattern memory requires *foraging* function in the central complex of *Drosophila*. *Learn. Mem.*, **15**, 133–142.
37. Goetz,K.G. (1980) Visual guidance in *Drosophila*. *Basic Life Sci.*, **16**, 391–407.
38. Coulom,H. and Birman,S. (2004) Chronic exposure to rotenone models sporadic Parkinson's disease in *Drosophila melanogaster*. *J. Neurosci.*, **24**, 10993–10998.
39. Burge,C. and Karlin,S. (1997) Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.*, **268**, 78–94.
40. Florea,L., Hartzell,G., Zhang,Z., Rubin,G.M. and Miller,W. (1998) A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.*, **8**, 967–974.
41. Reese,M.G., Kulp,D., Tammana,H. and Haussler,D. (2000) Genie–gene finding in *Drosophila melanogaster*. *Genome Res.*, **10**, 529–538.
42. Mercer,T.R., Wilhelm,D., Dinger,M.E., Soldà,G., Korbie,D.J., Glazov,E.A., Truong,V., Schwenke,M., Simons,C., Matthaei,K.I. *et al.* (2011) Expression of distinct RNAs from 3′ untranslated regions. *Nucleic Acids Res.*, **39**, 2393–2403.
43. Hsueh,Y.P. (2006) The role of the MAGUK protein CASK in neural development and synaptic function. *Curr. Med. Chem.*, **13**, 1915–1927.
44. Martianov,I., Ramadass,A., Serra Barros,A., Chow,N. and Akoulitchev,A. (2007) Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature*, **445**, 666–670.
45. Stevens,L.M., Frohnhöfer,H.G., Klingler,M. and Nüsslein-Volhard,C. (1990) Localized requirement for *torso-like* expression in follicle cells for development of terminal anlagen of the *Drosophil*a embryo. *Nature*, **346**, 660–663.
46. Savant-Bhonsale,S. and Montell,D.J. (1993) *torso-like* encodes the localized determinant of *Drosophila* terminal pattern formation. *Genes Dev.*, **7**, 2548–2555.
47. Martin,J.R., Raibaud,A. and Ollo,R. (1994) Terminal pattern elements in *Drosophila* embryo induced by the *torso-like* protein. *Nature*, **367**, 741–745.