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

Meixia Li¹, Shengyun Wen¹,², Xiangqian Guo³, Baoyan Bai³, Zhefeng Gong¹, Xiaojun Liu⁴, Yijin Wang¹,², Yanqiong Zhou¹,², Xiaowei Chen²,³, Li Liu¹,⁵,* and Runsheng Chen³,*

¹State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, ²Graduate University of the Chinese Academy of Sciences, Beijing 100039, ³Key Laboratory of Noncoding RNA, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, ⁴National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100005 and ⁵Key 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 (Ca²⁺/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...
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^{A1877} 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^{A1877}

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'-TATTAACGCGTCTTACGAGGACCGGTATGCC T-3',

5'-GGCCAGATCTTACCGAATTTTAAATACATAA G-3'.

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

5'-CCGGCTAGCTGTGTTGTGTATACATATTTTC T-3',

5'-TTAGGCGCGGCGGTGGGATGATGCCGCTGTTTG G-3'.

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

5'-CTAGTAGGGATAACAGGGTAAT-3',

5'-CTAGATTACCTGTATTCCATCATATTTC T-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 females with mosaic eyes were singly crossed with yw;ey-FLP;MKRS/TM2,y^+ males. The yw^+ offspring were selected and crossed with w^{1118},hs-I-Cre, Sh/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 yw^+ 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' - GTCTTTGCACGCATGCATGATTCA-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' - GCTGCAGATCTTTATTACGATTAGTTC-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'-AAGGAGACCAAAACGGAAT-3' and 3' primer 5'-ATAAATGCAGGCTGTGGCT-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'-CAGATGCGACGACAGGTTC-3' and 3' primer 5'-AGTTCACCTTGATGCCGT-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 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'-GCCACTTGGATCTTTATTCG-3'.

The primers used in 3' RACE for CASK were:

- outer primer: 5'-GATGCGCCTCAATTGGT-3',
- inner primer: 5'-CGAATAATGGTAGATTGTGTCTT-3'.

The primers used in 3' RACE for CRG were:

- outer primer: 5'-AGGAATGAGGAGATTGCAGGT-3'
- inner primer: 5'-CGAAAAACGGGCAATAACCGT-3'.

The primers used in 3' RACE for CASK were:

- outer primer: 5'-CAGATGCGACGACAGGTTC-3'
- inner primer: 5'-CCACTACGAGATGACAGT-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'-AGCCTCAATCGCCGTCATC-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, 1:1000) overnight at 4°C. All 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 GTTAGAATAA-3') and one corresponding to GFP (5'-ACTGGGTGCTCAGGTAGTG-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 = $2^{ΔCt}$, 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′-TATCTGGCTGTTTGGGCTTG-3′
- **qPCR-CRG low**: 5′-TAATGCAACGGCCTTTAT-3′
- **qPCR-CASK up**: 5′-GGGCACTAAAGAACCT-3′
- **qPCR-CASK low**: 5′-ACGACATAGGGGCTGA-3′
- **qPCR-actin up**: 5′-CAGGCGGTGCTTTCTCTA-3′
- **qPCR-actin low**: 5′-GGAATGCTAGCCAAATCCGAT-3′

**Luciferase assay**

The potential CASK promoter regions from the CASK transcription start site, 2 kb-up (−208/−1), 1 kb-up1 (−1000/−1), 1 kb-up2 (−208/−1000), 0.5 kb-up (−500/−1), 0.2 kb-up (−200/−1), 0.5 kb-down (−100/−190), 0.2 kb-down (−200/−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 plasmids to generate 10 recombinant reporter plasmids. The primers used to amplify these fragments were:

- **2 kb-up**
  - 5′-CAATAACCGTCACCTGAGTGTCGAGCCTT-3′
  - 5′-CAATACTCGAGTTGAGCGAAGACTGGTTG-3′
  - 1 kb-up1
  - 5′-GGTGAGGTACCGCTTGAGGGAATATC-3′
  - 5′-GGTTCGGCTAGCTTCGCTGAGGTTGTAAT-3′
  - 1 kb-up2
  - 5′-ATATAACCGTCACCTGAGTGTCGAGCCTT-3′
  - 5′-AGGCGTCTAGCTGCTGGGAGACTGGTTG-3′
  - 0.5 kb-up
  - 5′-GGTGAGGTACCGCTTGAGGGAATATC-3′
  - 5′-GGTTCGGCTAGCTTCGCTGAGGTTGTAAT-3′
  - 0.2 kb-up
  - 5′-GGAGAGGTACCTAAGAGCGAAGAAATACTGGTTG-3′
  - 5′-GGTGAGGTACCGCTTGAGGGAATATC-3′
  - 2 kb-down
  - 5′-GGTAAGGTACCGCTTGAGGGAATATC-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% CO₂ 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′-TTCCCTCTCCTGAGTTCATCACC-3′
  - 5′-CCCTCTCCTTATCATTGAGTG-3′
- **CHIP up2**
  - 5′-GGTGGCTGATACAGTGTGGTTG-3′
  - 5′-AGCACAACACACATTTCGAG-3′
CHIP up
5'-GAAGGTTGGACACCTAGTGGG-3'
5'-AACCGTACGGTGCAGTAATCTCT-3'

CHIP up4
5'-GTGTGTTGCTGTGGGCGT-3'
5'-CGGCGGACTCAAATGACTT-3'

CHIP up5
5'-GGGCGGAGAGAAGTAATG-3'
5'-GCTTTGAGCGTTTTCCTTCGG-3'

CHIP up6
5'-ACTTGGTTGGACATCTTG-3'
5'-GGCTATTGGAAGAAGTGTGGG-3'

CHIP up7
5'-CTTTGAGCGTTTTCCTTCGG-3'
5'-GCTTGAGTGAAGTCAAATGG-3'

CHIP down1
5'-CAATCGAGTCTTAATGCGA-3'
5'-GTGGACATCTTG-3'

CHIP down2
5'-TTGGCAATTTGCTGACACT-3'
5'-ACGAAAATGTGGGAACGAGATCGAAATG-3'

CHIP down3
5'-ACTACGAGAAGTCTTTC-3'
5'-AGGCGCTACTTTGTGCTTT-3'

CHIP down4
5'-CGAACAATCCCTTGAGAAG-3'
5'-ATGGGTGTCGACGGAAGAATG-3'

CHIP down5
5'-TCTTTCTCATTGTGCGA-3'
5'-GCCACTTAAAGGTGCGGCT-3'

CHIP down6
5'-GTGTTCTGCCATCAGATG-3'
5'-AGTTGAAAGGCCGCTAATCG-3'

CHIP down7
5'-GGGCGGACTCAAATGACTT-3'
5'-CCGCGGTTACCTTTTATTCAGCATTTA-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'-CAGCATTTAGCTAAAACAGC-3'
5'-GTCACCACTGCTTTCAGGTT-3'

CRG RIP2
5'-GTTCAGTTCTACATTTCGTGG-3'

CRG RIP3
5'-GGAAAAAACATAGGATAGGC-3'
5'-GTCTCTCTTGTGAGTTCCTCG-3'

CRG RIP4
5'-CAACTTAACTCGAGGTGAGA-3'
5'-ACATTTTGGTGTTGTCATTG-3'

CRG RIP5
5'-GTGGATTGGTTCGGAACTTGG-3'
5'-CTGTGTCGTTATGAATGACT-3'

CRG RIP6
5'-GAAACGCTAGTTACTTTGGG-3'
5'-GCTATCTTCTATCTTCTTC-3'

CRG RIP7
5'-AGTGTCAGCTTCTAGATCGAC-3'
5'-GTGTGGTGCATATACTCCTCG-3'

CRG RIP8
5'-CGATACCAGACAGAGATAC-3'
5'-CTCGGTATCATATCATCATCATC-3'

CRG RIP9
5'-TGCTTATGATCTCTCAATGTCC-3'
5'-TTCCTCATTCCAGATTTTCTCG-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'-CCGCGGGATACCTTTTTACATCAGGATTTACGATTA-3'
5'-CGACGCGCCGCGGCTTCTTGTGTTTTACTCGTTTT-3'

CRG with RIP4 deletion
5'-GGCGCGGATATCTTTATTCAGCATTTA-3'
5'-GTGGATTGGTTCGGAACTTGG-3'
5'-CTGTGTCGTTATGAATGACT-3'

CRG with RIP5 deletion
5'-GAAACGCTAGTTACTTTGGG-3'
5'-GCTATCTTCTATCTTCTTC-3'

CRG with RIP6 deletion
5'-GTCACCACTGCTTTCAGGTT-3'
tions. 1.2
reagent (Qiagen) according to the manufacturer's instruc-
plated on 6-well plates using a Attractene transfection
(Thermo). Cell transfection was performed with cells
six recombinant plasmids were transfected into
of flies at 1–2 days after eclosion was rapidly anesthetized
Buridan's paradigm as described earlier (37). Each group
The locomotor activity of the flies was measured using
Buridan's paradigm
CASK. All experiments were performed in triplicate.

CRG with RIP4-7 deletion
5'-GGCCGGATATCTTTTTATTACGATTTA-GTCA-3'
5'-GGCCGGTCTAGATTTCATGTTTACTCG-GTTT-3'
5'-CATTTCCATGATTTGGGTATCTTTATC-TCATC-3'
5'-GATGAGATGAAGATGCCCCAAAATCGAT-GAAATG-3'

CRG with RIP4-7 deletion
5'-GGCCGGATATCTTTTTATTACGATTTA-GTCA-3'
5'-GGCCGGTCTAGATTTCATGTTTACTCG-GTTT-3'
5'-CATTTCCATGATTTGGGTATCTTTATC-TCATC-3'
5'-GATGAGATGAAGATGCCCCAAAATCGAT-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 was 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 (CO2) 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 CO2 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 16302 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 12405 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 decibels, indicating that CRG is 1.137×10906 times more likely to be a non-coding transcript than protein-coding. In addition, in vitro translation assay
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 at various developmental stages from the embryo to the adult (Figure 2 and Supplementary Figure S3). Taken together, these results demonstrated 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,*1877* (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,*1877* by northern blotting, and no expression of CRG was detected in CRG,*1877* by quantitative RT-PCR (Figures 1E and 3A). Moreover, *in situ* hybridization revealed that CRG expression was absent in CRG,*1877* at various developmental stages from the embryo to the adult (Figure 2 and Supplementary Figure S3). Taken together, these results demonstrated that CRG,*1877* was a CRG-null mutant line.

![Figure 1](http://nar.oxfordjournals.org/)

**Figure 1.** Identification of CRG as a lncRNA. (A) Schematic representation of the CRG locus and transcript. The region of DNA deleted in CRG,*1877*; 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,*1877* flies. CRG,*1877*, CRG null mutant. (E) Northern blotting of CRG in WT and CRG,*1877* 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,*1877* flies. M, RNA marker. CRG,*1877*, CRG null mutant.
The CRG\textsuperscript{A1877} 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\textsuperscript{D1877} background. The CRG level was fully restored by elav-Gal4-driven over-expression (Figure 3A), and the locomotor activity defect in CRG\textsuperscript{A1877} was fully rescued by elav-Gal4-driven over-expression (Figure 3B). The climbing ability defect in CRG\textsuperscript{D1877} 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.

**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\textsuperscript{A1877} flies (right column). Scale bars, 100 μm.

**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\textsuperscript{A1877} 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\textsuperscript{A1877} 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\textsuperscript{A1877} 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\textsuperscript{A1877} 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...
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 promoter-induced luciferase expression, but had no effect on the luciferase expression induced by the four fragments (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
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 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.
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 IncRNAs 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 over-expression 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 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 IncRNAs 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 IncRNA. 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 CASK91877 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 mammalian 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 IncRNA 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 IncRNA could promote the dissociation of the transcription initiation complex from the neighboring DHFR promoter through the formation of a complex between the IncRNA, 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 CRG91877 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 CRG91877 mutants. However, CRG91877 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 IncRNA 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 Skogerbo 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


